Dissecting the Aggregate Market Elasticity*

Victor Duarte† Goutham Gopalakrishna‡ Mahyar Kargar§

Jiacui Li¶ Dejanir Silva‖

August 20, 2026

Abstract

We study aggregate market elasticity, the inverse price impact of portfolio flows, in general equilibrium with heterogeneous investors and passive demand. Once interest rates adjust, elasticity is governed by how flows shift the consumption–wealth relation, not by individual demand elasticities. At the efficient allocation, interest-rate movements offset risk-premium responses, making the aggregate claim infinitely elastic. Risk misallocation generates finite price impact, amplified by preference heterogeneity, leverage constraints, and cash-flow leverage. A perturbation method characterizes this mechanism analytically. Quantitatively, the baseline reaches the lower end of empirical estimates, while an extension with belief distortions raises price impact to the upper end.

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

JEL Classification: G12, G11, G21.

*We thank Andy Atkeson, Xavier Gabaix, Valentin Haddad, Tarek Hassan, Tim Johnson, Rohan Kekre, Leonid Kogan, Ralph Koijen, Arvind Krishnamurthy, Moritz Lenel, Hanno Lustig, Tyler Muir, Christian Opp, Stavros Panageas, Larry Schmidt, 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, SED 2026 Annual Meeting, NBER SI Macro, Money and Financial Frictions and Asset Pricing meetings, UIUC, UCLA macro-finance lunch, and University of Zurich for valuable comments and suggestions. We thank Rotman-FinHub for computational resources. 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). This evidence raises a deeper question: what determines the aggregate market's capacity to absorb portfolio flows?

The inelastic markets hypothesis of Gabaix and Koijen (2023) provides a natural starting point. In their framework, investment mandates and delegated asset management make individual investors' demands inelastic. In a tractable setting with a constant interest rate, risky-asset market clearing then maps these individual demand elasticities directly into aggregate price impact.[2] Using a granular instrumental-variable strategy, they estimate that a \$1 flow into stocks raises aggregate market value by about \$5. Their estimate provides the central empirical motivation for our analysis and lies near the upper end of the available estimates of aggregate price impact.

While a constant interest rate provides a clean benchmark, it is natural to allow both the risk premium and the interest rate to adjust when considering flows between stocks and bonds. We show that this fundamentally changes the determination of aggregate market elasticity. Risky-asset market clearing determines the risk premium required for investors to absorb the flow. But the asset price depends on the total discount rate, the sum of the interest rate and the risk premium. Because the asset price affects household wealth and consumption, goods-market clearing determines this total discount rate, and hence the equilibrium price. The aggregate price response is therefore governed by how portfolio flows shift the consumption–wealth relation, not directly by the slopes of individual portfolio-demand curves.

This distinction is the central point of the paper. The familiar economics of risk bearing determines the risk premium required to reallocate risk across investors. But aggregate market elasticity is an endogenous object whose determination resides in the consumption–saving margin and goods-market clearing. Individual demand elasticities can shape the risk-premium adjustment required to clear the asset market without determining the aggregate price response.

Throughout the paper, we define the *aggregate market elasticity* as the inverse of the equilibrium price impact of a flow between bonds and stocks. It measures the market's capacity to absorb a portfolio flow after all prices, including the interest rate, have adjusted. An elastic, deep market

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

[2] In the basic model of Gabaix and Koijen (2023), the risk-free rate is exogenous. In their behavioral general-equilibrium extension, the interest rate remains constant and therefore does not respond to portfolio flows.

absorbs a large flow with little change in aggregate value, whereas an inelastic market requires a large price movement.[3] The aggregate market elasticity is therefore an equilibrium object. It differs from an investor's elasticity of demand for the risky asset, which measures the investor's portfolio response holding other prices fixed.

We first characterize the aggregate market elasticity in a simple two-period model. In this setting, we establish an infinitely elastic benchmark: when portfolio flows leave the consumption–wealth relation unchanged, the aggregate claim on output has zero price impact. Consider an outflow from stocks to bonds. The outflow raises the risk premium, which by itself would lower the value of the risky claim and household wealth. The resulting decline in consumption is inconsistent with goods-market clearing when aggregate output is fixed. The interest rate therefore falls, exactly offsetting the effect of the higher risk premium on valuations. The aggregate price remains unchanged even though every investor has a finite, and potentially small, demand elasticity.

We then embed the mechanism in a dynamic general equilibrium model with heterogeneous investors, passive demand, and leverage constraints. To characterize price impact, we apply a perturbation method building on Kogan and Uppal (2001) and Silva (2026). The method is local in the size of exogenous shocks but global in the endogenous wealth distribution. We refer to this approach as *state-global perturbation*.[4] It yields closed-form expressions that allow us to dissect aggregate market elasticity into its different determinants.

A key lesson of the dynamic model is that aggregate price impact depends on how portfolio flows shift the consumption–wealth relation. This response, in turn, depends on the initial allocation of risk relative to the portfolios investors would choose if allowed to optimize freely. When initial risk exposures coincide with these counterfactual choices, a marginal change in risk exposure has no first-order effect on investors' saving incentives. The consumption–wealth relation therefore remains unchanged, reproducing the infinitely elastic benchmark. When risk is initially misallocated, however, an additional flow changes the aggregate risk-adjusted return on saving and hence the consumption–wealth ratio. Goods-market clearing then requires the valuation of the aggregate claim to adjust. Thus, risk misallocation, rather than low individual demand elasticity, generates finite aggregate market elasticity.

Beyond identifying the source of finite aggregate market elasticity, the dynamic model reveals the forces that amplify price impact. Preference heterogeneity makes the efficient allocation of risk depend on the wealth distribution. In bad times, the efficient allocation calls for passive investors to bear more risk, but their actual exposure may fail to adjust accordingly, compounding the effects of lower active wealth. Leverage constraints further concentrate risk among unconstrained

[3] We use "aggregate elasticity" and "macro elasticity" interchangeably. The inverse-price-impact definition parallels market depth in Kyle (1985), but our object concerns flows between the aggregate stock and bond markets.

[4] For a recent application of this method, see Kargar, Passadore, Silva, and Yang (2023).

investors. As a result, aggregate market elasticity is state dependent, with price impact rising when active risk-bearing capacity is low. Portfolio-flow shocks then generate excess volatility through their interaction with this finite, state-dependent elasticity: neither flows nor inelasticity alone are sufficient.

Our benchmark concerns the aggregate claim to output, whereas the empirical object of interest is the price impact of portfolio flows on equity. To connect the two, we split the aggregate claim into an equity claim and a corporate debt claim. Equity pays dividends that respond more strongly to aggregate shocks than output itself, while corporate debt receives the remaining cash flow. This cash-flow leverage makes the equity risk premium more sensitive to portfolio flows than the risk premium of the aggregate claim. Goods-market clearing generates an interest-rate response that largely offsets the movement in the aggregate claim's risk premium, but the same response cannot fully offset the larger movement in the equity risk premium. As a result, cash-flow leverage generates price impact for equity even at the benchmark at which the aggregate claim is infinitely elastic. Generating quantitatively large equity price impact, however, requires the aggregate claim itself to depart meaningfully from the infinitely elastic benchmark.

We next consider two extensions that clarify the roles of consumption behavior and monetary policy in the mechanism. First, we introduce wealthy hand-to-mouth investors who own risky assets but consume their dividends, so their consumption does not respond to changes in wealth. Under a natural benchmark, our price-impact results are unchanged even when the share of optimizing households becomes arbitrarily small: the weaker direct response of consumption to wealth is exactly offset by stronger amplification through goods-market clearing. Second, we let the central bank set the short-term interest rate in a production economy with sticky prices. An outflow from equities lowers wealth and aggregate demand, putting downward pressure on inflation. To stabilize inflation, the central bank lowers the policy rate along with the natural rate, implementing the same real-rate adjustment as in the baseline model and leaving the price-impact results unchanged.

For the quantitative analysis, we solve the full dynamic model using the neural-network method of Duarte, Duarte, and Silva (2024). We calibrate the model to U.S. Flow of Funds sectoral holdings, aggregate cash-flow volatility, and the magnitude and persistence of quarterly equity flows. Empirical estimates of equity price impact range from approximately 1.5 to 5.9 (Hartzmark and Solomon, 2025; Da, Larrain, Sialm, and Tessada, 2018; Gabaix and Koijen, 2023). The rational model generates a price impact of 2.2, near the lower end of this range, while also producing an equity premium of 5.9%, return volatility of 19.0%, and a broad match to sectoral balance sheets. The general equilibrium offset remains strong for the aggregate endowment claim, whose price impact of 0.10 corresponds to an aggregate elasticity of approximately 10.0. By contrast, cash-flow leverage reduces the elasticity of equity to approximately 0.5.

We next introduce a belief distortion under which non-fundamental portfolio flows change

investors' expectations of future cash-flow growth. We discipline the strength of this response using the estimated response of cash-flow expectations to non-fundamental price movements in Chaudhry (2025). In isolation, the belief distortion generates a price impact of 3.0. Combining it with preference heterogeneity and the leverage constraint raises price impact to 6.2, consistent with the upper end of the empirical estimates. Relative to the homogeneous rational benchmark, rational frictions account for 31% of the total price impact amplification, distorted beliefs for 46%, and their interaction for the remaining 23%. Although rational frictions and distorted beliefs operate through different channels, each makes the consumption–wealth relation respond to portfolio flows. Rational frictions do so by changing risk misallocation, while distorted beliefs operate through subjective expectations of future cash flows. Price impact also varies substantially with the risk-bearing capacity of active investors, and the model generates return predictability in line with the data. Taken together, our results provide a demand-based explanation of excess volatility: modest portfolio flows can generate large and state-dependent price movements, even when the interest rate adjusts endogenously.

Related literature

Our work relates to a growing literature on macro demand elasticity. The closest contribution to our paper is the inelastic markets hypothesis of Gabaix and Koijen (2023). Their granular instrumental-variable estimate of aggregate price impact provides the central empirical motivation for our analysis, and their model shows how delegated investment and fund mandates translate flows into large price movements. Their general equilibrium extension uses a behavioral consumption–saving specification under which the equilibrium interest rate does not respond to portfolio flows. We allow the interest rate to adjust and show that this shifts the determination of aggregate elasticity from individual demand curves to the consumption–wealth relation and goods-market clearing. Johnson (2006) also studies aggregate price impact in a representative-agent Lucas economy and finds finite elasticity in a frictionless benchmark.[5] His experiment changes the aggregate supply of the risky claim, whereas ours redistributes a fixed collection of claims across heterogeneous investors. van der Beck, Bretscher, and Fu (2026) derive a model-agnostic bound that links return volatility, flow volatility, and investor heterogeneity, and show that observed moments in the U.S. equity market imply a large aggregate price impact, providing a complementary motivation for the structural mechanism we develop. This macro elasticity literature is distinct from a larger body of work on micro elasticity, which examines relative price changes across individual stocks (e.g., Harris and Gurel, 1986; Shleifer, 1986; Chang, Hong, and Liskovich, 2015; Schmickler, 2020; Pavlova and Sikorskaya, 2023) and finds much larger elasticities, consistent with stocks being

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

closer substitutes for each other than for bonds.

We also build on demand-system asset pricing as developed by Koijen and Yogo (2019) and extended in work that incorporates cross-asset elasticities (Fuchs, Fukuda, and Neuhann, 2023; Haddad, He, Huebner, Kondor, and Loualiche, 2025a). Our framework distinguishes these investor-level demand elasticities from the aggregate market elasticity determined by goods-market clearing, and extends the analysis to dynamic settings in which persistent flows and the evolution of risk-bearing capacity shape equilibrium price impact (Binsbergen, David, and Opp, 2025; He, Kondor, and Li, 2025; Davis, Kargar, Li, and Silva, 2026).

Our quantitative belief-distortion mechanism is motivated by Chaudhry (2025), who finds that non-fundamental price increases raise analysts' cash-flow expectations. It is related to the mislearning-from-prices mechanism of Bastianello and Fontanier (2025) and to Bastianello (2026), who studies its connection with demand elasticity in a constant-interest-rate setting. We use this mechanism as a quantitatively complementary extension to the rational general-equilibrium forces that constitute the paper's central contribution.

The consumption–saving mechanism connects our analysis to evidence on wealth effects and to heterogeneous-agent macroeconomics. Stock-market wealth has economically significant effects on consumption and real activity (Chodorow-Reich, Nenov, and Simsek, 2021), consistent with the wealth-effect channel at the center of our mechanism. The offset between the risk premium and the risk-free rate echoes Tallarini (2000), in which greater risk aversion raises the price of risk and lowers the risk-free rate while leaving aggregate quantity dynamics nearly unchanged when intertemporal substitution is held fixed. We characterize when an analogous offset makes the aggregate claim infinitely elastic and when portfolio flows break this neutrality. A substantial fraction of households also exhibit wealthy hand-to-mouth behavior despite holding sizable illiquid assets (Kaplan and Violante, 2014), while the marginal propensity to consume out of capital gains is small relative to the response to dividend income (Di Maggio, Kermani, and Majlesi, 2020). Our neutrality result for wealthy hand-to-mouth investors is related to benchmark aggregation results in Werning (2015) and Bilbiie (2008). The monetary-policy extension connects to models in which asset revaluations and wealth effects shape the transmission of monetary policy (Caballero and Simsek, 2020; Kekre, Lenel, and Mainardi, 2025; Caramp and Silva, 2026). It also relates to evidence that policymakers respond to equity-market declines because of their consumption effects (Cieslak and Vissing-Jorgensen, 2021) and to recent work on financial-conditions targeting (Caballero, Caravello, and Simsek, 2024).

Our passive sector builds on Chien, Cole, and Lustig (2011), in which passive traders hold fixed stock–bond portfolios and active traders absorb the residual aggregate risk. Chien, Cole, and Lustig (2012) show how households that rebalance only intermittently shift aggregate risk toward a small active sector in downturns and thereby make risk compensation countercyclical. We share

this risk-concentration mechanism, but study its implications for the equilibrium price impact in a setting exposed to both fundamental and non-fundamental shocks. Related empirical work studies how passive investment affects elasticity and volatility. Haddad, Huebner, and Loualiche (2025b) show that indexing has made demand for individual stocks more inelastic, while Lochstoer and Muir (2026) find that the secular rise of indexing has substantially increased U.S. equity-market volatility. Parker, Schoar, and Sun (2023) find that target-date funds dampen volatility through more contrarian flows. Our model provides a general equilibrium mechanism connecting passive demand and portfolio flows to aggregate price impact and volatility.

More broadly, our paper contributes to the market-macrostructure literature, which studies how passive investors, levered intermediaries, and central banks shape asset prices (Haddad and Muir, 2025). We embed these investors and institutions in a dynamic general equilibrium in which both interest rates and the wealth distribution adjust to portfolio flows. Our paper also relates to preferred-habitat models (Vayanos and Vila, 2021; Kekre et al., 2025; Ray, Droste, and Gorodnichenko, 2024), which emphasize intermediary risk-bearing capacity and market segmentation. We complement this literature by showing how flow-induced wealth effects alter consumption–saving decisions and, through goods-market clearing, the aggregate price response.

The rest of the paper is organized as follows. Section 2 develops a two-period model that delivers the core intuition. Section 3 builds the dynamic heterogeneous-agent economy with passive and active investors. Section 4 characterizes the determinants of aggregate market elasticity using the state-global perturbation method. Section 5 studies wealthy hand-to-mouth investors and monetary policy. Section 6 presents the rational benchmark, the belief-distortion extension, state-dependent price impact, and return predictability. A short conclusion follows.

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-agent 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 riskless asset, which pays one unit of the consumption good in the second period, has price 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 $W_j = (P + Y)Q_{j,-1} > 0$, where $Q_{j,-1}$ denotes the investor's initial position in the risky asset.

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 (log) risk premium.[6]

Consumption is specified as[7]

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

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

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 $p \equiv \log \frac{P}{Y}$ 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

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

[6] Merton's formula corresponds to the special case $\overline{\alpha}_a(\pi) = \frac{\pi}{\gamma_a \sigma^2}$, given risk aversion γ_a and volatility σ.

[7] In this formulation, consumption responds to changes in wealth for both investor types. Appendix B considers an extension with wealthy hand-to-mouth investors whose consumption does not respond to changes in wealth.

Investor j's demand for the risky asset is

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

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

The market clearing condition for the risky asset is therefore

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

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 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. Goods-market clearing requires aggregate consumption to equal output, $\sum_j \omega_j C_j = Y$. Dividing both sides by aggregate cum-dividend wealth, $\sum_j \omega_j W_j = P + Y$, gives $\sum_j x_j \overline{c}_j = Y/(P + Y)$. Combining this condition with Equation (1), we obtain

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

where $r + \pi = \mu - p$ and $Y/(P + Y) = 1/(1 + e^p)$.

Condition (3) depends only on p and therefore determines the price-dividend ratio. Conditional on p, risky-asset market clearing in (2) determines r, and hence the division of the total expected return between the risk-free rate and the risk premium.

2.1 General equilibrium implications of portfolio flows

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

For a small perturbation, $\overline{\alpha}_p = \overline{\alpha}_p^* e^{\hat{\alpha}_p}$, we linearize demand around the initial equilibrium. Defining $\hat{p} \equiv p - p^*$, $\hat{r} \equiv r - r^*$, and $q_j \equiv \log(Q_j/Q_j^*)$, 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* to investor j's demand 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 $f_a = 0$. A portfolio-flow shock is therefore an asymmetric demand shock. Formally, it directly shifts only passive-investor demand; active investors rebalance in equilibrium as a consequence.

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. Therefore, the prices of the risky and risk-free assets must be jointly determined in equilibrium.

Equilibrium. Let $s_j \equiv \omega_j Q_j^*$ 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.[8]

[8] 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 $\frac{\partial \overline{c}_p}{\partial p}$ is not too large.

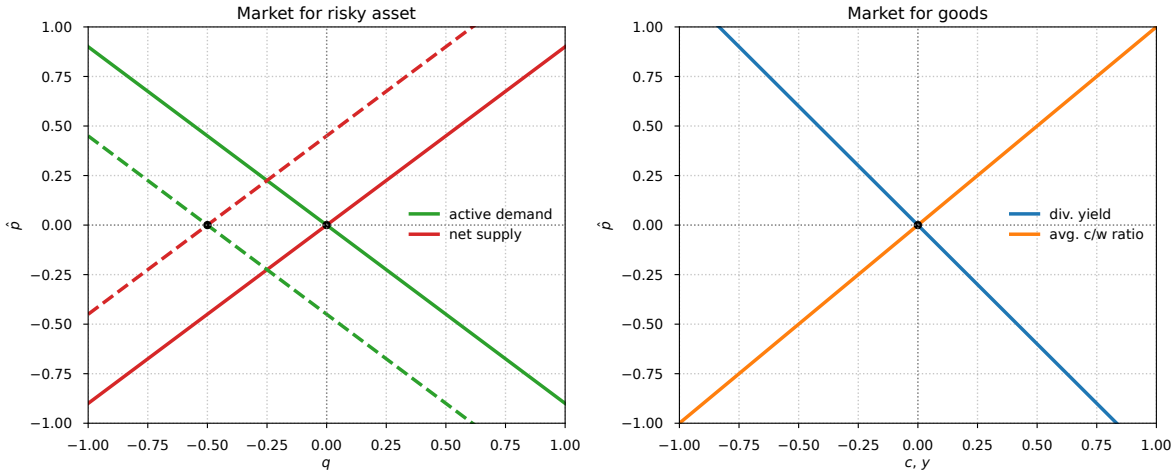


Figure 1. Equilibrium effects of a portfolio-flow shock in the risky-asset market (left) and goods market (right).

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

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

The right panel of Figure 1 plots the linearized changes in the two sides of the goods-market condition in (6): the desired average consumption-wealth ratio, $c \equiv \chi_p \hat{p}$, and the dividend-price ratio, $y \equiv -d_p \hat{p}$. Both are expressed per unit of aggregate cum-dividend wealth. We focus on $\chi_p > 0$, so c is increasing in p, while y is decreasing in p.

The aggregate market elasticity. We are interested in the market's ability to absorb the aggregate portfolio flow $f \equiv s_p f_p + s_a f_a$ between bonds and stocks.[9] We define the *aggregate market elasticity* as the inverse of the price impact generated by this flow:

$$\mathcal{E}_M \equiv \left[\frac{\partial \hat{p}}{\partial f}\right]^{-1}. \tag{7}$$

Hence, an elastic market can absorb a large flow with little price adjustment. If the flow is absorbed without any price adjustment, we say that the market is infinitely elastic. Conversely, in an inelastic market, even a small flow generates a large price movement.[10]

Aggregate market elasticity is a market-level equilibrium outcome. It is distinct from the price elasticities of individual demand curves, which describe investors' portfolio responses holding other prices fixed. It is also distinct from a micro-level elasticity, which concerns substitution and

[9] We focus on the case with no direct shift in active investors' demand, $f_a = 0$, so the aggregate flow is $f = s_p f_p$.

[10] This definition is closely related to the notion of market depth in Kyle (1985) and market liquidity in Johnson (2006).

relative price movements across individual stocks.

The inelastic markets hypothesis. Consider a version of the model in which the interest rate is held fixed at $r = r^*$ and goods-market clearing is ignored. Using the linearized risky-asset market condition (5) with $\hat{r} = 0$, we obtain

$$\hat{p} = \underbrace{\frac{1}{\zeta_p}}_{\text{price impact}} \times \underbrace{f}_{\text{flow shock}},$$

where

$$\zeta_p \equiv s_a \zeta_{a,p} + s_p \zeta_{p,p}, \qquad s_j \equiv \omega_j Q_j^*.$$

In this setting, the aggregate market elasticity equals the average of investors' own-price demand elasticities: $\mathcal{E}_M = \zeta_p$. Intuitively, markets are inelastic when individual demands are inelastic. This is the core implication of the *inelastic markets hypothesis* of Gabaix and Koijen (2023), who estimate large price responses to stock-market flows and study a behavioral general-equilibrium model in which investors have relatively inelastic demand and the interest rate is constant.

Portfolio flows in general equilibrium. With a constant interest rate, risky-asset market clearing alone determines the response of the risky-asset price to portfolio flows. Once goods-market clearing is imposed and the interest rate adjusts endogenously, the determination of aggregate market elasticity changes fundamentally, as shown by the next proposition.

Proposition 1 (Infinite market elasticity)**.** *Suppose that the aggregate cross-price elasticity* $\zeta_r \equiv s_a \zeta_{a,r} + s_p \zeta_{p,r}$ *is nonzero. Then a portfolio-flow shock has no price impact:*

$$\hat{p} = 0.$$

Portfolio flows instead have offsetting effects on the interest rate and the risk premium:

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

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

Proof. Equation (6) implies $\hat{p} = 0$. Given $\hat{p} = 0$ and $\zeta_r \neq 0$, Equation (5) implies $\hat{r} = f/\zeta_r$. Because $\hat{\pi} = -\hat{p} - \hat{r}$, it follows that $\hat{\pi} = -f/\zeta_r$. □

Goods-market clearing implies that a portfolio flow has no price impact, so the market is infinitely elastic. Provided that the aggregate cross-price elasticity ζ_r is nonzero, the interest

rate adjusts to clear the risky-asset market. Its magnitude determines the size of the offsetting movements in the interest rate and the risk premium. This result holds even when the aggregate own-price demand elasticity ζ_p is arbitrarily small. Thus, an inelastic risky-asset demand curve need not imply an inelastic aggregate market once the interest rate adjusts endogenously.

How can the market be infinitely elastic even when risky-asset demand is inelastic? Consider a positive portfolio flow into the risky asset. If the interest rate were held fixed, the flow would raise the price of the risky asset. The resulting increase in wealth would induce households to increase consumption. Because aggregate output is fixed, this increase in consumption is inconsistent with goods-market clearing. The interest rate therefore adjusts, shifting the demand curve for the risky asset until it intersects the new net-supply curve at the original price. Graphically, the interest-rate adjustment shifts the active-investor demand curve so that the new equilibrium has $\hat{p} = 0$.

The division of roles between the two market-clearing conditions makes this result transparent. The goods-market condition determines the price-dividend ratio through households' consumption-saving behavior. Conditional on that price-dividend ratio, the risky-asset market determines the interest rate and the risk premium.[11] Consequently, the aggregate own-price demand elasticity ζ_p does not determine aggregate market elasticity in this benchmark. In our setting, the aggregate cross-price elasticity ζ_r determines the magnitude of the offsetting movements in the interest rate and the risk premium. Because the risky-asset price does not move, own-price elasticities play no role in the equilibrium adjustment.

This mechanism is particularly transparent in models with unit EIS, in which the consumption-wealth ratio is constant and the price-dividend ratio is tied to the subjective discount rate. As long as the subjective discount rate remains unchanged, shocks are absorbed through offsetting movements in the interest rate and the risk premium, leaving the risky-asset price unchanged. Our framework generalizes this logic by allowing consumption-saving behavior to respond to expected returns. The central implication is that aggregate market elasticity is determined by the consumption-saving block of the economy rather than by the average own-price demand elasticity.

A fundamental shock. It is instructive to contrast the effects of a portfolio-flow shock with those of a fundamental shock. Consider a shock to expected dividend growth, $\hat{\mu} \equiv \mu - \mu^*$. The linearized demand system remains as in Equation (4), but the fundamental shock introduces the demand shifter $f_j \equiv \frac{\partial \log \overline{Q}_j}{\partial \mu} \hat{\mu}$ for investor j.

A change in expected dividend growth also affects households' incentives to save. The change

[11] The equilibrium can equivalently be described using the markets for risky assets and bonds rather than the markets for risky assets and goods. Appendix A.1 presents this alternative formulation.

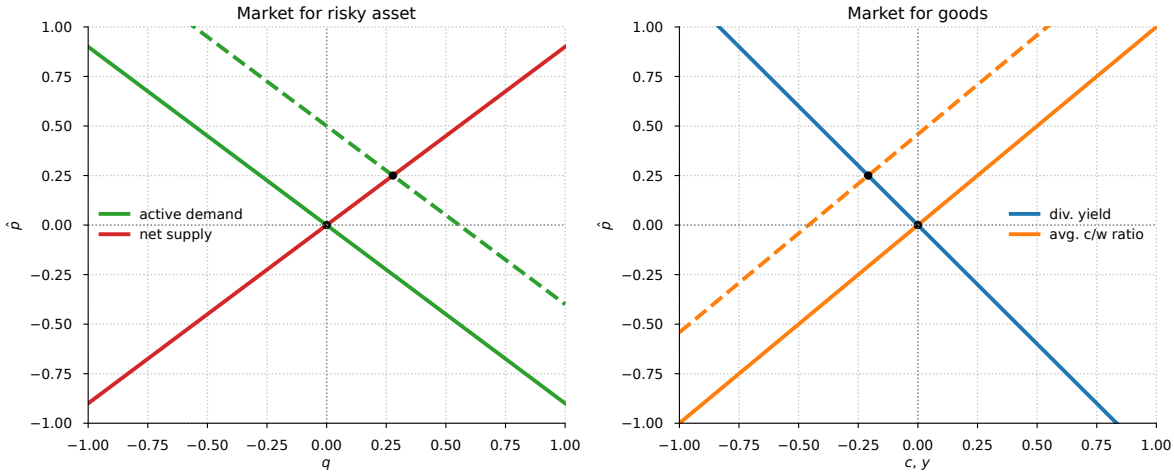


Figure 2. Equilibrium effects of a fundamental shock in the risky-asset market (left) and goods market (right).

in the average consumption-wealth ratio can be written as

$$\hat{c} = \chi_\mu \hat{\mu} + \chi_p \hat{p}, \qquad \chi_\mu = -\chi_p,$$

where the restriction $\chi_\mu = -\chi_p$ follows from Assumption 1. A fundamental shock therefore shifts both risky-asset demand and, when $\chi_\mu \neq 0$, the consumption-wealth schedule. The shift in risky-asset demand determines the equilibrium interest rate and risk premium conditional on the price-dividend ratio. The shift in the consumption-wealth schedule determines the response of the price-dividend ratio. Figure 2 illustrates these equilibrium adjustments.

Formally, using goods-market clearing, we obtain the price response to a fundamental shock:

$$\hat{p} = -\frac{\chi_\mu}{\chi_p + d_p}\hat{\mu}. \tag{8}$$

Hence, a fundamental shock changes the price-dividend ratio whenever it changes the consumption-wealth ratio, that is, whenever $\chi_\mu \neq 0$. Moreover, the magnitude of the price response is independent of the individual investors' demand elasticities $\zeta_{j,p}$ and $\zeta_{j,r}$.

The unit-EIS case, in which $\chi_p = \chi_\mu = 0$, shows that the logic behind Proposition 1 is not specific to portfolio flows. With unit EIS, neither a portfolio-flow shock nor cash-flow news shifts the goods-market condition. The price-dividend ratio therefore remains unchanged in both cases. In particular, positive cash-flow news changes the equilibrium interest rate and risk premium rather than the price-dividend ratio.

This discussion also clarifies when portfolio flows generate nonzero price impact. When the consumption-wealth ratio responds directly to portfolio flows, the flow shifts the goods-market condition that determines the price-dividend ratio. The interest rate then no longer fully offsets the

risk-premium response, and aggregate market elasticity is finite.

The need for a micro-founded demand system. The analysis identifies the consumption-saving response as the key determinant of aggregate market elasticity. In the reduced-form model, a portfolio flow shifts risky-asset demand but, by assumption, does not directly affect the average consumption-wealth ratio. In this case, the interest rate adjusts to clear the risky-asset market, leaving the price-dividend ratio unchanged. If portfolio flows instead shift the consumption-wealth ratio, as does the fundamental shock considered above, the price-dividend ratio responds and aggregate market elasticity is finite.

Determining whether and how portfolio flows affect consumption-saving behavior requires a micro-founded framework. Such a framework must jointly connect portfolio reallocation to risky-asset demand, wealth, and saving decisions. Section 3 develops this framework.

3 A Dynamic Asset Pricing Model with Passive Investors

We study a dynamic heterogeneous-agent economy with passive and active investors in which portfolio demand and consumption-saving behavior are jointly derived from preferences and financial constraints. Passive investors hold an exogenous, stochastic share of their wealth in risky assets. There are two types of active investors who optimally choose their portfolios subject to financial frictions: a more risk-averse unconstrained investor and a less risk-averse investor subject to an occasionally binding leverage constraint.

The representative firm issues equity, which is a levered claim on the aggregate endowment, and risky corporate debt. The distinction between the equity and endowment claims will be quantitatively important for the price impact of portfolio flows.

3.1 Environment

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

$$\frac{dY_t}{Y_t} = \mu\, dt + \sigma\, dZ_t, \tag{9}$$

where $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}^{1\times d}$ are constants.[12] 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, $d = 2$, capturing two sources of risk: (i) a fundamental shock to aggregate endowment and (ii) a non-fundamental portfolio-flow shock to passive investors.

[12] Throughout the paper, we adopt the convention that loadings on the Brownian motion are row vectors.

We consider a representative firm that issues equity and risky corporate debt. Let D_t and B_t denote the respective payoff flows, which satisfy $D_t + B_t = Y_t$. We assume that the dividend process is given by

$$\frac{dD_t}{D_t} = \mu\, dt + \phi\sigma\, dZ_t, \tag{10}$$

where $\phi \geq 1$ is the cash-flow leverage parameter. This specification captures the well-known fact that dividends are more volatile than consumption (see, e.g., Abel 1999; Campbell 1999). The corporate-debt payoff is determined residually as $B_t = Y_t - D_t$. When $B_t < 0$, we interpret the negative net payout as corporate saving rather than borrowing.

Let $P_{D,t}$ and $P_{B,t}$ denote the values of equity and corporate debt, respectively. The value of the firm is given by $P_{Y,t} = P_{D,t} + P_{B,t}$. Let $\mathcal{D}_{k,t}$ denote the payout flow of claim $k \in \{D, B, Y\}$, so that $\mathcal{D}_{D,t} = D_t$, $\mathcal{D}_{B,t} = B_t$, and $\mathcal{D}_{Y,t} = Y_t$. The return on claim k is

$$dR_{k,t} = \frac{dP_{k,t} + \mathcal{D}_{k,t}\, dt}{P_{k,t}} = \mu_{R_k,t}\, dt + \sigma_{R_k,t}\, dZ_t.$$

Stochastic discount factor. Absence of arbitrage implies the existence of a stochastic discount factor (SDF), Λ_t, with dynamics

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

where $\eta_t \in \mathbb{R}^{1\times d}$ is the vector of market prices of risk.

Define the vector of risk premia as $\pi_t \equiv [\pi_{D,t}, \pi_{B,t}]^\top$, where $\pi_{k,t} \equiv \mu_{R_k,t} - r_t$ for $k \in \{D, B\}$. The no-arbitrage restrictions can be written compactly as $\pi_t = \Sigma_t \eta_t^\top$, where Σ_t collects the risk exposures of the two risky assets:

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

We assume that Σ_t is nonsingular, so the equity and corporate-debt claims dynamically span the two aggregate shocks.

Demography and preferences. The economy consists of three investor types, indexed by $j \in \mathcal{J} = \{p, h, \ell\}$, with population masses ω_j that sum to one. Type p investors are passive, while types h and ℓ are active. Investors die with Poisson intensity κ, and at each instant a mass $\kappa\omega_j$ of type-j investors is born, leaving the total population constant. The estates of deceased investors are distributed across newborns, who enter with average wealth. This redistribution leaves aggregate resources unchanged and induces mean reversion in the wealth distribution across investor types.

Investors have recursive preferences as in Duffie and Epstein (1992). Type j has relative risk aversion γ_j, while all types have a common elasticity of intertemporal substitution (EIS) ψ and

effective discount rate $\rho \equiv \widehat{\rho} + \kappa$, where $\widehat{\rho}$ is the subjective discount rate. The recursive utility aggregator is

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

Each investor chooses consumption and, subject to the restrictions described below, portfolio allocations across the risk-free bond, equity, and corporate debt.

Passive investors. Passive investors allocate a fraction $\overline{\alpha}_{p,t}$ of their wealth to a *market index* fund and a fraction $1 - \overline{\alpha}_{p,t}$ to the risk-free asset. The market index holds the equity and corporate debt in proportion to their market-capitalization weights.

The share of wealth invested in the market index, $\overline{\alpha}_{p,t}$, follows an Ornstein–Uhlenbeck process:

$$d\overline{\alpha}_{p,t} = \kappa_p \left(\overline{\alpha} - \overline{\alpha}_{p,t}\right) dt + \sigma_{\alpha_p} dZ_t, \tag{11}$$

where $\overline{\alpha} > 0$ is its long-run mean, $\kappa_p > 0$ governs mean reversion, and $\sigma_{\alpha_p} \in \mathbb{R}^{1\times d}$ is its exposure to aggregate shocks. Innovations to $\overline{\alpha}_{p,t}$ change the fraction of passive wealth allocated to the market index and therefore induce portfolio flows between the risk-free asset and risky claims. We refer to these innovations as *portfolio-flow shocks*. As in Section 2, these shocks exogenously affect the demand of passive investors and trigger an endogenous reallocation of risky assets across investors.

Portfolio flows may comove with cash-flow shocks or arise independently. We impose the following structure on the shock exposures:

$$\sigma = (\sigma_1, 0), \qquad \sigma_{\alpha_p} = (\sigma_{\alpha_p,1}, \sigma_{\alpha_p,2}).$$

The first innovation corresponds to a fundamental shock to aggregate endowment growth. The second innovation corresponds to a non-fundamental portfolio-flow shock, orthogonal to cash-flow shocks. This structure allows portfolio flows to be correlated with aggregate endowment shocks while retaining distinct non-fundamental flow and cash-flow innovations.

The exposure of the passive portfolio share to aggregate shocks is consistent with both *imperfect rebalancing* and *performance-induced flows*. With imperfect rebalancing, a positive risky return mechanically raises the risky share when $\overline{\alpha}_{p,t} < 1$. With performance-induced flows, investors actively increase their desired risky exposure following strong returns. The independent component of portfolio flows captures reallocations that are not driven by contemporaneous fundamentals, such as changes in portfolio rules or default investment options. These channels are consistent with evidence that households rebalance infrequently (Ameriks and Zeldes, 2004; Brunnermeier and Nagel, 2008), that flows respond to past performance (Chevalier and Ellison, 1997; Dannhauser and

Pontiff, 2024), and that changes in default investment options can generate substantial reallocations toward equities (Parker, Schoar, Cole, and Simester, 2025).

Active investors. Investors of types h and ℓ are *active investors* who optimally rebalance their portfolios. Let $\alpha_{j,t} \equiv [\alpha_{j,D,t}, \alpha_{j,B,t}]^\top$ denote the portfolio allocations of active investor $j \in \{h, \ell\}$ across equity and corporate debt. The investor's total risk exposure is

$$\sigma_{j,t} \equiv \alpha_{j,t}^\top \Sigma_t = \alpha_{j,D,t}\sigma_{R_D,t} + \alpha_{j,B,t}\sigma_{R_B,t}.$$

Type h investors have high risk aversion and face no portfolio constraint. Type ℓ investors have lower risk aversion, $\gamma_\ell < \gamma_h$, and therefore desire greater risk exposure. Their risk exposure is subject to the leverage constraint

$$\|\sigma_{\ell,t}\| \leq \overline{\sigma},$$

where $\|\cdot\|$ denotes the Euclidean norm. This constraint resembles a Value-at-Risk (VaR) limit commonly applied to banks and leveraged institutions (see, e.g., Adrian and Shin 2014 and Brunnermeier and Pedersen 2009). The constraint binds when the desired risk exposure of type ℓ investors exceeds $\overline{\sigma}$ and is otherwise slack.

Investors' problem. Each investor chooses consumption, and active investors also choose their portfolio allocations. We represent these choices uniformly by allowing the investment set $\Omega_{j,t}$ to depend on investor type. Investor $j \in \mathcal{J}$ solves

$$V_{j,t} = \max_{C_{j,t},\alpha_{j,t}} \mathbb{E}_t \left[\int_t^\infty f_j(C_{j,s}, V_{j,s})\, ds \right] \tag{12}$$

subject to

$$\frac{dW_{j,t}}{W_{j,t}} = \left[r_t + \pi_t^\top \alpha_{j,t} - \frac{C_{j,t}}{W_{j,t}} \right] dt + \sigma_{j,t}\, dZ_t,$$

with $W_{j,t} \geq 0$ and $\alpha_{j,t} \in \Omega_{j,t}$, where $\sigma_{j,t} \equiv \alpha_{j,t}^\top \Sigma_t$. Here, $W_{j,t}$ denotes wealth and $\Omega_{j,t}$ denotes the type-specific investment set.

The investment sets capture differences in portfolio behavior and financial constraints across investors.[13] For passive investors, the investment set is the singleton implementing their prescribed allocation to the market index:

$$\Omega_{p,t} = \left\{ \alpha_{p,t} : \ \alpha_{p,D,t} = \overline{\alpha}_{p,t} \frac{P_{D,t}}{P_{Y,t}}, \ \alpha_{p,B,t} = \overline{\alpha}_{p,t} \frac{P_{B,t}}{P_{Y,t}} \right\},$$

[13] The differences in investment sets here are reminiscent of differences in investment universes emphasized by Koijen and Yogo (2019).

Type h investors are unconstrained, so $\Omega_{h,t} = \mathbb{R}^2$. The investment set of type ℓ investors imposes the leverage constraint:

$$\Omega_{\ell,t} = \left\{\alpha_{\ell,t} : \|\Sigma_t^\top \alpha_{\ell,t}\| \leq \overline{\sigma}\right\}.$$

Market clearing and equilibrium.

Definition 1. *A competitive equilibrium consists of adapted processes (P_D, P_B, r, η) and, for every $j \in \mathcal{J}$, allocations (C_j, α_j) and wealth W_j, such that:*

(i) The aggregate endowment Y, dividends D, and the passive portfolio share $\overline{\alpha}_p$ evolve according to (9), (10), and (11), respectively.

(ii) The firm's payoff flows satisfy $B_t = Y_t - D_t$, and the value of the firm is $P_{Y,t} = P_{D,t} + P_{B,t}$.

(iii) Asset returns satisfy the no-arbitrage restrictions: $\pi_t = \Sigma_t \eta_t^\top$.

(iv) Given prices and $\overline{\alpha}_{p,t}$, $(C_{j,t}, \alpha_{j,t})$ solves the problem in (12) subject to the type-specific investment set $\Omega_{j,t}$ for every $j \in \mathcal{J}$.

(v) The goods, equity, corporate-debt, and risk-free-bond markets clear:

$$\sum_{j\in\mathcal{J}} \omega_j C_{j,t} = Y_t, \qquad \sum_{j\in\mathcal{J}} \omega_j Q_{j,k,t} = 1, \qquad \sum_{j\in\mathcal{J}} \omega_j A_{j,t} = 0,$$

for $k \in \{D, B\}$. Here,

$$Q_{j,k,t} = \frac{\alpha_{j,k,t} W_{j,t}}{P_{k,t}}, \qquad A_{j,t} = (1 - \alpha_{j,D,t} - \alpha_{j,B,t}) W_{j,t}.$$

3.2 Equilibrium characterization

We now characterize equilibrium in the dynamic model. We focus on a Markov equilibrium in which prices and return dynamics are functions of an aggregate state vector $X_t \in \mathbb{R}^N$. The state vector follows

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

where $\mu_{X,t} \in \mathbb{R}^N$ and $\sigma_{X,t} \in \mathbb{R}^{N\times d}$ denote its drift and diffusion. We specify the components of X_t and characterize their dynamics below.

3.2.1 Equilibrium conditions

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

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

The function $c_{j,t} = c_j(X_t)$ denotes the consumption-wealth ratio, so that $C_{j,t}/W_{j,t} = c_{j,t}$.

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

$$c_{j,t} = \psi\rho + (1-\psi)\left[r_t + \eta_t\, \sigma_{j,t}^{\top} - \frac{\gamma_j}{2}\|\sigma_{j,t}\|^2\right] + \xi_{j,t}, \tag{13}$$

where $\sigma_{j,t} = \alpha_{j,t}^{\top}\Sigma_t$ is the investor's risk exposure. The first term reflects time preference, while the second term captures the impact of current investment opportunities on consumption.

The term $\xi_{j,t}$ is given by

$$\xi_{j,t} = \mu_{c_j,t} + (1-\gamma_j)\sigma_{c_j,t}\sigma_{j,t}^{\top} + \frac{\psi-\gamma_j}{1-\psi}\frac{\|\sigma_{c_j,t}\|^2}{2}.$$

This term captures the effect of changes in future investment opportunities on consumption.

Notice that $\eta_t\sigma_{j,t}^{\top} = \pi_t^{\top}\alpha_{j,t}$, so the consumption-wealth ratio depends on both the interest rate r_t and the expected excess return on the investor's portfolio, as in Section 2. Importantly, it also depends on the investor's portfolio choice $\alpha_{j,t}$. Finally, $\xi_{j,t}$ acts as a state-dependent shifter, and it is equal to zero when investment opportunities are constant.

Optimal risk exposure. The optimal risk exposure for active investors $j \in \{h, \ell\}$ is given by

$$\sigma_{j,t} = \frac{1}{1+\nu_{j,t}}\left[\frac{\eta_t}{\gamma_j} + \frac{1-\gamma_j^{-1}}{\psi-1}\,\sigma_{c_j,t}\right], \tag{14}$$

where $\nu_{j,t}$ is the Lagrange multiplier on the leverage constraint, so $\nu_{j,t} \geq 0$ and $\nu_{h,t} = 0$. When the leverage constraint is not binding, the optimal risk exposure is the sum of two components. The first is the myopic demand, $\frac{\eta_t}{\gamma_j}$, which depends on the investor's risk tolerance and the vector of market prices of risk η_t. This term captures the compensation per unit of exposure to each risk factor: $\eta_t^{\top} = \Sigma_t^{-1}\pi_t$. The second is the hedging demand, $\frac{1-\gamma_j^{-1}}{\psi-1}\,\sigma_{c_j,t}$, which reflects the investor's desire to hedge fluctuations in future investment opportunities. When the constraint binds, the multiplier $\nu_{j,t} > 0$ scales down exposure uniformly across all risk factors to satisfy $\|\sigma_{j,t}\| \leq \overline{\sigma}$. This acts like an endogenous increase in effective risk aversion.

The Lagrange multiplier $\nu_{\ell,t}$ satisfies the complementary slackness condition:

$$\nu_{\ell,t}\left(\|\sigma_{\ell,t}\| - \overline{\sigma}\right) = 0.$$

For passive investors, risk exposure is given by $\sigma_{p,t} = \overline{\alpha}_{p,t}\sigma_{R_Y,t}$, where $\sigma_{R_Y,t}$ is the risk exposure of the market index fund. Hence, $\overline{\alpha}_{p,t}$ controls the overall risk exposure of passive investors.

Given an investor's risk exposure, the corresponding portfolio shares are recovered from

$$\alpha_{j,t} = \left(\sigma_{j,t}\,\Sigma_t^{-1}\right)^\top,$$

which maps risk exposures into asset holdings.

Market clearing conditions. Market clearing implies that aggregate wealth equals the price of the aggregate claim:

$$\sum_{j\in\mathcal{J}} \omega_j W_{j,t} = P_{Y,t}.$$

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 market clearing conditions for goods and risky assets are given by:

$$\sum_{j\in\mathcal{J}} x_{j,t} c_{j,t} = \frac{1}{p_{Y,t}}, \qquad \sum_{j\in\mathcal{J}} x_{j,t}\sigma_{j,t} = \sigma_{R_Y,t},$$

where $p_{Y,t} \equiv P_{Y,t}/Y_t$ is the price-dividend ratio of the aggregate claim.

These conditions echo those in the simple model. The first condition states that the wealth-weighted average consumption-wealth ratio equals the dividend yield of the aggregate claim. The second condition states that the wealth-weighted average risk exposure equals the risk exposure of the economy's endowment claim.

Market price of risk. We can write the market clearing condition for risky assets as follows:

$$\underbrace{\sum_{j\in\{h,\ell\}} \frac{x_{j,t}}{x_{a,t}}\sigma_{j,t}}_{\substack{\text{active demand}\\ \text{for risk}}} = \underbrace{\chi_{a,t}\sigma_{R_Y,t}}_{\substack{\text{net supply}\\ \text{of risk}}},$$

where $\chi_{a,t} \equiv \frac{1-x_{p,t}\overline{\alpha}_{p,t}}{1-x_{p,t}}$ captures the share of aggregate risk that must be absorbed by active investors, and $x_{a,t} \equiv 1 - x_{p,t}$ is the wealth share of active investors.

This is the dynamic counterpart of the risk market clearing condition in Section 2. The expression above states that active investors absorb the net supply of risk, defined as total risk minus the portion held by passive investors.

Substituting optimal risk exposures into the market clearing condition yields the market price of risk:

$$\eta_t = \gamma_{a,t} \left[\chi_{a,t} \sigma_{R_Y,t} - h_{a,t} \right], \tag{15}$$

where $h_{a,t} = \sum_{j \in \{h,\ell\}} \frac{x_{j,t}}{x_{a,t}} \frac{1-\gamma_j^{-1}}{\psi-1} \frac{\sigma_{c_j,t}}{1+\nu_{j,t}}$ is the average hedging demand of active investors, adjusted for leverage constraints, and $\gamma_{a,t}$ is the average effective risk aversion:

$$\gamma_{a,t} \equiv \left[\sum_{j \in \{h,\ell\}} \frac{x_{j,t}}{x_{a,t}} \frac{1}{\gamma_j (1+\nu_{j,t})} \right]^{-1}.$$

Equation (15) shows that the market price of risk is determined by three forces. First, the average effective risk aversion of active investors, $\gamma_{a,t}$. This varies over time with the wealth distribution and the extent to which leverage constraints bind. Second, the net supply of risk to active investors, captured by $\chi_{a,t}\sigma_{R_Y,t}$. As passive investors increase their holdings of the market index, the amount of risk that must be absorbed by active investors falls, reducing the compensation for risk. Third, the average hedging demand of active investors, $h_{a,t}$. Higher hedging demand implies active investors are willing to bear more risk, lowering the equilibrium compensation for risk.

Equation (15) makes clear that portfolio flows affect risk premia through their impact on the net supply of risk absorbed by active investors.

Interest rate. Combining goods market clearing with the expression for the consumption-wealth ratio yields the equilibrium interest rate:

$$r_t = \rho + \psi^{-1} \left(\mu_{P_Y,t} + \xi_t \right) + (1 - \psi^{-1}) \sum_{j \in \mathcal{J}} x_{j,t} \frac{\gamma_j}{2} \|\sigma_{j,t}\|^2 - \pi_{Y,t}, \tag{16}$$

where $\mu_{P_Y,t}$ is the expected capital-gain rate on the aggregate claim, $\pi_{Y,t} \equiv \sigma_{R_Y,t}\eta_t^\top$ is its risk premium, and $\xi_t \equiv \sum_{j \in \mathcal{J}} x_{j,t}\xi_{j,t}$.

The first term reflects time preference. The second term captures intertemporal substitution. The remaining terms capture the effects of risk allocation and risk premia on the interest rate.

This expression highlights that the interest rate is determined jointly by risk premia and the allocation of risk across investors, providing the key general equilibrium link between portfolio

flows, risk allocation, and saving behavior.

Pricing condition. Define the valuation ratios $p_{D,t} \equiv P_{D,t}/D_t$ and $p_{Y,t} \equiv P_{Y,t}/Y_t$ for the equity and endowment claims, respectively. The pricing condition for each claim $k \in \{Y, D\}$ is given by

$$\frac{1}{p_{k,t}} = r_t + \eta_t(\phi_k\sigma + \sigma_{p_k,t})^\top - \left(\mu + \mu_{p_k,t} + \phi_k\sigma\,\sigma_{p_k,t}^\top\right), \tag{17}$$

where $\phi_Y = 1$ and $\phi_D = \phi$, using the fact that $\sigma_{R_k,t} = \phi_k\sigma + \sigma_{p_k,t}$. The terms $\mu_{p_k,t}$ and $\sigma_{p_k,t}$ denote the drift and diffusion of the valuation ratios, respectively.

3.2.2 Markov equilibrium

State variables. We can express the normalized aggregate equilibrium variables as functions of the state vector $X_t = (x_t, \overline{\alpha}_{p,t})$, where $x_t = (x_{h,t}, x_{\ell,t})$ is the vector of wealth shares of active investors and $\overline{\alpha}_{p,t}$ is the passive portfolio share. To obtain the law of motion of the aggregate state vector, $\mu_{X,t}$ and $\sigma_{X,t}$, we need to solve for the drift and diffusion of the endogenous state variables $(x_{h,t}, x_{\ell,t})$.

By Itô's lemma, the law of motion for $x_{j,t}$, $j \in \{h, \ell\}$, is given by

$$\frac{dx_{j,t}}{x_{j,t}} = \left[(\eta_t - \sigma_{R_Y,t})\left(\sigma_{j,t} - \sigma_{R_Y,t}\right)^\top - \left(c_{j,t} - \frac{1}{p_{Y,t}}\right) + \kappa\frac{\omega_j - x_{j,t}}{x_{j,t}}\right]dt + (\sigma_{j,t} - \sigma_{R_Y,t})\,dZ_t.$$

Endogenous volatility. Return volatility has both a cash-flow and a valuation component, $\sigma_{R_k,t} = \phi_k\sigma + \sigma_{p_k,t}$, where $\sigma_{p_k,t}$ arises from movements in the valuation ratio $p_{k,t}$ for asset $k \in \{D, Y\}$. By Itô's lemma, the endogenous component of return volatility is given by:

$$\sigma_{p_k,t} = \frac{p_{k,x}(X_t)}{p_k(X_t)}\sigma_{x,t} + \frac{p_{k,\overline{\alpha}_p}(X_t)}{p_k(X_t)}\sigma_{\overline{\alpha}_p,t} \tag{18}$$

Equation (18) shows that endogenous volatility depends on the sensitivity of the valuation ratio to both the wealth distribution and the passive portfolio share. The first term is standard in heterogeneous-agent models (see Panageas, 2020). The second term arises only in the presence of exogenous portfolio-flow shocks. This second term is quantitatively important when markets are sufficiently inelastic, that is, when price-dividend ratios are highly sensitive to changes in passive demand. Thus, (18) provides a direct link between market elasticity and return volatility, as lower elasticity increases endogenous volatility.

The PDE system. Equations (15) and (16) allow us to express the market price of risk η_t and the interest rate r_t as functions of the consumption-wealth ratios $c_j(X_t)$ and the valuation ratio for the

aggregate endowment $p_Y(X_t)$. Using Itô's lemma, the drift and diffusion terms $(\mu_{c_j,t}, \mu_{p_k,t})$ and $(\sigma_{c_j,t}, \sigma_{p_k,t})$ can be written as functions of the state dynamics $(\mu_{X,t}, \sigma_{X,t})$.

Substituting these expressions into the HJB equation (13) and the pricing condition (17), together with the laws of motion for the state variables, yields a system of four partial differential equations for the consumption-wealth ratios $\{c_j(X_t)\}_{j\in\mathcal{J}}$ and $p_Y(X_t)$.

Proposition 2 (PDE system)**.** *The consumption-wealth ratios $\{c_j(X_t)\}_{j\in\mathcal{J}}$ and the valuation ratio for the aggregate endowment $p_Y(X_t)$ are functions of the aggregate state vector X_t and satisfy the following system of PDEs:*

$$
\begin{aligned}
c_j &= \psi\rho + (1-\psi)\left[r + \eta\,\sigma_j^\top - \frac{\gamma_j}{2}\|\sigma_j\|^2\right] + \mu_{c_j} + (1-\gamma_j)\sigma_{c_j}\sigma_j^\top + \frac{\psi-\gamma_j}{1-\psi}\frac{\|\sigma_{c_j}\|^2}{2} \\
\frac{1}{p_Y} &= r + \eta\,(\sigma+\sigma_{p_Y})^\top - \left(\mu + \mu_{p_Y} + \sigma\,\sigma_{p_Y}^\top\right),
\end{aligned}
\tag{19}
$$

where $(r, \eta, \sigma_j, \mu_{c_j}, \sigma_{c_j}, \mu_{p_Y}, \sigma_{p_Y})$ are all functions of X_t, (c_j, p_Y) and their derivatives.

The PDE system closes in $\{c_j(X_t)\}_{j\in\mathcal{J}}$ and $p_Y(X_t)$ without requiring the equity valuation ratio $p_D(X_t)$. We can therefore characterize the equilibrium in two steps. First, we solve for the $\{c_j(X_t)\}_{j\in\mathcal{J}}$ and $p_Y(X_t)$ as functions of the aggregate state vector X_t. As we can express (r, η) as functions of (c_j, p_Y) and their derivatives, this step allows us to recover the economy's stochastic discount factor. Second, we can solve for $p_{D,t}$ by solving the pricing condition (17) for the dividend claim. Given D_t and Y_t, we can recover both $P_{D,t}$ and $P_{Y,t}$, as well as the value of corporate debt, $P_{B,t} = P_{Y,t} - P_{D,t}$.

This completes the characterization of the equilibrium. We now turn to an analytical characterization of the aggregate market elasticity in this economy.

4 The Determinants of the Aggregate Market Elasticity

This section characterizes the determinants of aggregate market elasticity. We find that the aggregate claim is infinitely elastic to first order around the efficient benchmark, even when individual investors have finite demand elasticities. Second-order departures from the benchmark generate finite aggregate market elasticity by changing consumption–saving decisions. Thus, the neutrality result from Section 2 extends to the dynamic model, and portfolio frictions are essential for generating finite aggregate market elasticity.

To characterize aggregate market elasticity analytically, we extend the perturbation method of Kogan and Uppal (2001). This method enables us to derive asymptotic closed-form approximations in an environment with heterogeneous agents and rich financial frictions.

4.1 Perturbation method

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 aggregate claim at all times, $\overline{\alpha}_{p,t} = 1$. In this case, the economy collapses to a Lucas economy, as shown in Lemma 1. Thus, the benchmark corresponds to an efficient frictionless allocation. Throughout, a superscript $*$ 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,*

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

$c_j^* = (p_Y^*)^{-1}$, *and* $\sigma_j^* = \sigma$ *for* $j \in \mathcal{J}$.

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 several dimensions. When $\epsilon = 0$, the economy coincides with the benchmark Lucas economy. As ϵ increases, the economy moves away from the benchmark toward a heterogeneous-agent economy with frictions.

First, investors have heterogeneous risk aversion:

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

so that γ is the population-weighted average and $\hat{\gamma}_j$ measures deviations from it.

Second, passive investors need not be fully invested in the aggregate claim. Their portfolio share is $\overline{\alpha}_{p,t} = 1 + \hat{\alpha}_{p,t}\,\epsilon$, where $\hat{\alpha}_{p,t}$ follows an Ornstein-Uhlenbeck process independent of ϵ:

$$d\hat{\alpha}_{p,t} = -\kappa_p\hat{\alpha}_{p,t}dt + \hat{\sigma}_p dZ_t,$$

which requires $\overline{\alpha} = 1$ and $\sigma_{\alpha_p} = \hat{\sigma}_p\epsilon$ to be consistent with Equation (11). When $\epsilon = 0$, passive investors are fully invested in the risky asset at all times, recovering the benchmark case.

Third, the leverage constraint coefficient is perturbed according to

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

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

Equilibrium objects depend on the parameter ϵ and the aggregate state variable X_t. As shown in Proposition 2, the aggregate equilibrium closes in the consumption-wealth ratios and the valuation

ratio of the aggregate endowment claim, which depend on $X_t = (x_t, \overline{\alpha}_{p,t})$. We expand them to second order in ϵ as follows:

$$p_Y(x, \overline{\alpha}_p; \epsilon) = p_{Y,0}(x, \hat{\alpha}_p) + p_{Y,1}(x, \hat{\alpha}_p)\epsilon + p_{Y,2}(x, \hat{\alpha}_p)\epsilon^2 + O(\epsilon^3),$$
$$c_j(x, \overline{\alpha}_p; \epsilon) = c_{j,0}(x, \hat{\alpha}_p) + c_{j,1}(x, \hat{\alpha}_p)\epsilon + c_{j,2}(x, \hat{\alpha}_p)\epsilon^2 + O(\epsilon^3),$$

for $j \in \mathcal{J}$.

Unlike standard DSGE perturbations, we do not assume that endowment volatility σ is small. Moreover, the approximation is local in the exogenous state variable $\overline{\alpha}_p$, but global in the endogenous state variables x: we solve directly for functions of x rather than coefficients around a deterministic steady state. This type of *state-global perturbation* builds on the perturbation method of Kogan and Uppal (2001) and, more closely, its extension in Silva (2026).[14]

This approach is important for two reasons. First, it allows us to capture state-dependent effects and time-varying risk premia in a tractable way.[15] Second, it 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 $p_{Y,1}(x, \hat{\alpha}_p)$ and $c_{j,1}(x, \hat{\alpha}_p)$. These capture the effects of a small deviation from the frictionless benchmark. Second, we solve for the second-order terms $p_{Y,2}(x, \hat{\alpha}_p)$ and $c_{j,2}(x, \hat{\alpha}_p)$. A second-order approximation is necessary to capture how the different frictions interact to determine aggregate market elasticity. Zeroth-order terms are given by Lemma 1, i.e., $p_{Y,0}(x, \hat{\alpha}_p) = p_Y^*$ and $c_{j,0}(x, \hat{\alpha}_p) = c_j^*$.

4.2 The first-order demand system

We start by characterizing the first-order corrections to equilibrium prices and allocations. In particular, we derive the first-order demand system and the consumption-wealth ratio, which jointly determine how risk premia and interest rates respond to portfolio flows.

Endogenous volatility and hedging demands. An important implication of the perturbation assumption is that the drift and diffusion of the state variables are first-order in ϵ: $\mu_X = O(\epsilon)$ and $\sigma_X = O(\epsilon)$. As shown in Appendix A.3, the derivatives of p_Y and c_j with respect to the state variables are also first-order in ϵ: $p_{Y,X}(X) = O(\epsilon)$ and $c_{j,X}(X) = O(\epsilon)$.

[14] See Kargar et al. (2023) for an application of state-global perturbations to a model with search frictions and a detailed comparison with standard perturbation methods.

[15] In contrast, first-order perturbations around a deterministic steady state feature no risk premia or state dependence, while second-order perturbations only generate constant risk corrections (see, e.g., Schmitt-Grohé and Uribe 2004).

Endogenous volatility and hedging demand are then second order in ϵ:

$$\sigma_{p_Y} = \frac{p_{Y,X}(X)}{p_Y(X)}\,\sigma_X(X) = O(\epsilon^2), \qquad \sigma_{c_j} = \frac{c_{j,X}(X)}{c_j(X)}\,\sigma_X(X) = O(\epsilon^2),$$

so $\sigma_{R_Y} = \sigma + O(\epsilon^2)$ and $h_a = O(\epsilon^2)$. A similar argument implies $\mu_{p_Y} = O(\epsilon^2)$ and $\mu_{c_j} = O(\epsilon^2)$.

Demand for the risky claim. Up to first order, each active investor's optimal risk exposure is proportional to the exposure of the aggregate claim, while passive investors hold the aggregate claim by assumption. The optimal allocation can therefore be implemented, up to first order, with equal holdings of equity and corporate debt:

$$Q_{j,D,t} = Q_{j,B,t} \equiv Q_{j,t}.$$

We refer to their common value $Q_{j,t}$ as investor j's demand for the risky claim. Define the proportional deviation of risky-claim demand from the corresponding benchmark allocation Q_j^* as $q_{j,t} \equiv \frac{Q_{j,t}-Q_j^*}{Q_j^*}$. Under this implementation, investor j's risk exposure satisfies

$$\sigma_{j,t} = (1 + q_{j,t})\sigma + O(\epsilon^2).$$

Expanding the optimal risk exposure in Equation (14) therefore yields, for active investors,

$$q_{j,t} = \frac{\hat{\pi}_{Y,t}}{\gamma\|\sigma\|^2} - (\hat{\gamma}_j + \hat{\nu}_{j,t})\,\epsilon + O(\epsilon^2),$$

where $\hat{\pi}_{Y,t} \equiv \pi_{Y,t} - \pi_Y^*$ denotes the deviation of the risk premium from the benchmark and $\nu_{j,t} = \hat{\nu}_{j,t}\epsilon + O(\epsilon^2)$. Hence, since hedging demand is second order in ϵ, optimal demand for the risky claim depends only on myopic demand up to first order.

Using the pricing condition in Equation (19), we can express the risk premium in terms of the interest rate and the price of the aggregate claim:

$$\hat{\pi}_{Y,t} = -\hat{r}_t - d_p\hat{p}_{Y,t} + O(\epsilon^2),$$

where $d_p \equiv 1/p_Y^*$ is the benchmark dividend yield. Here $\hat{r}_t \equiv r_t - r^*$ and $\hat{p}_{Y,t} \equiv \frac{p_{Y,t}-p_Y^*}{p_Y^*}$ denote deviations from the benchmark.

Substituting into the expression for $q_{j,t}$ yields the first-order demand system:

$$q_{j,t} = -\zeta_{j,p}\,\hat{p}_{Y,t} - \zeta_{j,r}\,\hat{r}_t + f_{j,t} + O(\epsilon^2), \qquad j \in \mathcal{J}, \tag{20}$$

where, for unconstrained investors,

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

For passive investors and active investors whose leverage constraints bind, own- and cross-price elasticities vanish, $\zeta_{j,p} = \zeta_{j,r} = 0$. Their portfolio demand depends only on the demand shifters: $f_{j,t} = \hat{\sigma} \, \epsilon$ for constrained investors and $f_{j,t} = \hat{\alpha}_{p,t} \, \epsilon$ for passive investors.

Equation (20) shows that the first-order approximation to the dynamic heterogeneous-agent model in Section 3 provides a microfoundation 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. This reflects the fact that investors care about the risk premium, which depends on the price of the risky claim and the risk-free rate.

Importantly, the system is exposed to portfolio-flow shocks, that is, shocks that shift the demand of a subset of investors and induce portfolio reallocation. Shocks to $\overline{\alpha}_{p,t}$ move the passive investors' demand shifter $f_{p,t}$, the dynamic analog of the comparative statics in Section 2. In addition, the model generates endogenous demand shifts from occasionally binding leverage constraints.

Consumption-wealth ratio. A first-order expansion of (13) yields

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

Hence, the risk-free rate and the risk premium enter symmetrically, consistent with Assumption 1. Substituting $\hat{\pi}_{Y,t} = -\hat{r}_t - d_p \hat{p}_{Y,t}$ into the expression above yields

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

with coefficients

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

Conditional on the price-dividend ratio, the consumption-wealth ratio is therefore independent of the interest rate, as in Section 2. Its slope with respect to $p_{Y,t}$ depends on the EIS: the consumption-wealth ratio rises with $p_{Y,t}$ when $\psi > 1$, the standard case in macro-finance models. Most importantly, to first order the consumption-wealth ratio is unaffected by $\overline{\alpha}_{p,t}$. This property will be central for the determination of aggregate market elasticity.

Market clearing. Given the first-order demand system, equilibrium prices follow from market clearing in the risky claim and goods markets:

$$\sum_{j\in\mathcal{J}} x_{j,t}\, q_{j,t} = 0, \qquad \sum_{j\in\mathcal{J}} x_{j,t}\, \hat{c}_{j,t} = -d_p\, \hat{p}_{Y,t}.$$

Aggregating the demand system across investors yields

$$\begin{bmatrix} \zeta_{p,t} & \zeta_{r,t} \\ \chi_p + d_p & 0 \end{bmatrix} \begin{bmatrix} \hat{p}_{Y,t} \\ \hat{r}_t \end{bmatrix} = \begin{bmatrix} f_t \\ -\chi_{0,t} \end{bmatrix},$$

where $\zeta_{k,t} \equiv \sum_{j\in\mathcal{J}} x_{j,t}\zeta_{j,k}$ for $k \in \{p, r\}$, $f_t \equiv \sum_{j\in\mathcal{J}} x_{j,t} f_{j,t}$, and $\chi_{0,t} \equiv \sum_{j\in\mathcal{J}} x_{j,t}\chi_{j,0}$.

Aggregate market elasticity. As in Section 2, define the aggregate market elasticity of the aggregate claim as the inverse of its equilibrium price impact:

$$\varepsilon_{Y,t} \equiv \left[\frac{\partial \hat{p}_{Y,t}}{\partial f_t}\right]^{-1}.$$

The market is infinitely elastic when it can absorb portfolio flows without any price impact.

First-order impact of portfolio flows. We are now ready to characterize the first-order impact of portfolio flows. The second row of the demand system shows that the aggregate price-dividend ratio is pinned down by goods market clearing:

$$\left(\chi_p + d_p\right) \hat{p}_{Y,t} = -\chi_{0,t}.$$

Since $\chi_{0,t}$ is independent of the portfolio-flow shock f_t at first order, portfolio flows do not affect $\hat{p}_{Y,t}$. Instead, the risk premium and interest rate adjust to clear the risky-claim market. This result extends the well-known unit-EIS case, where $\chi_p = \chi_{0,t} = 0$ and the price-dividend ratio is constant. In our economy with heterogeneous investors and $\psi \neq 1$, the price-dividend ratio moves with changes to the average risk aversion in the economy, as captured by the coefficient $\chi_{0,t}$, but it is unaffected by portfolio flows. The next proposition formalizes this result.

Proposition 3 (First-order impact of portfolio flows)**.** *Consider a first-order approximation of the demand system. Then:*

(i) The price impact on the aggregate claim is zero:

$$\varepsilon_{Y,t}^{-1} = \frac{\partial \hat{p}_{Y,t}}{\partial f_t} = 0.$$

(ii) The interest rate and risk premium responses satisfy

$$\frac{\partial \hat{r}_t}{\partial f_t} = \frac{1}{\zeta_{r,t}}, \qquad \frac{\partial \hat{\pi}_{Y,t}}{\partial f_t} = -\frac{1}{\zeta_{r,t}}.$$

Proposition 3 shows that the main lesson from the simple model in Section 2 extends to the dynamic economy. Portfolio inflows into the risky claim reduce the risk premium, while the shift out of the risk-free asset raises the interest rate. To first order, these effects offset exactly, leaving the aggregate price-dividend ratio unchanged. The aggregate market is therefore infinitely elastic.

An important implication is that aggregate market elasticity is independent of both the own-price and cross-price elasticities of individual demand, $\zeta_{j,p}$ and $\zeta_{j,r}$. The aggregate cross-price elasticity, $\zeta_{r,t}$, determines the interest-rate and risk-premium responses required to clear the risky-claim market, but these responses offset in the valuation of the aggregate claim. Thus, individual investors can be made arbitrarily inelastic by increasing γ, while the aggregate market remains infinitely elastic.

Why is the market infinitely elastic? The disconnect between individual and aggregate market elasticity is inherently a general equilibrium phenomenon. Given a shift in demand for the risky claim, its market-clearing condition determines the required adjustment in expected excess returns. It does not, by itself, determine whether that adjustment occurs through the price-dividend ratio or the interest rate. For this, we need to know how the consumption-wealth ratio responds to the portfolio-flow shock.

As discussed in Section 2, there is no price impact 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 do not affect the consumption-wealth ratio to first order, differentiate (13) with respect to $\sigma_{j,t}$:

$$\frac{\partial c_{j,t}}{\partial \sigma_{j,t}} = (1-\psi)\,\gamma_j \left[\frac{\eta_t}{\gamma_j} + \frac{1-\gamma_j^{-1}}{\psi - 1}\,\sigma_{c_j,t} - \sigma_{j,t}\right].$$

At the benchmark, all investors hold their privately optimal risk exposure, that is, the exposure they would choose if allowed to optimize their portfolios freely. For passive investors, the prescribed portfolio coincides with this counterfactual choice, while the leverage constraint does not distort the low-risk-aversion investor's benchmark allocation. The term in brackets is therefore zero for every investor at the point around which we perturb.

By the envelope theorem, first-order deviations in risk exposure have only second-order effects

on the consumption-wealth ratio. Thus, the consumption-wealth ratio does not respond to portfolio flows at first order, and goods-market clearing leaves the price-dividend ratio unchanged. In the simple model of Section 2, this property was imposed as an *assumption*; here it emerges as a *result* of expanding around the efficient allocation.

This reasoning also suggests that an *inefficient* initial allocation of risk plays an important role in delivering finite aggregate market elasticity. Capturing these effects requires a second-order approximation of the demand system.

4.3 Determinants of aggregate market elasticity

We next turn to a second-order approximation of the demand system to study the determinants of aggregate market elasticity. To isolate the mechanism that generates inelastic markets more clearly, we consider the special case without preference heterogeneity or leverage constraints.

Second-order price response. Up to second order, the price-dividend ratio for the aggregate claim takes the form:

$$\hat{p}_{Y,t} = a_Y(x_t)\frac{(\overline{\alpha}_{p,t}-1)^2}{2} + b_Y(x_t)(\overline{\alpha}_{p,t}-1) + c_Y(x_t) + O(\epsilon^3), \tag{21}$$

where $a_Y(x)$, $b_Y(x)$, and $c_Y(x)$ are functions to be determined.

A passive flow changes $x_{p,t}(\overline{\alpha}_{p,t}-1)$. Thus, the inverse aggregate market elasticity defined above is given by:[16]

$$\varepsilon_{Y,t}^{-1} \equiv \frac{1}{x_{p,t}}\frac{\partial \hat{p}_{Y,t}}{\partial \overline{\alpha}_{p,t}} = \frac{a_Y(x_t)}{x_{p,t}}(\overline{\alpha}_{p,t}-1) + \frac{b_Y(x_t)}{x_{p,t}} + O(\epsilon^2).$$

Hence, characterizing aggregate market elasticity amounts to determining $a_Y(x_t)$ and $b_Y(x_t)$.

Risk misallocation and consumption. To understand why the second-order allocation of risk matters for prices, consider its effect on consumption-saving decisions. Aggregating Equation (13) across investors in the case without preference heterogeneity and using market clearing gives

$$c_t = \psi\rho + (1-\psi)\underbrace{\left[r_t + \pi_{Y,t} - \frac{\gamma}{2}\|\sigma_{R_Y,t}\|^2 \sum_{j\in\mathcal{J}} x_{j,t}(1+q_{j,t})^2\right]}_{\text{aggregate risk-adjusted expected return}} + \xi_t + O(\epsilon^3),$$

[16] In this definition, we consider the effect of a change in $\overline{\alpha}_{p,t}$ holding $x_{p,t}$ fixed. Appendix E.2 shows that the endogenous response of $x_{p,t}$ to a portfolio-flow shock affects prices only at higher order. Hence, it does not contribute to aggregate market elasticity up to second order.

where $c_t \equiv \sum_{j\in\mathcal{J}} x_{j,t} c_{j,t}$, and we used the fact that $\sigma_{j,t} = (1 + q_{j,t})\sigma + O(\epsilon^2)$. The expression in brackets captures the current return to saving, adjusted for the risk borne by investors.

Because risky-claim market clearing implies $\sum_{j\in\mathcal{J}} x_{j,t} q_{j,t} = 0$,

$$\sum_{j\in\mathcal{J}} x_{j,t}(1 + q_{j,t})^2 = 1 + \sum_{j\in\mathcal{J}} x_{j,t} q_{j,t}^2.$$

Thus, the first-order effects of reallocating risk cancel, but the second-order effects do not. At the frictionless allocation, all investors hold the aggregate risk exposure, so $q_{j,t} = 0$ for every j. We therefore interpret $\sum_{j\in\mathcal{J}} x_{j,t} q_{j,t}^2$ as a measure of *risk misallocation*.

A passive flow changes this dispersion by shifting risk between passive and active investors. Without preference heterogeneity or leverage constraints,

$$q_{p,t} = \overline{\alpha}_{p,t} - 1, \qquad q_{j,t} = -\frac{x_{p,t}}{x_{a,t}}(\overline{\alpha}_{p,t} - 1), \quad j \in \{h, \ell\},$$

where $x_{a,t} = 1 - x_{p,t}$ is the wealth share of active investors. Consequently,

$$\sum_{j\in\mathcal{J}} x_{j,t} q_{j,t}^2 = \frac{x_{p,t}}{x_{a,t}}(\overline{\alpha}_{p,t} - 1)^2 + O(\epsilon^3).$$

When $\psi > 1$, an increase in risk misallocation reduces the risk-adjusted return and weakens the incentive to save, raising the aggregate consumption-wealth ratio.

The second-order expansion of the consumption-wealth ratio $\hat{c}_t \equiv c_t - c^*$ is given by

$$\hat{c}_t = \chi_{0,t} + \chi_p \hat{p}_{Y,t} + O(\epsilon^3),$$

where $\chi_p = (\psi - 1) d_p$ and $\chi_{0,t}$ is given by

$$\chi_{0,t} = (\psi - 1)\frac{\gamma\|\sigma\|^2}{2}\frac{x_{p,t}}{x_{a,t}}(\overline{\alpha}_{p,t} - 1)^2 - \psi\left[\mu_{p_Y,t} + (1 - \gamma)\sigma_{p_Y,t}\sigma^\top\right].$$

The first term captures risk misallocation, while the second captures the endogenous drift and volatility of the price-dividend ratio. Both terms respond to portfolio flows.

Thus, at second order, portfolio-flow shocks shift not only demand for the risky claim but also the consumption-wealth ratio schedule. This shift reflects both changes in risk misallocation and the endogenous dynamics of the price-dividend ratio.

The aggregate market elasticity. As in the first-order analysis, risky-claim market clearing determines the risk-premium response, while goods-market clearing determines the price-dividend

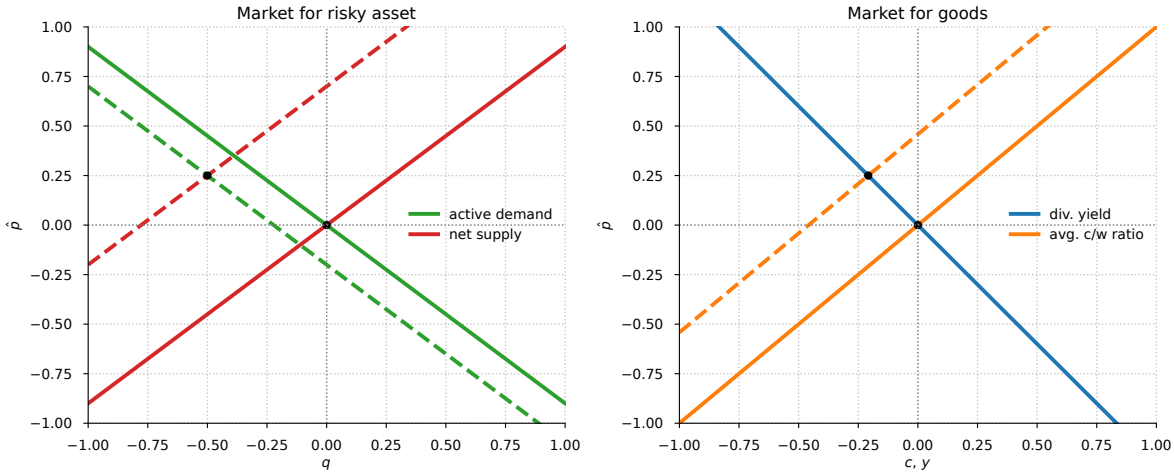


Figure 3. Equilibrium with inefficient risk allocation: risky-claim (left) and goods (right) markets.

ratio. Goods-market clearing therefore requires the price-dividend ratio to satisfy

$$\chi_{0,t} + \chi_p \hat{p}_{Y,t} = -d_p \hat{p}_{Y,t} + O(\epsilon^3) \Rightarrow \hat{p}_{Y,t} = -\frac{\chi_{0,t}}{\chi_p + d_p} + O(\epsilon^3).$$

Thus, price impact is governed by how portfolio flows affect consumption-saving decisions, not directly by the individual demand elasticities $\zeta_{j,p}$ and $\zeta_{j,r}$. The expression above is the analog of the price response to a fundamental shock in the simple model of Section 2.

Using the quadratic expansion in Equation (21) to evaluate $\mu_{p_Y,t}$ and $\sigma_{p_Y,t}$, and matching coefficients in goods-market clearing, yields the following result.

Proposition 4 (Aggregate elasticity under passive demand)**.** *Let $x_{a,t} = 1 - x_{p,t}$ denote the wealth share of active investors. Without preference heterogeneity or leverage constraints, the inverse aggregate market elasticity is given by:*

$$\varepsilon_{Y,t}^{-1} = \frac{1-\psi^{-1}}{d_p + 2\kappa_p} \frac{\gamma \|\sigma\|^2}{x_{a,t}} \left[1 - \overline{\alpha}_{p,t} + \frac{\gamma - 1}{d_p + \kappa_p} \sigma_{\alpha_p} \sigma^\top \right] + O\left(\epsilon^2\right). \tag{22}$$

Proposition 4 provides our main analytical result, a characterization of the aggregate market elasticity in an economy subject to portfolio-flow shocks. Several important implications follow from Equation (22).

First, the market is infinitely elastic in the unit EIS case, $\psi = 1$. As is well known, the price-dividend ratio is essentially pinned down by the subjective discount rate in this case, so movements in risk premia are offset by opposite movements in the interest rate. The consumption-wealth ratio is then not affected by portfolio flows, causing the aggregate market elasticity to be infinite.

Second, in the case $\sigma_{\alpha_p} = 0$, aggregate market elasticity is finite as long as the initial allocation of risk is inefficient, that is, $\overline{\alpha}_{p,t} \neq 1$. As in the analysis of a fundamental shock in Section 2,

portfolio-flow shocks simultaneously shift demand for the risky claim and the consumption–wealth ratio schedule, generating positive price impact. Figure 3 illustrates the mechanism. If initially $\overline{\alpha}_{p,t} < 1$, a portfolio inflow into the risky claim reduces risk misallocation. When $\psi > 1$, this increases the incentive to save, reducing the consumption–wealth ratio.[17] This raises the price–dividend ratio, as shown in the right panel of Figure 3. In equilibrium, the interest rate only partially offsets the risk-premium response, as shown in the left panel of Figure 3.

Third, when $\gamma > 1$, price impact is positive even if the initial allocation of risk is efficient, $\overline{\alpha}_{p,t} = 1$, when portfolio-flow shocks are positively correlated with fundamental shocks, $\sigma_{\alpha_p}\sigma^\top > 0$. The mechanism again operates through the consumption-wealth ratio. The equilibrium price impact generates endogenous return volatility:

$$\sigma_{R_Y,t} = \sigma + x_{p,t}\varepsilon_{Y,t}^{-1}\sigma_{\alpha_p} + O(\epsilon^3).$$

Thus, correlated portfolio-flow shocks change the endogenous volatility of the risky-claim return. When $\psi > 1$ and $\gamma > 1$, a portfolio inflow from the efficient allocation reduces this volatility contribution, lowers the consumption–wealth ratio, and raises the price–dividend ratio.[18]

Fourth, the price impact is state-dependent. In a numerical illustration, Figure 4 shows that the price impact is larger when the active wealth share is low, reflecting their limited risk-bearing capacity. Hence, the market is endogenously more inelastic in bad times, when active investors are undercapitalized. Moreover, a reduction in active wealth share raises the price impact by more when passive investors are further from the efficient allocation. For simplicity, we focus on the case $\sigma_{\alpha_p} = 0$, so there is no price impact when the passive portfolio is initially efficient.

Finally, market elasticity depends on the persistence of portfolio-flow shocks, governed by κ_p. Faster mean reversion reduces price impact, and in the limit in which flows are transitory, $\kappa_p \to \infty$, the market becomes infinitely elastic. This comparative static echoes Gabaix and Koijen (2023), who show that faster mean reversion in flows lowers the multiplier. Binsbergen et al. (2025) likewise show that persistence is central to both the identification of demand elasticities and the dynamic interventions required to sustain a given price path. The common lesson is that persistence is an essential part of the mapping from quantities to prices. These effects arise from equilibrium reallocations of aggregate risk over time rather than from the short-horizon trading frictions studied in market microstructure. In this respect, our mechanism is closer to the literature on *market macrostructure* (see, e.g., Haddad and Muir, 2025).

[17] This mirrors the mechanism in representative-agent models in which an increase in uncertainty reduces risky asset prices only when $\psi > 1$ (see, e.g., Bansal and Yaron, 2004).

[18] In the case $\sigma_{\alpha_p}\sigma^\top = 0$, the impact of σ_{α_p} on return variance is negligible up to second order, as flow risk is essentially diversifiable. The logic is reminiscent of why zero-beta assets have zero excess returns in the CAPM.

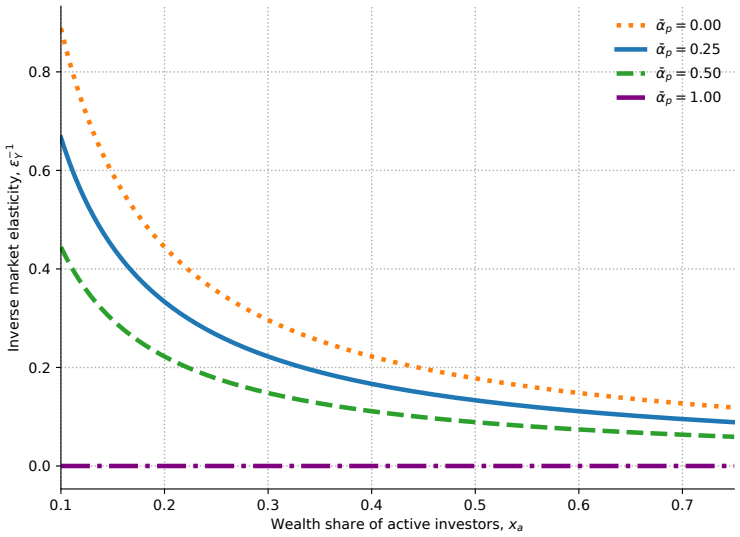


Figure 4. Price impact: inefficient passive demand.

Risk premium and interest rate. Having characterized the price response to portfolio flows, we now describe the corresponding adjustments in the market price of risk, the risk premium, and the interest rate.

Proposition 5 (Market price of risk, risk premium, and interest rate)**.** *Consider the case without preference heterogeneity or leverage constraints. Then:*

1. *The market price of risk is given by*

$$\eta_t = \gamma \chi_{a,t} \sigma + \left[1 - H_{a,t}\right] \gamma x_{p,t} \varepsilon_{Y,t}^{-1} \sigma_{\alpha_p} + O(\epsilon^3), \tag{23}$$

where

$$\chi_{a,t} \equiv \frac{1 - x_{p,t}\overline{\alpha}_{p,t}}{1 - x_{p,t}}, \qquad H_{a,t} \equiv \frac{1 - \gamma^{-1}}{\psi - 1}\left[\frac{\psi}{x_{a,t}} - 1\right].$$

2. *The risk premium is given by*

$$\pi_{Y,t} = \eta_t \sigma^\top + \gamma x_{p,t} \varepsilon_{Y,t}^{-1} \sigma_{\alpha_p} \sigma^\top + O(\epsilon^3). \tag{24}$$

3. *The interest rate is given by*

$$r_t = r^* + (1 - \psi^{-1}) \frac{\gamma \|\sigma\|^2}{2} \frac{x_{p,t}}{x_{a,t}} (\overline{\alpha}_{p,t} - 1)^2 - (\eta_t - \eta^*) \sigma^\top + O(\epsilon^3).$$

Equation (23) shows that the market price of risk depends on two terms. The first captures the net supply of risk absorbed by active investors, summarized by $\chi_{a,t}\sigma$. When passive investors

hold more of the risky claim, less aggregate risk must be absorbed by active investors, reducing the market price of risk.

The second term captures the effect of endogenous volatility, $x_{p,t}\varepsilon_{Y,t}^{-1}\sigma_{\alpha_p}$. The magnitude and sign of this effect are modified by the hedging demand of active investors, summarized by the coefficient $H_{a,t}$.

The risk premium satisfies $\pi_{Y,t} = \eta_t \sigma_{R_Y,t}^{\top}$. In Equation (24), the first term prices the fundamental exposure σ, while the second arises from the endogenous return exposure generated by portfolio flows. Portfolio-flow shocks therefore contribute directly to the risk premium on the aggregate claim when they covary with fundamental risk, $\sigma_{\alpha_p}\sigma^{\top} \neq 0$.

The risk premium incorporates compensation for portfolio-flow risk, in line with noise-trader models (see, e.g., De Long, Shleifer, Summers, and Waldmann, 1990) and the preferred-habitat framework of Vayanos and Vila (2021). Unlike these frameworks, however, our model explicitly incorporates wealth effects in portfolio choice, so the wealth distribution plays a central role in determining risk premia.

The interest rate offsets the change in the market price of fundamental risk through the term $-(\eta_t - \eta^*)\sigma^{\top}$. Risk misallocation, however, contributes an additional term by changing consumption-saving incentives. The interest-rate adjustment therefore does not perfectly offset the risk-premium response, consistent with the general equilibrium mechanism characterized in Proposition 4.

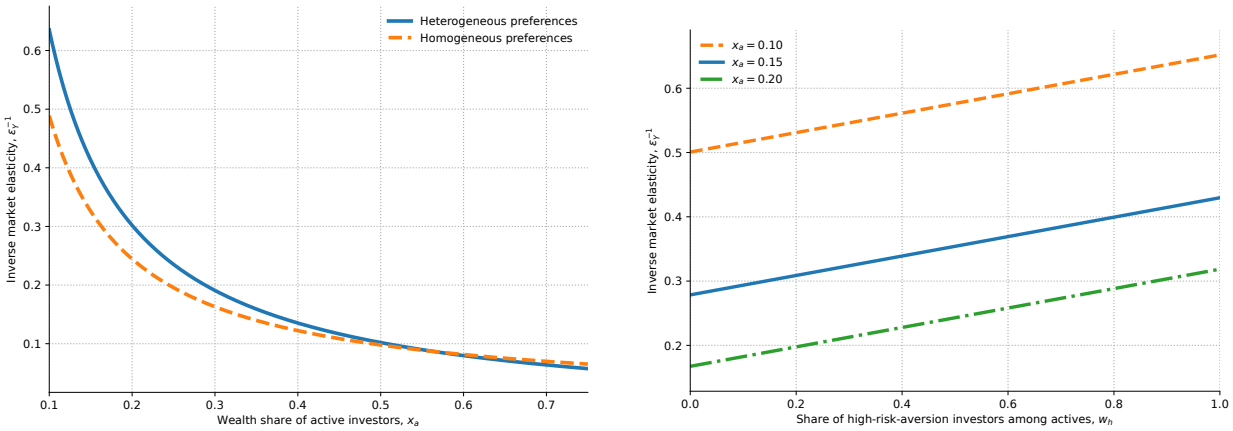
4.4 Amplification mechanisms

We now enrich the analysis along two dimensions. Preference heterogeneity and leverage constraints affect aggregate price impact by changing the risk-bearing capacity of marginal investors. The presence of multiple claims introduces an additional distinction between the price impact on the aggregate claim and on equity.

Preference heterogeneity and leverage constraints. The next proposition shows how preference heterogeneity and binding leverage constraints affect aggregate market elasticity.

Proposition 6 (Aggregate market elasticity with heterogeneity and leverage constraints)**.** *Suppose* $\overline{\alpha} = 1 + \hat{\alpha}\epsilon$ *and* $\gamma_j = \gamma(1 + \hat{\gamma}_j\epsilon)$ *for each* $j \in \mathcal{J}$*. In the regime in which the leverage constraint binds, suppose also that* $\overline{\sigma} = \|\sigma\|(1 + \hat{\sigma}\epsilon)$*. Then, the inverse aggregate market elasticity is given by*

$$\varepsilon_{Y,t}^{-1} = \left(1 - \psi^{-1}\right) \frac{\gamma \|\sigma\|^2}{d_p + 2\kappa_p} \frac{1 - x_{c,t}}{x_{u,t}} \left[\overline{\alpha}_{p,t}^{opt} - \overline{\alpha}_{p,t} + \frac{\kappa_p(\overline{\alpha}_{p,t}^{opt} - \overline{\alpha})}{d_p + \kappa_p} + \frac{(\gamma - 1)\sigma_{\alpha_p}\sigma^{\top}}{d_p + \kappa_p}\right] + O\left(\epsilon^2\right), \tag{25}$$



(a) Homogeneous vs. heterogeneous preferences

(b) Role of wealth distribution of active investors

Figure 5. 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 for different passive portfolio shares. Leverage constraints are slack.

where $x_{c,t}$ is the wealth share of active investors whose constraint binds and $x_{u,t} \equiv x_{a,t} - x_{c,t}$ is the wealth share of unconstrained active investors. The term $\overline{\alpha}_{p,t}^{\text{opt}}$ is the passive investors' optimal portfolio share, given by

$$\overline{\alpha}_{p,t}^{\text{opt}} = 1 - \frac{x_{u,t}}{1 - x_{c,t}} \frac{\gamma_p - \mathbb{E}_t^u[\gamma_j]}{\gamma} - \frac{x_{c,t}}{1 - x_{c,t}} \left(\frac{\overline{\sigma}}{\|\sigma\|} - 1 \right),$$

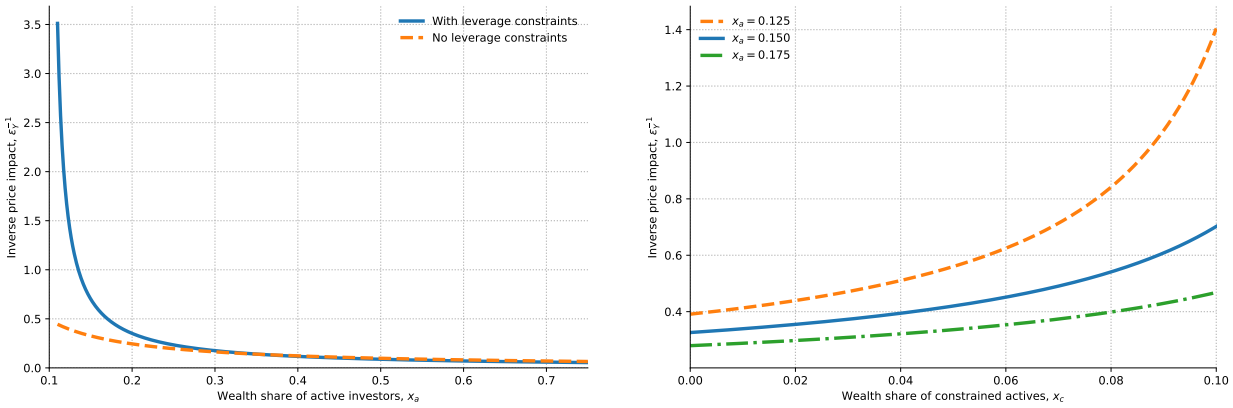
with $\mathbb{E}_t^u[\gamma_j]$ denoting the wealth-weighted average risk aversion of unconstrained active investors.

Proposition 6 characterizes aggregate market elasticity in this richer environment. The price impact now depends on the difference between the optimal and actual passive portfolio shares, $\overline{\alpha}_{p,t}^{\text{opt}} - \overline{\alpha}_{p,t}$. The optimal passive share is the portfolio share that passive investors would choose if they were allowed to optimize their portfolios freely. In the absence of preference heterogeneity and binding leverage constraints, the optimal passive share is $\overline{\alpha}_{p,t}^{\text{opt}} = 1$. We then recover Proposition 4 after imposing $\overline{\alpha} = 1$.

Preference heterogeneity and leverage constraints introduce additional state dependence through the distribution of wealth and which constraints bind. Consider first the case in which leverage constraints are slack and passive investors are more risk averse than active investors, $\gamma_p > \mathbb{E}_t^u[\gamma_j]$. Holding the composition of active investors fixed, the optimal passive share rises when active investors become less capitalized:

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

Under these conditions, shocks that reduce active wealth make the market endogenously more



(a) Role of leverage constraints *(assumes* $x_{\ell,t} = 0.10$*)*

(b) Varying the wealth share of constrained actives, $x_{c,t}$

Figure 6. Price impact under binding leverage constraints. Panel (a) compares binding and slack leverage constraints while holding $x_{\ell,t} = 0.10$; hence, $x_{c,t} = x_{\ell,t}$ in the binding regime and $x_{c,t} = 0$ in the slack regime. Panel (b) plots the price impact as a function of $x_{c,t}$ for different values of $x_{a,t}$.

inelastic. Active investors' risk-bearing capacity declines, while the gap between the optimal and actual passive portfolio shares widens. Together, these forces generate a sharp increase in price impact as $x_{a,t}$ declines, as shown in the left panel of Figure 5.

The distribution of wealth among active investors also matters. An increase in the share of high risk-aversion active investors, $w_{h,t} \equiv x_{h,t}/x_{a,t}$, raises the average risk aversion of active investors, $\mathbb{E}_t^u[\gamma_j]$, reducing aggregate risk-bearing capacity and making the market more inelastic, as shown in the right panel of Figure 5.

When leverage constraints bind, portfolio-flow shocks must be absorbed by a smaller set of unconstrained investors. Their remaining risk-bearing capacity becomes a key determinant of price impact. The left panel of Figure 6 plots price impact against the active wealth share $x_{a,t} = 1 - x_{p,t}$. To isolate the role of constraints, we set $\gamma_p = \mathbb{E}_t^u[\gamma_j]$, removing pure preference heterogeneity effects. As $x_{a,t}$ declines, price impact rises more sharply under binding constraints. The right panel holds $x_{a,t}$ fixed and shows that price impact rises with the constrained share $x_{c,t}$.

Price impact of the equity claim. So far, we have focused on how portfolio flows affect the price of the aggregate claim. Empirically, however, the main object of interest is the equity market. Because dividends are more exposed to fundamental risk than aggregate output, the general equilibrium offset between the interest rate and the aggregate risk premium does not extend to equity. The next proposition characterizes the resulting equity price impact.

Proposition 7 (Market elasticity of the equity claim)**.** *Define the inverse elasticity of the equity*

claim by $\varepsilon_{D,t}^{-1} \equiv \frac{1}{x_{p,t}} \frac{\partial \log p_{D,t}}{\partial \overline{\alpha}_{p,t}}$. *Then,*

$$\varepsilon_{D,t}^{-1} = \frac{\phi - 1}{d_D + \kappa_p} \zeta_{r,t}^{-1} + O(\epsilon),$$

where $\zeta_{r,t}^{-1}$ *is the first-order risk-premium response defined in Proposition 3.*

Proposition 7 shows that equity price impact arises at first order, even though the price impact on the aggregate claim is zero at first order. A portfolio inflow raises the interest rate and lowers the risk premium on the aggregate claim by the same amount, leaving its valuation ratio unchanged. Equity cash flows load on fundamental risk with exposure $\phi\sigma$, so the decline in the equity premium is ϕ times as large as the decline in the aggregate risk premium. The interest-rate response therefore offsets only part of the equity-premium response when $\phi > 1$.

The magnitude of equity price impact increases with cash-flow leverage, ϕ, and with the inverse cross-elasticity, $\zeta_{r,t}^{-1}$. It decreases with κ_p, because transitory portfolio flows have smaller effects on equity valuations. When $\phi = 1$, equity is proportional to the aggregate claim and the first-order equity price impact disappears.

Taking stock. This section provides a detailed characterization of market elasticity in a general equilibrium setting with frictions. For the aggregate claim, portfolio flows affect prices only when they alter consumption–saving decisions. This occurs when risk is initially misallocated or when the economy is systematically exposed to portfolio-flow shocks. Preference heterogeneity and leverage constraints further amplify this price impact in bad times by reducing effective risk-bearing capacity. Equity exhibits an additional source of price impact because its cash flows are more exposed to fundamental risk than aggregate output. Cash-flow leverage breaks the general equilibrium offset between the interest rate and the aggregate risk premium, generating equity price impact already at first order. We next examine the robustness of the general equilibrium mechanism before turning to its quantitative implications.

5 Robustness of the General Equilibrium Mechanism

The baseline mechanism operates through two margins: household consumption responds to flow-induced changes in wealth, and the interest rate adjusts to clear the goods market. This section examines the robustness of the mechanism along each of these margins. First, we introduce wealthy hand-to-mouth investors whose consumption does not respond to wealth and show that, under a natural benchmark, equilibrium prices and normalized allocations remain unchanged even when the fraction of optimizing investors is arbitrarily small. Second, we introduce production, sticky prices,

and monetary policy and show that an inflation-targeting central bank implements the endogenous interest-rate adjustment of the baseline model.

5.1 Wealthy hand-to-mouth consumers

The analysis so far assumes that households choose consumption optimally. In particular, their consumption responds to changes in wealth, so asset-price movements affect goods demand. This wealth-effect channel is consistent with evidence that stock-market wealth has economically significant effects on consumption and real activity (Chodorow-Reich et al., 2021).

Not all asset holders, however, adjust consumption in response to changes in the market value of their wealth. A substantial fraction of households hold sizable illiquid assets while consuming their current income, a behavior commonly described as *wealthy hand-to-mouth* (Kaplan and Violante, 2014). Moreover, the marginal propensity to consume out of capital gains is small relative to the marginal propensity to consume out of dividend income (Di Maggio et al., 2020). To capture this behavior, Appendix B introduces investors who own a fixed fraction of aggregate assets and consume their current payouts rather than choosing consumption optimally.

Perhaps surprisingly, introducing these non-optimizing investors does not attenuate the equilibrium effects of portfolio flows. The appendix establishes an exact *neutrality result*. For any path of the passive portfolio share, introducing wealthy hand-to-mouth investors leaves equilibrium prices, interest rates, risk premia, and the normalized allocations of all other investors unchanged. The general equilibrium mechanism therefore does not require all investors to choose consumption optimally. It remains operative even when the fraction of optimizing investors is arbitrarily small.

Why does the equilibrium response remain unchanged as the hand-to-mouth sector grows? Two opposing forces are at work. Because fewer households adjust consumption when wealth changes, the direct wealth effect becomes weaker. At the same time, the goods-market adjustment is concentrated among a smaller optimizing sector, amplifying its consumption response in exactly the same proportion. When the hand-to-mouth sector consumes the payouts on its asset holdings, these two forces exactly cancel.[19] This logic parallels neutrality results in heterogeneous-agent New Keynesian models (Werning, 2015; Bilbiie, 2008).[20]

5.2 Monetary policy

Our characterization of aggregate market elasticity uses an endowment economy in which the interest rate adjusts endogenously to clear the goods market. Because short-term interest rates are

[19] Appendix B uses the simple model of Section 2 to characterize departures from this neutrality benchmark when the hand-to-mouth sector's consumption share differs from its wealth share.

[20] For recent evidence that this neutrality benchmark provides a good empirical approximation, see, e.g., Bilbiie, Galaasen, Gürkaynak, Mæhlum, and Molnar (2025).

controlled by central banks, one might instead expect the constant-interest-rate benchmark to be relevant. More generally, one might expect policy rates to be governed by inflation dynamics and therefore independent of equity portfolio flows. Does central-bank control eliminate the endogenous interest-rate adjustment at the center of our mechanism?

Appendix C addresses this question by introducing production, sticky prices, and monetary policy into the model of Section 3. The flexible-price equilibrium is equivalent to the baseline endowment economy, with the equilibrium interest rate becoming the *natural rate of interest*. The natural rate therefore coincides with the equilibrium interest rate in the baseline model, and aggregate market elasticity is unchanged.

With sticky prices, a central bank pursuing price stability implements the flexible-price allocation by moving the policy rate with the natural rate. An inflation-targeting central bank therefore responds to persistent portfolio flows in exactly the same way as the interest rate in the baseline model. Central-bank control of the short-term rate does not make the interest rate independent of financial conditions.

To understand this result, consider a persistent portfolio outflow from equities into the risk-free asset. If the central bank held the real interest rate fixed, the resulting decline in equity prices would reduce household wealth and consumption. The contraction in aggregate demand would create an output gap and deflationary pressure. Maintaining price stability therefore requires the central bank to cut the policy rate to the natural rate implied by the baseline model. This mechanism is consistent with evidence that policymakers respond to stock-market declines in large part because of their potential effects on consumption (Cieslak and Vissing-Jorgensen, 2021).[21]

An important implication is that even a central bank concerned only with inflation has an incentive to track financial conditions and respond to asset-price movements induced by portfolio flows. This implication connects our analysis to recent work on financial-conditions targeting (Caballero et al., 2024; Caballero and Simsek, 2026). In that work, stabilizing financial conditions can be desirable beyond their direct effects on output and inflation because policy also affects private risk-bearing capacity. We show that, even without this additional motive, price stability itself requires monetary policy to respond when flow-induced asset-price movements affect aggregate demand through household wealth.

6 Quantitative Implications

We now evaluate the quantitative implications of the dynamic model of Section 3. We first describe the calibration and evaluate the model's implications for sectoral balance sheets, portfolio

[21] For discussions of monetary models with wealth effects from asset-price revaluations, see, e.g., Caballero and Simsek (2020) and Caramp and Silva (2026).

flows, aggregate price impact, and standard asset pricing moments. The baseline model generates price impact for the dividend claim consistent with the lower end of the empirical estimates. We then extend the model by allowing non-fundamental portfolio-flow shocks to affect investors' expectations of future cash-flow growth, consistent with recent evidence on expectations. This belief-distortion mechanism further amplifies price impact. Finally, we examine the model's implications for state-dependent price impact and return predictability.

We solve the full dynamic system numerically using deep neural networks to approximate the equilibrium functions, as in Duarte et al. (2024) and Gopalakrishna and Wu (2026). Appendix G.3 describes the numerical procedure in detail.

6.1 Calibration

We calibrate the model using U.S. data on aggregate cash flows, sectoral portfolio holdings, and portfolio flows. The calibration has three blocks: technology, preferences and the cross section of investors, and passive demand. Table 1 reports the resulting parameter values. We evaluate the resulting model-implied price impact against the empirical range discussed below.

Technology. We set the endowment growth rate to $\mu = 1.8\%$ and its volatility to $\|\sigma\| = 3.0\%$, consistent with U.S. aggregate consumption data (Campbell and Cochrane, 1999). We set the cash-flow leverage parameter to $\phi = 4.6$, implying annual dividend-growth volatility of 13.8%, matching the full-sample CRSP estimate in Atkeson, Heathcote, and Perri (2026).

Preferences and the cross section of investors. We aggregate Flow of Funds (FoF) sectors into three ultimate sectors: a passive sector, a high-risk-aversion active sector, and a low-risk-aversion active sector.[22] We set a common elasticity of intertemporal substitution $\psi = 1.5$, in line with Bansal and Yaron (2004), and set the subjective discount rate to $\widehat{\rho} = 0.001$. The risk-aversion coefficients are $\gamma_p = 12.4$, $\gamma_h = 25.0$, and $\gamma_\ell = 4.8$. We jointly choose the sector masses, risk-aversion coefficients, and leverage-constraint tightness to match sectoral asset shares and asset-to-equity ratios in the FoF data. The resulting post-burn-in wealth-weighted aggregate risk aversion is approximately 8.1.

We set the agent turnover rate to $\kappa = 0.0145$ so that the simulated wealth distribution is close to its empirical counterpart. Together, $\widehat{\rho}$ and κ imply an effective discount rate of $\rho = \widehat{\rho} + \kappa = 0.0155$. The low-risk-aversion active investors are subject to the occasionally binding leverage constraint. Its tightness is $\overline{\sigma} = 0.072$, chosen to match the asset-to-equity ratio of their sector.

[22] We follow the aggregation in Koijen, Richmond, and Yogo (2024). Appendix G.1 describes the sector mapping and the treatment of pass-through vehicles.

	Parameter	Choice
	Technology	
μ	Endowment growth rate	0.018
$\|\|\sigma\|\|$	Endowment volatility	0.030
ϕ	Cash-flow leverage parameter	4.60
	Preferences & distribution	
ψ	EIS	1.50
ρ	Effective discount rate	0.0155
κ	Agent turnover rate	0.0145
γ_p	Passive-sector risk aversion	12.37
γ_h	High active-sector risk aversion	25.00
γ_ℓ	Low active-sector risk aversion	4.80
ω_p	Share of passive agents	0.65
ω_h	Share of high-risk-aversion active agents	0.30
ω_ℓ	Share of low-risk-aversion active agents	0.05
$\overline{\sigma}$	Tightness of the leverage constraint	0.072
	Passive demand	
$\overline{\alpha}$	Mean	0.52
κ_p	Mean reversion parameter	0.10
$\|\|\sigma_{\alpha_p}\|\|$	Passive-demand volatility	0.032
ν_p	Corr. of passive demand and fundamental shock	0.08

Table 1. Parameter values.
This table reports the parameter values used to calibrate the model. Technology parameters are disciplined by aggregate consumption and CRSP dividend growth. Risk-aversion parameters and sectoral masses are chosen jointly to match sectoral asset shares and asset-to-equity ratios. The leverage-constraint tightness targets the asset-to-equity ratio of the low-risk-aversion active sector. Passive-demand parameters are disciplined by the average passive portfolio share, household portfolio inertia, the volatility of quarterly aggregate equity flows, and the correlation between annual changes in the passive portfolio share and annual dividend growth.

Passive demand. The long-run portfolio share of passive investors $\overline{\alpha} = 0.52$ is the household-weighted average implied by our FoF aggregation. This value lies between the one-third estimate of Chinco and Sammon (2024), inferred from index-reconstitution events, and the roughly two-thirds estimate of Koijen et al. (2024), based on active share measures in the spirit of Cremers and Petajisto (2009).

We discipline portfolio-share volatility $\|\sigma_{\alpha_p}\|$ using the volatility of quarterly aggregate equity flows. We construct the flow series from FoF sectoral holdings, scale each sector's dollar equity flow by lagged aggregate market capitalization, remove mechanical revaluation, and aggregate gross flows across sectors, dividing by two to avoid double counting. This procedure replicates the quarterly flow series in Figure D.7 of Gabaix and Koijen (2023). Appendix G.2 describes the construction and time-series properties of the flow series.

The persistence parameter κ_p is informed by evidence on household portfolio inertia. Using

PSID data, Brunnermeier and Nagel (2008) estimate portfolio-inertia coefficients that imply $\kappa_p \in [0.06, 0.14]$. We set $\kappa_p = 0.10$, the midpoint of this range.

Finally, $\nu_p \equiv \sigma\sigma_{\alpha_p}^{\top}/(\|\sigma\|\,\|\sigma_{\alpha_p}\|)$ is the correlation between passive-demand innovations and the aggregate fundamental shock. We set $\nu_p = 0.08$, matching the contemporaneous correlation between annual changes in the passive portfolio share implied by our FoF aggregation and annual real dividend growth from CRSP.

6.2 Baseline results

We next examine the unconditional implications of the calibrated economy. We simulate 16 independent monthly paths of 1,400 years each, discard the first 25% of every path, and report the resulting moments in Table 2. These include standard asset pricing moments for the dividend and endowment claims, sectoral balance sheets, and the price impact of both claims. The risk premium, return volatility, and price impact are not direct calibration targets.

Price impact regression. To estimate the model-implied price impact of the dividend claim, we regress the change in the log price-dividend ratio on the innovation to passive demand, scaled by the passive sector's lagged wealth share:

$$\Delta \log(P/D)_t = a_0 + b_1 f_t + e_t, \qquad f_t \equiv \overline{\alpha}^{s}_{p,t} x_{p,t-1}, \tag{26}$$

where $\overline{\alpha}^{s}_{p,t}$ denotes the innovation to passive demand. Because passive demand innovations may be correlated with fundamental shocks, we instrument f_t with its component orthogonal to the aggregate fundamental shock. The resulting coefficient isolates the response of prices to non-fundamental portfolio flows.

Equation (26) is the model counterpart of the price-impact regression in Gabaix and Koijen (2023). Scaling by the lagged passive wealth share $x_{p,t-1}$ converts the per-unit-mass shock into an aggregate flow. Using the orthogonal flow component as an instrument parallels the GIV strategy of isolating variation in aggregate flows that is unrelated to common fundamentals.[23] We estimate the price impact of the endowment claim analogously.

Empirical evidence on price impact. Empirical estimates of aggregate equity price impact span a meaningful range. Using predictable flows generated by dividend payouts in the U.S., Hartzmark and Solomon (2025) estimate price impacts between 1.5 and 2.3. Using coordinated

[23] Gabaix and Koijen (2023), following Gabaix and Koijen (2024), remove observed and latent common factors from sector-level flows and construct a granular instrument from the remaining idiosyncratic variation. They use this instrument to address the simultaneity of prices and aggregate flows.

Moment	Model	Data
Panel A. Dividend-claim moments		
Risk premium (%)	5.9	6.3
Risk-free rate (%)	1.8	0.9
Return volatility (%)	19.0	19.4
Quarterly gross-flow volatility (%)	0.5	0.5
Price impact	2.2	1.5–5.9
Panel B. Endowment-claim moments		
Risk premium (%)	1.1	—
Return volatility (%)	3.6	—
Price impact	0.1	—
Panel C. Sector balance-sheet moments		
Assets/Equity (high-risk-aversion active)	0.2	0.5
Assets/Equity (low-risk-aversion active)	1.9	1.6
Asset share (passive, %)	42.5	44.0
Asset share (high-risk-aversion active, %)	17.8	22.8
Asset share (low-risk-aversion active, %)	39.7	33.1

Table 2. Unconditional moments.
This table reports model-implied and empirical unconditional moments for the dividend and endowment claims, together with sectoral balance-sheet moments, in the baseline economy. The empirical moments for the risk premium, risk-free rate, and return volatility are from Bansal and Yaron (2004). Model moments pool 16 independent monthly simulations of 1,400 years each, after discarding the first 25% of every path as burn-in. Model price impact is estimated by instrumenting aggregate portfolio flows with their component orthogonal to the fundamental shock. Asset/Equity denotes the ratio of risky asset holdings to equity for the high- and low-risk-aversion active sectors. Empirical asset shares are normalized by the combined holdings of the three modeled sectors and therefore sum to one. Quarterly gross-flow volatility is the standard deviation of the gross reallocation measure constructed by summing absolute sectoral flows and dividing by two. Return volatility is the mean instantaneous diffusion volatility of returns across simulated states. Returns and interest rates are annualized; portfolio flows are quarterly.

reallocations by Chilean pension fund investors, the estimates of Da et al. (2018) imply a price impact of approximately 2.2. Across alternative GIV specifications, Gabaix and Koijen (2023) obtain estimates between 4.7 and 5.9. These exercises isolate different sources of portfolio flows, but together suggest an empirical range of approximately 1.5 to 5.9. Table 2 reports this range rather than selecting a single empirical target.

Price impact and asset pricing moments. The baseline economy delivers a price impact of approximately 2.2, in line with the estimate of Da et al. (2018) and near the lower end of the empirical range. Thus, heterogeneity, risk misallocation, and leverage constraints generate substantial price impact, although they do not by themselves reach the upper end of the range.

For the dividend claim, the model delivers a risk premium of 5.9%, a risk-free rate of 1.8%, and return volatility of 19.0%. The risk premium and return volatility are close to their empirical counterparts, while the risk-free rate is modestly above its empirical target.

The distinction between the dividend and endowment claims is quantitatively important. Panel B of Table 2 shows that the price impact for the aggregate endowment claim is only 0.10, roughly 22 times smaller than that for the dividend claim. Equivalently, the model-implied aggregate elasticity is roughly 10.0 for the endowment claim but only 0.5 for the dividend claim. Thus, while heterogeneity and frictions generate finite price impact for the aggregate claim, cash-flow leverage substantially amplifies the price impact of equity.

The model also roughly matches sectoral balance sheets, as shown in Panel C of Table 2. The low-risk-aversion active sector is levered, with an asset-to-equity ratio of 1.9 in the model versus 1.6 in the data, whereas the high-risk-aversion active sector is conservative, with a ratio of 0.2 versus 0.5. The distribution of risky assets across the passive, high-risk-aversion active, and low-risk-aversion active sectors is also broadly in line with the data.

While the model is able to generate realistic asset pricing moments with a substantial price impact, it does not reach the upper end of the empirical estimates. We show next that introducing a belief-distortion mechanism can close this gap.

6.3 The belief-distortion mechanism

Recent evidence by Chaudhry (2025) indicates that an increase in prices unrelated to cash flows raises analysts' cash-flow expectations. We extend the baseline model by allowing investors' cash-flow expectations to respond to non-fundamental portfolio-flow shocks. We focus on the main changes relative to the baseline model, and relegate a detailed description to Appendix D.

The model with belief distortions. As in the baseline model, we assume that cash flows evolve according to Equations (9) and (10) under the objective measure. We allow subjective beliefs to differ from objective beliefs, with the subjective cash-flow process given by

$$\frac{dY_t}{Y_t} = (\mu + m_t)\,dt + \sigma\,d\widetilde{Z}_t, \qquad \frac{dD_t}{D_t} = (\mu + \phi m_t)\,dt + \phi\sigma\,d\widetilde{Z}_t,$$

where $\widetilde{Z}_t \in \mathbb{R}^2$ is a Wiener process under the subjective measure.[24]

The term m_t captures the belief distortion, and it is allowed to respond to non-fundamental portfolio-flow shocks. We assume that these beliefs are common across all investors. Formally, m_t evolves according to

$$dm_t = -\kappa_m m_t\,dt + \theta\,dF_t,$$

where $dF_t = \sigma_{\alpha_p,2}\,dZ_{2,t}$ is the component of the passive-demand innovation orthogonal to the

[24] This formulation is consistent with the following change of measure from the objective to the subjective measure: $d\widetilde{Z}_t = dZ_t - h_t^\top\,dt$, where $h_t = \left(\frac{m_t}{\sigma_1}, 0\right)$.

Moment	Model	Data
Risk premium (%)	5.1	6.3
Risk-free rate (%)	1.9	0.9
Return volatility (%)	15.0	19.4
Quarterly gross-flow volatility (%)	0.5	0.5
Price impact	3.0	1.5–5.9

Table 3. Belief distortions without preference heterogeneity or leverage constraints.
This table reports model-implied and empirical unconditional moments for the dividend claim in the economy with belief distortions, homogeneous risk aversion, and slack leverage constraints. Empirical targets, simulation conventions, and the construction of price impact and quarterly gross-flow volatility follow Table 2.

fundamental shock. A positive flow innovation raises investors' expectations of future endowment and dividend growth even though it does not change actual cash flows. This specification captures in reduced form the mislearning from prices studied by Bastianello and Fontanier (2025).[25] The parameter θ governs the initial belief response, while κ_m governs how quickly the distortion dissipates. The rest of the model is unchanged, so we recover the baseline model when $\theta = m_0 = 0$.

Calibrating the belief-distortion mechanism. We set the response of beliefs to portfolio flows to $\theta = 0.08$. To discipline this parameter, we consider an economy without preference heterogeneity or leverage constraints, which isolates the belief-distortion mechanism from the rational mechanisms emphasized in Section 4. In this economy, the calibration implies that a one-percent increase in prices induced by non-fundamental flows raises expected cash-flow growth by approximately 0.12 percentage points, coinciding with the average treatment effect estimated by Chaudhry (2025). We set the mean-reversion parameter to $\kappa_m = 0.20$, corresponding to a half-life of approximately three and a half years. This persistence is consistent with Chaudhry's evidence that the response of expectations does not reverse over the following year but dissipates over longer horizons. The belief state m_t adds a fourth state variable to the dynamic model. We solve the resulting system using the same neural network method as in the baseline economy, as described in Appendix G.3.

Quantitative implications. We first study the behavioral mechanism in an economy with homogeneous risk aversion and slack leverage constraints. Table 3 reports the resulting unconditional moments. The belief-distortion mechanism generates a substantial dividend-claim price impact of 3.0, but still below the upper end of the empirical range. The mechanism is reminiscent of the fundamental shock in Section 2: although current resources do not move, perceived cash-flow news shifts the consumption-wealth schedule and thereby increases the price response required for goods-market clearing. Both the rational and behavioral mechanisms therefore operate through

[25] See also Bastianello (2026) for a discussion of the connection between this mechanism and demand elasticity in the context of a model with a constant risk-free return.

Economy	Price impact	P/D volatility (%)	Flow-driven volatility (%)	Return volatility (%)	Risk premium (%)
Panel A. Dividend claim					
Rational, no frictions	0.31	0.70	0.47	14.31	4.82
Rational, full frictions	2.16	6.05	2.90	18.96	5.89
Behavioral, no frictions	3.03	4.57	4.45	14.98	5.10
Behavioral, full frictions	6.19	9.63	8.03	20.17	6.37
Panel B. Endowment claim					
Rational, no frictions	0.02	0.08	0.03	3.07	1.03
Rational, full frictions	0.10	0.66	0.13	3.64	1.13
Behavioral, no frictions	0.30	0.45	0.43	3.10	1.05
Behavioral, full frictions	0.54	0.98	0.70	3.73	1.16

Table 4. Financial frictions and belief distortions.
This table compares rational and behavioral economies with and without financial frictions. *No frictions* denotes homogeneous risk aversion and slack leverage constraints; *full frictions* introduces preference heterogeneity and the leverage constraint. In each no-frictions economy, common risk aversion equals the post-burn-in wealth-weighted average in the corresponding full-frictions economy: 8.08 under rational beliefs and 7.87 under distorted beliefs. Price-dividend-ratio volatility is the annualized standard deviation of changes in the log price-dividend ratio. Flow-driven volatility is the product of price impact and the annualized volatility of the component of aggregate portfolio flows orthogonal to the fundamental shock. Panel A reports results for the dividend claim, whose cash-flow volatility is 13.8%; Panel B reports results for the endowment claim, whose cash-flow volatility is 3.0%.

the consumption-wealth ratio. Under the rational mechanism, flows change the degree of risk misallocation and hence risk-adjusted expected returns. Under the behavioral mechanism, flows change investors' expectations of future cash flows.

The behavioral economy generates an equity premium of 5.1% and return volatility of 15.0%. Thus, although belief distortions generate substantial price impact, they are not sufficient to match the observed volatility or the upper range of price-impact estimates. We consider next the interaction of the belief-distortion mechanism with the financial frictions emphasized in the baseline model.

6.4 Price impact with financial frictions and belief distortions

We have seen that financial frictions and belief distortions each generate substantial price impact in isolation, but neither reaches the upper end of the empirical estimates. We now study their joint impact. To separate the two mechanisms, we compare four economies that combine objective or distorted beliefs with homogeneous and unconstrained investors or with preference heterogeneity and a leverage constraint. We hold the parameters governing belief distortions fixed at the values used in Section 6.3. Table 4 reports the results for the dividend claim in Panel A and the endowment claim in Panel B.

Table 4 delivers three main results. First, each mechanism independently generates substantial price impact. Under objective beliefs, introducing preference heterogeneity and the leverage

constraint raises dividend-claim price impact from 0.31 to 2.16. With homogeneous and unconstrained investors, introducing distorted beliefs raises price impact from 0.31 to 3.03. Second, the mechanisms are quantitatively complementary. Combining them raises price impact to 6.19, above the value obtained by adding their separate incremental effects. This value exceeds even the upper end of the GIV point estimates in Gabaix and Koijen (2023), while remaining within the 95% confidence interval around their directly estimated GIV multiplier.[26]

A factorial decomposition quantifies this complementarity. Relative to the homogeneous rational benchmark, rational frictions account for 31% of the total amplification, belief distortions for 46%, and their interaction for the remaining 23%. Thus, rational forces and belief distortions make similarly important contributions, and their interaction provides additional amplification.

Third, the two mechanisms play different quantitative roles. Preference heterogeneity and the leverage constraint raise both return volatility and price-dividend-ratio volatility. Belief distortions instead make valuations substantially more sensitive to flows while having little effect on return volatility or the equity premium in the economy without financial frictions: they raise price impact from 0.31 to 3.03 and flow-driven volatility from 0.47% to 4.45%. When belief distortions and financial frictions operate together, flow-driven volatility reaches 8.03%, a magnitude comparable to total price-dividend-ratio volatility of 9.63%. Financial frictions broaden the effects on valuations and returns, while belief distortions amplify the transmission of flows into prices. Together, they allow non-fundamental portfolio flows to generate substantial valuation fluctuations without changes in objective cash-flow fundamentals, providing a flow-driven explanation for excess volatility.

Panel B shows that both mechanisms also increase the price impact of the aggregate endowment claim. Cash-flow leverage substantially magnifies the price impact of the dividend claim, but cannot generate a realistically large response by itself. A large price impact for the dividend claim also requires the endowment claim to depart meaningfully from the infinitely elastic benchmark. Thus, understanding the elasticity of the aggregate endowment claim remains essential even when the ultimate object of interest is the stock market.

Dynamic response to portfolio inflows. Figure 7 traces the response to an inflow driven by the component of portfolio flows orthogonal to fundamentals. The upper-left panel shows that an inflow raises the price-dividend ratio in all four economies. Financial frictions and belief distortions each amplify this response, and their combination produces the largest and most persistent increase in valuations.

The remaining discount-rate panels explain this valuation response. The inflow lowers both the endowment and equity risk premia, while raising the risk-free rate. Because cash-flow leverage

[26]The two-principal-component GIV specification in Gabaix and Koijen (2023, Table 2) directly estimates a multiplier of 5.28 with a Newey–West standard error of 1.10, implying an approximate 95% confidence interval of $[3.12, 7.44]$.

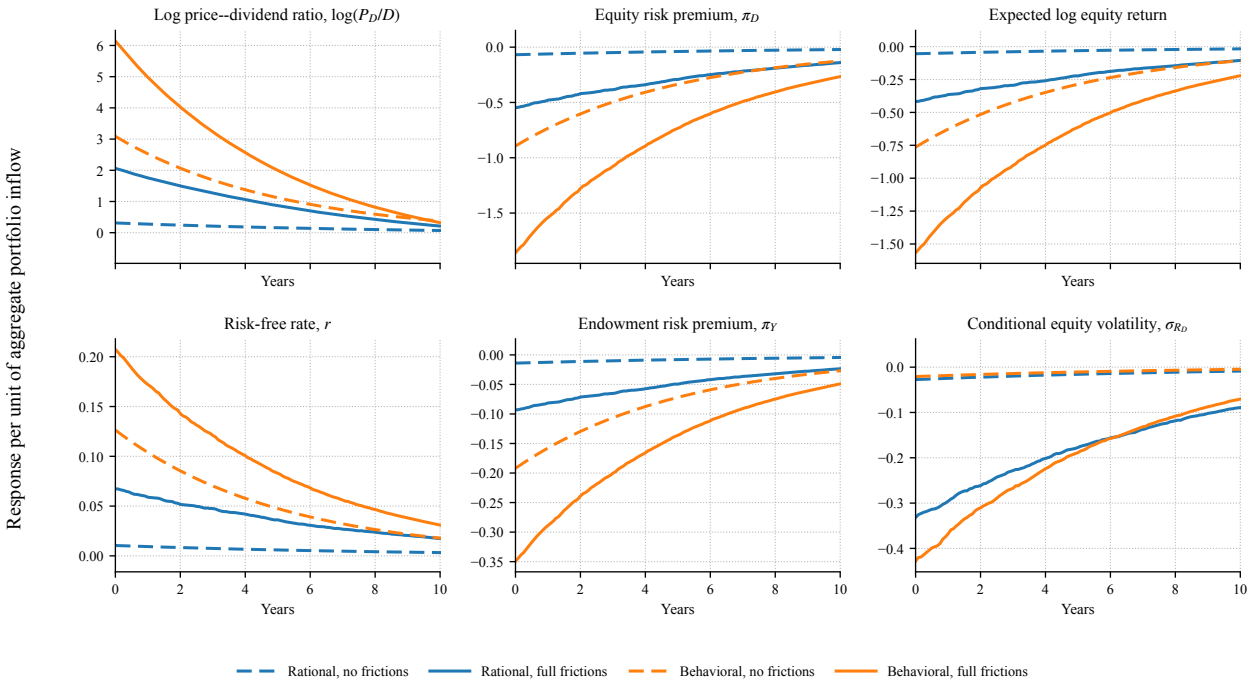


Figure 7. Dynamic response to portfolio inflows.
This figure reports generalized impulse responses to an inflow driven by the component of portfolio flows orthogonal to fundamentals. Responses are expressed per unit of initial aggregate flow and averaged across states drawn from each economy's ergodic distribution. Dashed lines denote economies with homogeneous risk aversion and slack leverage constraints; solid lines introduce preference heterogeneity and the leverage constraint. Blue lines report economies with rational beliefs, and orange lines report economies with belief distortions. The expected log equity return is $r_t + \pi_{D,t} - \frac{1}{2}\|\sigma_{R_D,t}\|^2$.

makes the equity risk premium considerably more responsive than the endowment risk premium, the increase in the risk-free rate only partially offsets the decline in the equity risk premium. Expected log equity returns therefore fall, especially in the economy combining financial frictions and belief distortions, as shown in the top right panel. These responses are the quantitative counterpart of the analytical mechanism in Proposition 7.

The relative importance of these forces can be summarized using a Campbell–Shiller decomposition. Because the portfolio-flow innovation is orthogonal to fundamentals, objective expected dividend growth does not change. We compute the discounted present values over a 30-year horizon, with the remaining valuation response captured by the terminal term. The average initial response of the log price-dividend ratio in the full behavioral economy can therefore be written as[27]

$$\begin{aligned}\Delta \log(P_D/D)_0 &\approx -\Delta\,\mathrm{PDV}(r) - \Delta\,\mathrm{PDV}(\pi_D) + \frac{1}{2}\Delta\,\mathrm{PDV}\left(\|\sigma_{R_D}\|^2\right) + \text{terminal term} \\ &= -0.85 + 7.52 - 0.35 - 0.13 = 6.19.\end{aligned}$$

[27]Here, $\Delta\,\mathrm{PDV}(x)$ denotes the change induced by the portfolio-flow innovation in the discounted present value of the expected path of x, relative to its no-shock path. All changes are expressed per unit of the initial aggregate flow.

The decline in the equity risk premium is the dominant source of the valuation response. The increase in the risk-free rate and the decline in conditional variance partially offset this effect. The small terminal term indicates that the decomposition captures nearly all of the long-run valuation response.

Finally, the lower-right panel shows that an inflow reduces conditional equity volatility, reproducing the negative price–volatility relation documented by Black (1976). This flow-driven leverage effect is substantially stronger in the economies with financial frictions. The decline in volatility attenuates the increase in valuations through the Jensen correction in expected log returns, $-\frac{1}{2}\|\sigma_{R_D,t}\|^2$, but is not large enough to offset the decline in the equity risk premium.

Interpreting the evidence. The results provide a possible interpretation of the range of price-impact estimates in the literature. Dividend payout variation, as studied by Hartzmark and Solomon (2025), may carry relatively little perceived information about cash flows, leaving less scope for confusing non-fundamental and fundamental shocks. The rational mechanism would then be the main source of price impact, and the model generates an estimate in line with their evidence. The variation isolated by the GIV of Gabaix and Koijen (2023) may leave greater scope for such confusion, bringing both mechanisms into play. The joint model generates a price impact of 6.19, above the upper end of their GIV point estimates but within the 95% confidence interval around their estimated multiplier. Thus, the differences between the lower and upper ends of the empirical estimates may reflect different combinations of rational general-equilibrium forces and the cash-flow information investors attribute to flows.

6.5 State dependence and return predictability

We conclude the quantitative analysis with two dynamic implications of the model: state-dependent price impact and return predictability.

State-dependent price impact. Figure 8 examines whether the price impact of the same portfolio-flow shock varies with the distribution of active wealth. We sort simulated observations into bins based on the wealth share of high-risk-aversion active investors, $x_{h,t}$, and estimate price impact within each bin using the component of portfolio flows orthogonal to fundamentals. Price impact varies substantially across states, ranging from roughly 0.6 to 2.9 in the rational model and from 4.7 to 7.1 in the behavioral model.

The hump-shaped pattern reflects two opposing effects of $x_{h,t}$ on risk-bearing capacity. At low and intermediate values, a higher $x_{h,t}$ is associated with more frequent binding of the leverage constraint faced by low-risk-aversion investors, limiting the sector best able to absorb additional

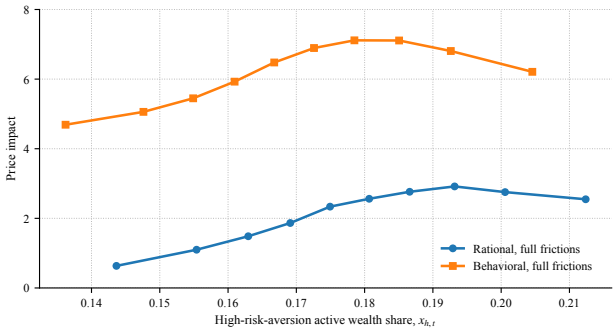


Figure 8. State-dependent price impact

	Rational	Behavioral	Data
Coefficient	−0.118	−0.164	−0.079
S.e.	–	–	(0.035)
R^2	0.013	0.030	0.044

Table 5. Return predictability

Note: The figure reports price-impact estimates within bins of the high-risk-aversion active wealth share, $x_{h,t}$. The rational and behavioral series both correspond to economies with preference heterogeneity and the leverage constraint. Price impact is estimated using the component of portfolio flows orthogonal to fundamentals. The table reports estimates of equation (27) in the rational model, the behavioral model, and the data. The empirical standard error is Newey–West with eight lags.

risk and raising price impact. At sufficiently high values, however, the additional wealth of the unconstrained high-risk-aversion sector expands its capacity to absorb risk, and price impact declines. This interaction between sectoral wealth and the leverage constraint makes price impact largest at intermediate values of $x_{h,t}$.

Belief distortions raise price impact throughout the wealth distribution but preserve the same hump-shaped pattern. They therefore amplify the level of price impact, while the distribution of active risk-bearing capacity remains an important source of its state dependence.

Return predictability. We next examine whether the flow-induced movements in valuations generate return predictability. We estimate

$$\sum_{j=1}^{4} \delta^{j-1} r^e_{t+j} = a + b \log(P_{D,t}/D_t) + u_{t+4}, \tag{27}$$

where r^e_{t+j} is the quarterly log excess equity return and δ is the Campbell–Shiller discount factor. Table 5 compares the model coefficients with the corresponding estimate in the data. The coefficient is negative in both versions of the model: a high price–dividend ratio predicts low subsequent excess returns. The rational and behavioral models generate coefficients of −0.118 and −0.164, respectively, roughly in line with the empirical coefficient of −0.079.

Taken together, the quantitative model is consistent with evidence on sectoral balance sheets, the equity premium, return volatility, price impact, and return predictability. It also generates substantial state dependence in price impact, which is highest when active risk-bearing capacity is impaired. The model therefore connects the level and dynamics of aggregate market elasticity to a range of asset pricing and portfolio choice facts.

7 Conclusion

This paper studies aggregate market elasticity, defined as the inverse of the equilibrium price impact of a portfolio flow between bonds and stocks, in a general equilibrium model with heterogeneous investors, passive demand, and leverage constraints. In the fixed-interest-rate benchmark, aggregate market elasticity is governed by an aggregation of individual demand elasticities. Once the interest rate adjusts, however, goods-market clearing makes the equilibrium price response depend on how portfolio flows shift the consumption–wealth relation. At the efficient allocation, a marginal flow leaves this relation unchanged, and the interest-rate response fully offsets the movement in the risk premium. Risk misallocation breaks this neutrality by changing saving incentives and gives the aggregate claim finite price impact.

Preference heterogeneity and leverage constraints make this price impact state dependent. When the investors best able to absorb risk are poorly capitalized or constrained, a portfolio flow produces a larger price response. Equity is less elastic than the aggregate claim because cash-flow leverage makes its risk premium more responsive to flows, so the interest-rate adjustment cannot fully offset it.

The mechanism does not depend on every household choosing consumption optimally or on the interest rate being determined without monetary policy. It survives the introduction of wealthy hand-to-mouth investors and, in a sticky-price production economy, central-bank control of the short-term interest rate.

Our state-global perturbation method characterizes the rational mechanism in closed form, while the numerical solution quantifies its importance. The rational calibration broadly matches sectoral balance sheets and standard asset-pricing moments while generating equity price impact near the lower end of the empirical estimates. A belief distortion, disciplined by evidence on cash-flow expectations, raises price impact toward the upper end. Rational frictions and distorted beliefs are quantitatively complementary because each makes the consumption–wealth relation respond to portfolio flows, the former through risk misallocation and the latter through subjective cash-flow expectations. The model also generates state-dependent price impact and return predictability consistent with the data. Understanding why markets are inelastic therefore requires looking beyond individual demand curves to how portfolio flows affect consumption, saving, risk allocation, and interest rates.

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.

Atkeson, Andrew, Jonathan Heathcote, and Fabrizio Perri, 2026, A macroeconomic perspective on stock market valuation ratios, NBER Working Paper No. 34748.

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.

Bastianello, Francesca, 2026, Expectations and inelastic markets, Working paper.

Bastianello, Francesca, and Paul Fontanier, 2025, Expectations and learning from prices, *Review of Economic Studies* 92, 1341–1374.

Bigio, Saki, Dejanir Silva, and Eduardo Zilberman, 2025, Heterogeneous beliefs, asset prices, and business cycles, NBER Working Paper No. 33672.

Bilbiie, Florin O., 2008, Limited asset markets participation, monetary policy and (inverted) aggregate demand logic, *Journal of Economic Theory* 140, 162–196.

Bilbiie, Florin O., Sigurd Mølster Galaasen, Refet S. Gürkaynak, Mathis Mæhlum, and Krisztina Molnar, 2025, HANKSSON, CEPR Discussion Paper No. 20090.

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

Black, Fischer, 1976, Studies of stock price volatility changes, in *Proceedings of the 1976 Meetings of the Business and Economic Statistics Section*, 177–181 (American Statistical Association, Washington, DC).

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.

Caballero, Ricardo J., Tomás E. Caravello, and Alp Simsek, 2024, Financial conditions targeting, NBER Working Paper No. 33206.

Caballero, Ricardo J, and Alp Simsek, 2020, A risk-centric model of demand recessions and speculation, *The Quarterly Journal of Economics* 135, 1493–1566.

Caballero, Ricardo J., and Alp Simsek, 2026, Financial conditions targeting in a multi-asset open economy, NBER Working Paper No. 34974.

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).

Campbell, John Y, and John H Cochrane, 1999, By force of habit: A consumption-based explanation of aggregate stock market behavior, *Journal of political Economy* 107, 205–251.

Caramp, Nicolas, and Dejanir H Silva, 2026, Monetary policy and wealth effects: The role of risk and heterogeneity, *Journal of Finance* 81, 1011–1052.

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.

Chaudhry, Aditya, 2025, The impact of prices on analyst cash flow expectations: Reconciling subjective beliefs data with rational discount rate variation, Working paper.

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, 2011, A multiplier approach to understanding the macro implications of household finance, *The Review of Economic Studies* 78, 199–234.

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.

Chinco, Alex, and Marco Sammon, 2024, The passive ownership share is double what you think it is, *Journal of Financial Economics* 157, 103860.

Chodorow-Reich, Gabriel, Plamen T. Nenov, and Alp Simsek, 2021, Stock market wealth and the real economy: A local labor market approach, *American Economic Review* 111, 1613–1657.

Cieslak, Anna, and Annette Vissing-Jorgensen, 2021, The economics of the fed put, *Review of Financial Studies* 34, 4045–4089.

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

Cremers, KJ Martijn, and Antti Petajisto, 2009, How active is your fund manager? a new measure that predicts performance, *Review of Financial Studies* 22, 3329–3365.

Da, Zhi, Borja Larrain, Clemens Sialm, and José Tessada, 2018, Destabilizing financial advice: Evidence from pension fund reallocations, *Review of Financial Studies* 31, 3720–3755.

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

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

De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann, 1990, Noise trader risk in financial markets, *Journal of political Economy* 98, 703–738.

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

Di Maggio, Marco, Amir Kermani, and Kaveh Majlesi, 2020, Stock market returns and consumption, *Journal of Finance* 75, 3175–3219.

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.

Gabaix, Xavier, and Ralph SJ Koijen, 2024, Granular instrumental variables, *Journal of Political Economy* 132, 2274–2303.

Gopalakrishna, Goutham, and Yuntao Wu, 2026, Aliens for continuous time economies, Working paper.

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

Haddad, Valentin, Paul Huebner, and Erik Loualiche, 2025b, How competitive is the stock market? theory, evidence from portfolios, and implications for the rise of passive investing, *American Economic Review* 115, 975–1018.

Haddad, Valentin, and Tyler Muir, 2025, Market macrostructure: Institutions and asset prices, *Annual Review of Financial Economics* 17, 133–150.

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.

Kaplan, Greg, and Giovanni L. Violante, 2014, A model of the consumption response to fiscal stimulus payments, *Econometrica* 82, 1199–1239.

Kargar, Mahyar, Juan Passadore, Dejanir Silva, and Yucheng Yang, 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.

Kyle, Albert S., 1985, Continuous auctions and insider trading, *Econometrica* 1315–1335.

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.

Lochstoer, Lars A., and Tyler Muir, 2026, Actively passive: The rise of market volatility, Working paper, UCLA and NBER.

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 Cole, and Duncan Simester, 2025, Household portfolios and retirement saving over the life cycle, *Journal of Finance* 80, 2739–2787.

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.

Schmitt-Grohé, Stephanie, and Martın Uribe, 2004, Solving dynamic general equilibrium models using a second-order approximation to the policy function, *Journal of economic dynamics and control* 28, 755–775.

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, 2026, The risk channel of unconventional monetary policy, *Review of Financial Studies* Forthcoming.

Tallarini, Jr., Thomas D., 2000, Risk-sensitive real business cycles, *Journal of Monetary Economics* 45, 507–532.

van der Beck, Philippe, Lorenzo Bretscher, and Julie Zhiyu Fu, 2026, A bound on price impact and investor heterogeneity, Working Paper.

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

Werning, Iván, 2015, Incomplete markets and aggregate demand, NBER Working Paper No. 21448.

Appendix

A Derivations

A.1 Derivations for Section 2

The bond demand view. Define the normalized bond demand $b_j \equiv \frac{R_f^{-1} B_j}{W_j} = 1 - c_j(r,\pi) - \alpha_j(1 - c_j(r,\pi))$ as the bond-to-wealth ratio:

$$b_j = (1-\alpha_j)(1-c_j(r,\pi)) = 1 - c_j(r,\pi) - \frac{PQ_j}{W_j}.$$

Notice that the bond demand depends not only on the interest rate, but also on the price of the risky asset, that is, the cross-elasticity will generically be non-zero.

The market clearing condition for bonds can be expressed as follows

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

where $x_j \equiv \frac{\omega_j W_j}{\omega_a W_a + \omega_p W_p} = \omega_j Q_{j,-1}$ denotes the wealth share of investor j.

Notice that $\frac{PQ_j}{W_j} = \frac{P}{Y+P}\frac{\omega_j Q_j}{x_j}$. We can then write the bond market-clearing condition as follows:

$$1 - \overline{c}(\mu - p) - \frac{e^p}{1+e^p}\left[\omega_a Q_a + \omega_p Q_p\right] = 0 \iff \underbrace{\frac{1}{1+e^p} - \overline{c}(\mu - p)}_{\text{excess supply of goods}} - \frac{e^p}{1+e^p}\left[\omega_a Q_a + \omega_p Q_p - 1\right] = 0$$

The left-hand side of the two expressions above gives you total bond demand. In the constant-interest-rate case, the risk premium adjusts to clear the market for risky assets, so $\omega_a Q_a + \omega_p Q_p = 1$, keeping the interest rate fixed. But, due to the cross-elasticity in bond demand, a change in the price of the risky asset — keeping the interest rate fixed — would create a disequilibrium in the bond market. For instance, in the case $\chi_p \equiv -\overline{c}'(\mu - p) > 0$, a decrease in the price of the risky asset would create an insufficient demand for goods and an excess demand for bonds. The excess demand for bonds would create an incentive to reduce the interest rate.

This shows that movements in the price of the risky asset in response to portfolio-flow shocks, keeping the interest rate fixed, would induce a disequilibrium in the bond market. Equilibrium in the market can be restored by allowing the interest rate to adjust. The change in the interest rate, for a fixed price of the risky asset, causes the adjustment in the risk premium required to clear the market for risky assets, while keeping the aggregate bond demand unchanged.

A.2 Derivations for Section 3

A.2.1 Investors' problem

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

$$0 = \max_{C_j, \alpha_j} f_j(C_j, V_j) + V_{j,W}\left[rW_j + \pi^\top \alpha_j W_j - C_j\right] + V_{j,X}\mu_X + \frac{1}{2}V_{j,WW}W_j^2\|\sigma_j\|^2 + V_{j,WX}W_j\sigma_X\sigma_j^\top + \frac{1}{2}\sum_{k=1}^{d}\sigma_{X,k}^\top V_{j,XX}\sigma_{X,k},$$

subject to

$$\alpha_j = \overline{\alpha}_p[\alpha_{D,t}^m, \alpha_{B,t}^m]^\top \qquad \text{for } j = p, \qquad\qquad \sigma_j\sigma_j^\top \leq \overline{\sigma}^2 \qquad \text{for } j = \ell.$$

Here, $\sigma_j = \alpha_j^\top\Sigma$, where $\alpha_{k,t}^m \equiv \frac{P_{k,t}}{P_{Y,t}}$ for $k \in \{D, B\}$. 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 an $N \times 1$ vector, the diffusion σ_X is an $N \times d$ matrix, while σ_R is a $1 \times d$ vector. The notation $\sigma_{X,k}$ denotes the kth column of σ_X, that is, the exposure to the kth Brownian motion.

The optimality conditions for consumption and portfolio shares are given by

$$f_{j,C}(C_j, V_j) = V_{j,W},$$
$$V_{j,W}\pi = -V_{j,WW}W_j\Sigma\Sigma^\top\alpha_j - \Sigma(V_{j,WX}\sigma_X)^\top + \tilde{\nu}_{j,t}\Sigma\Sigma^\top\alpha_j.$$

The second optimality condition holds for $j \in \{h, \ell\}$, and $\tilde{\nu}_{j,t}$ is the Lagrange multiplier on the leverage constraint.

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 exposure for an active investor is given by

$$\sigma_j = -\frac{V_{j,W}}{V_{j,WW}W_j - \tilde{\nu}_{j,t}}\eta - \frac{V_{j,WX}}{V_{j,WW}W_j - \tilde{\nu}_{j,t}}\sigma_X.$$

Here, we use the fact that $\pi = \Sigma\eta^\top$.

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

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

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

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

The optimal risk exposure for active investors is given by

$$\sigma_{j,t} = \frac{1}{1+\nu_{j,t}}\left[\frac{\eta_t}{\gamma_j} + \frac{1-\gamma_j^{-1}}{\psi-1}\sigma_{c_j,t}\right],$$

where $\nu_{j,t} \equiv -\tilde{\nu}_{j,t}/(V_{j,WW,t}W_{j,t})$ is the normalized Lagrange multiplier on the leverage constraint.

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

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

Rearranging the expression above, we obtain

$$c_{j,t} = \psi\rho + (1-\psi)\left[r_t + \eta_t\sigma_{j,t}^\top - \frac{\gamma_j}{2}\|\sigma_{j,t}\|^2\right] + \mu_{c_j,t} + (1-\gamma_j)\sigma_{c_j,t}\sigma_{j,t}^\top + \frac{\psi-\gamma_j}{1-\psi}\frac{\|\sigma_{c_j,t}\|^2}{2},$$

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

$$\frac{dc_{j,t}}{c_{j,t}} = \mu_{c_j,t}dt + \sigma_{c_j,t}dZ_t,$$

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

$$\mu_{c_j,t} = \frac{c_{j,X}}{c_j}\mu_{X,t} + \frac{1}{2}\sum_{k=1}^{d}\sigma_{X,k,t}^\top\frac{c_{j,XX,t}}{c_{j,t}}\sigma_{X,k,t}, \qquad \sigma_{c_j,t} = \frac{c_{j,X}}{c_j}\sigma_{X,t}.$$

A.2.2 Pricing condition

Define $p_{D,t} = P_{D,t}/D_t$ and $p_{Y,t} = P_{Y,t}/Y_t$. For $k \in \{D, Y\}$, let $\phi_D = \phi$ and $\phi_Y = 1$. The cash-flow process for claim k is therefore

$$\frac{d\mathcal{D}_{k,t}}{\mathcal{D}_{k,t}} = \mu\, dt + \phi_k\sigma\, dZ_t.$$

Because $P_{k,t} = p_{k,t}\mathcal{D}_{k,t}$, Itô's lemma implies

$$\sigma_{R_k,t} = \phi_k\sigma + \sigma_{p_k,t}.$$

Absence of arbitrage then gives the pricing condition

$$\frac{1}{p_{k,t}} = r_t + \eta_t(\phi_k\sigma + \sigma_{p_k,t})^\top - \left(\mu + \mu_{p_k,t} + \phi_k\sigma\sigma_{p_k,t}^\top\right),$$

where Itô's lemma gives

$$\mu_{p_k,t} = \frac{p_{k,X,t}}{p_{k,t}}\mu_{X,t} + \frac{1}{2}\sum_{n=1}^{d}\sigma_{X,n,t}^{\top}\frac{p_{k,XX,t}}{p_{k,t}}\sigma_{X,n,t}, \qquad \sigma_{p_k,t} = \frac{p_{k,X,t}}{p_{k,t}}\sigma_{X,t}.$$

The corporate-debt price is recovered from $P_{B,t} = P_{Y,t} - P_{D,t}$. Since the combined equity and corporate-debt claims replicate the endowment claim, their return exposures satisfy

$$P_{Y,t}\sigma_{R_Y,t} = P_{D,t}\sigma_{R_D,t} + P_{B,t}\sigma_{R_B,t}.$$

A.2.3 Aggregate state variable

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

$$x_{j,t} \equiv \frac{\omega_j W_{j,t}}{\sum_{i\in\mathcal{J}}\omega_i W_{i,t}}.$$

We define the aggregate state vector as $X_t = (x_{h,t}, x_{\ell,t}, \overline{\alpha}_{p,t})$.

Market clearing conditions. We can write the market clearing condition for the equity claim, corporate bond claim, and short-term bonds as follows:

$$\sum_{j\in\{p,h,\ell\}}\omega_j W_{j,t}\alpha_{j,k,t} = P_{k,t}, \qquad \sum_{j\in\{p,h,\ell\}}\omega_j W_{j,t}(1-\alpha_{j,D,t}-\alpha_{j,B,t}) = 0,$$

for $k \in \{D, B\}$.

Using the fact that $P_{Y,t} = P_{D,t} + P_{B,t}$, we obtain:

$$\sum_{j\in\{p,h,\ell\}}\omega_j W_{j,t} = P_{Y,t}.$$

The market clearing condition for goods is given by

$$\sum_{j\in\{p,h,\ell\}} x_{j,t}c_{j,t} = \frac{1}{p_{Y,t}}.$$

Define the market-capitalization weights $\alpha_{k,t}^m = P_{k,t}/P_{Y,t}$ for $k \in \{D, B\}$. The risky-asset market clearing conditions can then be written as

$$\sum_{j\in\{p,h,\ell\}} x_{j,t}\alpha_{j,D,t} = \alpha_{D,t}^m, \qquad \sum_{j\in\{p,h,\ell\}} x_{j,t}\alpha_{j,B,t} = \alpha_{B,t}^m.$$

Let $\alpha_t^m = [\alpha_{D,t}^m, \alpha_{B,t}^m]^{\top}$. The two conditions can be written compactly as

$$\sum_{j\in\{p,h,\ell\}} x_{j,t}\alpha_{j,t} = \alpha_t^m.$$

Multiplying both sides by Σ_t and using the aggregation identity for returns gives

$$\sum_{j\in\{p,h,\ell\}} x_{j,t}\sigma_{j,t} = \sigma_{R_Y,t}.$$

Law of motion of the aggregate state variable. The law of motion of $\overline{\alpha}_{p,t}$ is given by (11). 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

$$\frac{dW_{j,t}}{W_{j,t}} = \left[r_t + \eta_t\sigma_{j,t}^\top - c_{j,t}\right]dt + \sigma_{j,t}dZ_t.$$

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

$$\frac{dx_{j,t}}{x_{j,t}} = \left[(\eta_t - \sigma_{R_Y,t})(\sigma_{j,t} - \sigma_{R_Y,t})^\top - \left(c_{j,t} - \frac{1}{p_{Y,t}}\right) + \kappa\frac{\omega_j - x_{j,t}}{x_{j,t}}\right]dt + (\sigma_{j,t} - \sigma_{R_Y,t})\,dZ_t.$$

A.2.4 Equilibrium PDE system

The aggregate equilibrium is characterized by four PDEs for the three consumption-wealth ratios $\{c_{j,t}(X_t)\}_{j\in\mathcal{J}}$ and the valuation ratio $p_{Y,t}(X_t)$. These functions depend on the three-dimensional state vector $X_t = (x_{h,t}, x_{\ell,t}, \overline{\alpha}_{p,t})$. Once the aggregate equilibrium and its SDF have been determined, the equity valuation ratio $p_{D,t}(X_t)$ is obtained from its pricing equation and corporate debt is valued residually. We now derive the aggregate PDE system stated in Proposition 2.

Lemma A.1. *The consumption-wealth ratios, $\{c_{j,t}(X_t)\}_{j\in\{p,h,\ell\}}$, and the price-dividend ratio for the aggregate claim, $p_Y(X_t)$, satisfy the following system of PDEs:*

$$c_{j,t} = \psi\rho + (1-\psi)\left[r_t + \eta_t\,\sigma_{j,t}^\top - \frac{\gamma_j}{2}\|\sigma_{j,t}\|^2\right] + \mu_{c_j,t} + (1-\gamma_j)\sigma_{c_j,t}\sigma_{j,t}^\top + \frac{\psi-\gamma_j}{1-\psi}\frac{\|\sigma_{c_j,t}\|^2}{2}$$

$$\frac{1}{p_{Y,t}} = r_t + \eta_t\,(\sigma + \sigma_{p_Y,t})^\top - \left(\mu + \mu_{p_Y,t} + \sigma\,\sigma_{p_Y,t}^\top\right),$$

where $(r_t, \eta_t, \sigma_{j,t}, \mu_{c_j,t}, \sigma_{c_j,t}, \mu_{p_Y,t}, \sigma_{p_Y,t})$ are all functions of X_t, (c_j, p_Y) and their derivatives.

Proof. We proceed in four steps. First, we show how to solve for the diffusion of the endogenous state variables, $\sigma_{x_h,t}$ and $\sigma_{x_\ell,t}$, in terms of X_t, (c_j, p_Y) and their derivatives. Second, we show how to solve for the drift of the endogenous state variables, $\mu_{x_h,t}$ and $\mu_{x_\ell,t}$, again in terms of X_t, (c_j, p_Y) and their derivatives. Third, we show how to solve for the interest rate, r_t. Finally, we show how to obtain the PDEs for (c_j, p_Y).

In the derivations below, we drop the explicit dependence on X for brevity.

Step 1: Risk exposure of endogenous state variables. Consider the expression for the risk exposure of the aggregate claim and the consumption-wealth ratio:

$$\sigma_{R_Y} = \sigma + \sum_{k=1}^{2} \frac{p_{Y,x_k}}{p_Y}\sigma_{x_k} + \frac{p_{Y,\overline{\alpha}_p}}{p_Y}\sigma_{\overline{\alpha}_p},$$
$$\sigma_{c_j} = \sum_{k=1}^{2} \frac{c_{j,x_k}}{c_j}\sigma_{x_k} + \frac{c_{j,\overline{\alpha}_p}}{c_j}\sigma_{\overline{\alpha}_p},$$

Notice that we can write σ_{R_Y} as follows:

$$\sigma_{R_Y} = \chi_{\sigma_{R_Y},0} + \chi_{\sigma_{R_Y},1}\sigma_{x_h} + \chi_{\sigma_{R_Y},2}\sigma_{x_\ell},$$

where the coefficients $\chi_{\sigma_{R_Y},k}$, $k = 0, 1, 2$, are functions of X given by

$$\chi_{\sigma_{R_Y},0} = \sigma + \frac{p_{Y,\overline{\alpha}_p}}{p_Y}\sigma_{\overline{\alpha}_p},$$
$$\chi_{\sigma_{R_Y},k} = \frac{p_{Y,x_k}}{p_Y}, k = 1, 2.$$

Similarly, we can write σ_{c_j} as follows:

$$\sigma_{c_j} = \chi_{\sigma_{c_j},0} + \chi_{\sigma_{c_j},1}\sigma_{x_h} + \chi_{\sigma_{c_j},2}\sigma_{x_\ell},$$

where the coefficients $\chi_{\sigma_{c_j},k}$, $k = 0, 1, 2$, are functions of X given by

$$\chi_{\sigma_{c_j},0} = \frac{c_{j,\overline{\alpha}_p}}{c_j}\sigma_{\overline{\alpha}_p},$$
$$\chi_{\sigma_{c_j},k} = \frac{c_{j,x_k}}{c_j}, k = 1, 2.$$

Let $x_a = x_h + x_\ell$ and define the average hedging demand of active investors as

$$h_a = \sum_{j\in\{h,\ell\}} \frac{x_j}{x_a}\frac{1-\gamma_j^{-1}}{\psi - 1}\frac{\sigma_{c_j}}{1+\nu_j}.$$

The market price of risk is

$$\eta = \gamma_a\left(\chi_a\sigma_{R_Y} - h_a\right),$$

where

$$\chi_a = \frac{1-(1-x_h-x_\ell)\overline{\alpha}_p}{x_a}, \qquad \gamma_a = \left[\sum_{j\in\{h,\ell\}} \frac{x_j}{x_a}\frac{1}{\gamma_j(1+\nu_j)}\right]^{-1}.$$

We can then write η as follows:

$$\eta = \chi_{\eta,0} + \chi_{\eta,1}\sigma_{x_h} + \chi_{\eta,2}\sigma_{x_\ell},$$

where the coefficients $\chi_{\eta,k}$, $k = 0, 1, 2$, are functions of X given by

$$\chi_{\eta,k} = \gamma_a \left[\chi_a \chi_{\sigma_{R_Y},k} - \sum_{j \in \{h,\ell\}} \frac{x_j}{x_a} \frac{1 - \gamma_j^{-1}}{\psi - 1} \frac{\chi_{\sigma_{c_j},k}}{1 + \nu_j} \right].$$

The risk exposure of investor $j \in \{p, h, \ell\}$ is given by

$$\begin{aligned} \sigma_p &= \overline{\alpha}_p \sigma_{R_Y} \\ \sigma_j &= \frac{1}{1 + \nu_j} \left[\frac{\eta}{\gamma_j} + \frac{1 - \gamma_j^{-1}}{\psi - 1} \sigma_{c_j} \right], \qquad j \in \{h, \ell\}. \end{aligned}$$

We can then solve for σ_j in terms of the risk exposure of the endogenous state variables:

$$\sigma_j = \chi_{\sigma_j,0} + \chi_{\sigma_j,1} \sigma_{x_h} + \chi_{\sigma_j,2} \sigma_{x_\ell},$$

where the coefficients $\chi_{\sigma_j,k}$, $k = 0, 1, 2$, are functions of X given by

$$\begin{aligned} \chi_{\sigma_p,k} &= \overline{\alpha}_p \chi_{\sigma_{R_Y},k} \\ \chi_{\sigma_j,k} &= \frac{1}{1 + \nu_j} \left[\frac{\chi_{\eta,k}}{\gamma_j} + \frac{1 - \gamma_j^{-1}}{\psi - 1} \chi_{\sigma_{c_j},k} \right], \qquad j \in \{h, \ell\}. \end{aligned}$$

The risk exposure of the endogenous state variables is given by

$$\sigma_{x_j} = x_j \left[\sigma_j - \sum_{i \in \mathcal{J}} x_i \sigma_i \right],$$

using the fact that $\sigma_{R_Y} = \sum_{i \in \mathcal{J}} x_i \sigma_i$. Define $\overline{\chi}_{\sigma,m} = \sum_{i \in \mathcal{J}} x_i \chi_{\sigma_i,m}$ for $m \in \{0, 1, 2\}$.

This gives us a system of equations determining $\sigma_{x_j,t}$:

$$\begin{bmatrix} 1 - x_h \left(\chi_{\sigma_h,1} - \overline{\chi}_{\sigma,1} \right) & -x_h \left(\chi_{\sigma_h,2} - \overline{\chi}_{\sigma,2} \right) \\ -x_\ell \left(\chi_{\sigma_\ell,1} - \overline{\chi}_{\sigma,1} \right) & 1 - x_\ell \left(\chi_{\sigma_\ell,2} - \overline{\chi}_{\sigma,2} \right) \end{bmatrix} \begin{bmatrix} \sigma_{x_h} \\ \sigma_{x_\ell} \end{bmatrix} = \begin{bmatrix} x_h \left(\chi_{\sigma_h,0} - \overline{\chi}_{\sigma,0} \right) \\ x_\ell \left(\chi_{\sigma_\ell,0} - \overline{\chi}_{\sigma,0} \right) \end{bmatrix}.$$

Solving the system above, we obtain σ_{x_j} as a function of X and the functions c_j and p_Y and their derivatives. Given σ_{x_j}, we can solve for σ_j, η, σ_{c_j}, and σ_{R_Y} using the expressions above. The risk premium on the aggregate claim is given by $\pi_Y = \sigma_{R_Y} \eta^\top$.

Step 2: Drift of endogenous state variables. Given $(\sigma_j, \eta, \sigma_{R_Y})$, we can solve for the drift of the endogenous state variables:

$$\mu_{x_j} = x_j \left[(\eta - \sigma_{R_Y}) \left(\sigma_j - \sigma_{R_Y} \right)^\top - \left(c_j - \frac{1}{p_Y} \right) \right] + \kappa(\omega_j - x_j),$$

This gives us μ_X and σ_X in terms of X, (c_j, p_Y) and their derivatives. We can then solve for

the drift of c_j and p_Y using Ito's lemma:

$$\mu_{c_j} = \frac{c_{j,X}}{c_j}\mu_X + \frac{1}{2}\sum_{k=1}^{d} \sigma_{X,k}^\top \frac{c_{j,XX}}{c_j}\sigma_{X,k},$$
$$\mu_{p_Y} = \frac{p_{Y,X}}{p_Y}\mu_X + \frac{1}{2}\sum_{k=1}^{d} \sigma_{X,k}^\top \frac{p_{Y,XX}}{p_Y}\sigma_{X,k}.$$

Step 3: Interest rate. Given the expressions derived in the previous steps, we can solve for the shifter of the consumption-wealth ratio $\xi_{j,t}$:

$$\xi_j = \mu_{c_j} + (1-\gamma_j)\sigma_{c_j}\sigma_j^\top + \frac{\psi-\gamma_j}{1-\psi}\frac{\|\sigma_{c_j}\|^2}{2}.$$

We can then solve for $\xi = \sum_{j\in\{p,h,\ell\}} x_j\xi_j$ and the interest rate:

$$r = \rho + \psi^{-1}\left(\mu_{p_Y} + \mu + \sigma\sigma_{p_Y}^\top + \xi\right) + (1-\psi^{-1})\sum_{j\in\{p,h,\ell\}} x_j\frac{\gamma_j}{2}\|\sigma_j\|^2 - \pi_Y.$$

Step 4: PDEs for the consumption-wealth ratio and price-dividend ratio. From the expression for the consumption-wealth ratio and the pricing condition for the aggregate claim, we can write the PDE for c_j and p_Y as follows:

$$c_j = \psi\rho + (1-\psi)\left[r + \eta\,\sigma_j^\top - \frac{\gamma_j}{2}\|\sigma_j\|^2\right] + \mu_{c_j} + (1-\gamma_j)\sigma_{c_j}\sigma_j^\top + \frac{\psi-\gamma_j}{1-\psi}\frac{\|\sigma_{c_j}\|^2}{2}$$
$$\frac{1}{p_Y} = r + \eta\,\sigma_{R_Y}^\top - \left(\mu + \mu_{p_Y} + \sigma\,\sigma_{p_Y}^\top\right),$$

Notice that $(r, \eta, \sigma_j, \mu_{c_j}, \sigma_{c_j}, \mu_{p_Y}, \sigma_{p_Y})$ are all functions of X, (c_j, p_Y) and their derivatives. Hence, these equations form a system of PDEs for (c_j, p_Y). □

Recovering asset positions and individual claim prices. Because Σ_t is nonsingular, each investor's desired risk exposure uniquely determines the portfolio shares invested in equity and corporate debt:

$$\alpha_{j,t} = \left(\sigma_{j,t}\Sigma_t^{-1}\right)^\top.$$

Thus, unlike in the previous formulation, active investors need not hold equity and corporate debt in market-capitalization weights. The market-capitalization portfolio remains special only for passive investors, for whom $\sigma_{p,t} = \overline{\alpha}_{p,t}\sigma_{R_Y,t}$.

Conditional on the aggregate solution for (r_t, η_t), the equity valuation ratio solves

$$\frac{1}{p_{D,t}} = r_t + \eta_t(\phi\sigma + \sigma_{p_D,t})^\top - \left(\mu + \mu_{p_D,t} + \phi\sigma\sigma_{p_D,t}^\top\right).$$

The value of corporate debt then follows from

$$P_{B,t} = P_{Y,t} - P_{D,t} = p_{Y,t}Y_t - p_{D,t}D_t.$$

This residual valuation is consistent with the payoff identity $B_t = Y_t - D_t$.

A.2.5 Closed-form solution to individual claims

Suppose that $r_t = r^*$ and $\eta_t = \eta^*$ are constant. The valuation ratios of the aggregate endowment and equity claims are then constant and satisfy

$$p_Y^* = \frac{1}{r^* + \eta^*\sigma^\top - \mu}, \qquad p_D^* = \frac{1}{r^* + \phi\eta^*\sigma^\top - \mu},$$

provided both denominators are positive. Consequently,

$$P_{Y,t} = p_Y^* Y_t, \qquad P_{D,t} = p_D^* D_t, \qquad P_{B,t} = p_Y^* Y_t - p_D^* D_t.$$

The corresponding return exposures are

$$\sigma_{R_Y} = \sigma, \qquad \sigma_{R_D} = \phi\sigma, \qquad \sigma_{R_B,t} = \frac{p_Y^* Y_t - p_D^* D_t \phi}{p_Y^* Y_t - p_D^* D_t}\sigma.$$

This one-factor limit does not require the two-asset exposure matrix to be nonsingular. It illustrates that corporate debt is generally risky even when the SDF has constant coefficients.

A.3 Derivations for Section 4: Perturbation Solution

In this section, we discuss the derivations of the perturbation solution in detail. The aggregate equilibrium is expressed in terms of the state vector $X = (x, \overline{\alpha}_p)$, where $x = (x_h, x_\ell)$ collects the wealth shares of active investors.

The perturbation is defined by the following parameter restrictions:

$$\gamma_j = \gamma(1 + \hat{\gamma}_j\epsilon), \qquad \overline{\sigma} = \|\sigma\|(1 + \hat{\sigma}\epsilon), \qquad \kappa = \hat{\kappa}\epsilon,$$

where $\sum_{j\in\mathcal{J}} \omega_j\hat{\gamma}_j = 0$.

We also make assumptions about the dynamics of the exogenous state variables. First, the portfolio share of passive investors is written as $\overline{\alpha}_p = 1 + \hat{\alpha}_p\epsilon$, where $\hat{\alpha}_p$ follows the process

$$d\overline{\alpha}_p = \kappa_p(\overline{\alpha} - \overline{\alpha}_p)dt + \hat{\sigma}_p\epsilon dZ_t,$$

where $\overline{\alpha} = 1 + \hat{\alpha}\epsilon$. Equivalently, $d\hat{\alpha}_p = \kappa_p(\hat{\alpha} - \hat{\alpha}_p)dt + \hat{\sigma}_p dZ_t$.

The case $\epsilon = 0$ corresponds to the benchmark economy, in which there is no preference heterogeneity and passive investors are fully invested in the market portfolio.

We expand the consumption-wealth ratio of investor $j \in \mathcal{J}$ and the valuation ratio of the

aggregate claim as follows:

$$c_j(x, \overline{\alpha}_p; \epsilon) = c_{j,0}(x, \hat{\alpha}_p) + c_{j,1}(x, \hat{\alpha}_p)\epsilon + c_{j,2}(x, \hat{\alpha}_p)\epsilon^2 + O(\epsilon^3),$$
$$p_Y(x, \overline{\alpha}_p; \epsilon) = p_{Y,0}(x, \hat{\alpha}_p) + p_{Y,1}(x, \hat{\alpha}_p)\epsilon + p_{Y,2}(x, \hat{\alpha}_p)\epsilon^2 + O(\epsilon^3).$$

The following lemma summarizes some basic properties of the perturbation solution that will be useful for the derivations below.

Lemma A.2 (Basic properties of the perturbation solution)**.** *The perturbation solution satisfies the following properties:*

1. *The zeroth-order terms are constant:*

$$c_{j,0}(x, \hat{\alpha}_p) = c_j^*, \qquad p_{Y,0}(x, \hat{\alpha}_p) = p_Y^*,$$

   for constants c_j^* and p_Y^* independent of $(x, \hat{\alpha}_p)$.

2. *The drift and diffusion of the state variables are first-order in ϵ:*

$$\mu_X = O(\epsilon), \qquad \sigma_X = O(\epsilon).$$

3. *The first-order and second-order corrections to the consumption-wealth ratio and the price-dividend ratio for the aggregate claim can be written as:*

$$c_{j,1}(x, \hat{\alpha}_p) = \psi_{j,0}(x) + \psi_{j,\hat{\alpha}_p}(x)\hat{\alpha}_p, \qquad c_{j,2}(x, \hat{\alpha}_p) = \Psi_{j,0}(x) + \Psi_{j,\hat{\alpha}_p}(x)\hat{\alpha}_p + \Psi_{j,\hat{\alpha}_p^2}(x)\hat{\alpha}_p^2,$$
$$p_{Y,1}(x, \hat{\alpha}_p) = \psi_{Y,0}(x) + \psi_{Y,\hat{\alpha}_p}(x)\hat{\alpha}_p, \qquad p_{Y,2}(x, \hat{\alpha}_p) = \Psi_{Y,0}(x) + \Psi_{Y,\hat{\alpha}_p}(x)\hat{\alpha}_p + \Psi_{Y,\hat{\alpha}_p^2}(x)\hat{\alpha}_p^2,$$

   where $\psi_{j,l}(x)$ and $\Psi_{j,l}(x)$, for $j \in \mathcal{J} \cup \{Y\}$ and $l \in \{0, \hat{\alpha}_p, \hat{\alpha}_p^2\}$, are functions of x.

These properties follow from the definition of the perturbation solution. When $\epsilon = 0$, the exogenous state variables are constant and investors have identical preferences. Hence, the zeroth-order terms are independent of $(x, \hat{\alpha}_p)$. The fact that the drift and diffusion of the state variables are first-order in ϵ comes from the assumptions on the dynamics of the exogenous state variables, and the fact that the zeroth-order terms are constant. The final property describes a solution that is global in the endogenous wealth shares and local in the passive portfolio share.

Our goal is to solve for the coefficients in this perturbation expansion. We proceed in steps. We first characterize the benchmark economy. We then derive the first- and second-order corrections to the consumption-wealth ratios and the valuation ratio of the aggregate claim.

A.3.1 Proof of Lemma 1

Proof. The assumption $\epsilon = 0$ implies that there is no preference heterogeneity and that passive investors are fully invested in the aggregate risky claim. We guess and verify that, in this benchmark economy, expected returns are constant and the leverage constraint has a zero multiplier. In particular, the wealth distribution plays no role in equilibrium. This implies that $\mu_{c_j,0}(X) = \mu_{p_Y,0}(X) = 0$ and $\sigma_{c_j,0}(X) = \sigma_{p_Y,0}(X) = 0_{1\times d}$.

Risk exposure and market price of risk. The risk exposure vector for the aggregate claim is given by

$$\sigma_{R_Y,0}(X) = \sigma.$$

The market price of risk is therefore

$$\eta_0(X) = \gamma\sigma,$$

using the fact that $h_{a,t} = 0$ and $\chi_{a,t} = 1$.

The risk exposure of investor j is therefore

$$\sigma_{j,0}(X) = \sigma.$$

Because $\overline{\sigma} = \|\sigma\|$, the leverage constraint is satisfied in the benchmark economy.

Risk premium and interest rate. The risk premium on the aggregate claim is given by

$$\pi_{Y,0}(X) = \eta_0(X)\sigma_{R_Y,0}(X)^\top = \gamma\|\sigma\|^2,$$

where $\|\sigma\|^2 = \sigma\sigma^\top$.

The interest rate is given by

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

using the fact that $\xi_t = 0$, $\sigma_{j,t} = \sigma$, $\mu_{p_Y,t} = 0$, and $\sigma_{p_Y,t} = 0_{1\times d}$ in the benchmark economy.

Consumption-wealth ratio and price-dividend ratio. The consumption-wealth ratio $c_{j,0}(X)$ is given by

$$c_{j,0}(X) = \psi\rho + (1-\psi)\left[r_0(X) + \eta_0(X)\,\sigma^\top - \frac{\gamma}{2}\|\sigma\|^2\right].$$

Plugging in the expression for $r_0(X)$ and $\eta_0(X)$, we obtain

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

We assume $\rho > \left(1-\psi^{-1}\right)\left(\mu - \frac{\gamma\|\sigma\|^2}{2}\right)$, so the benchmark consumption-wealth ratio is positive.

The goods-market-clearing condition then implies

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

Drift and diffusion of wealth shares. The drift and diffusion of the wealth shares are given by

$$\mu_{x_j,0}(X) = 0, \qquad \sigma_{x_j,0}(X) = 0_{1\times d}, \qquad \forall j \in \{h,\ell\}.$$

Here, we used the fact that $\kappa = 0$, $\sigma_{j,0}(X) = \sigma_{R_Y,0}(X)$, and $c_{j,0}(X) = \frac{1}{p_{Y,0}(X)}$.

At $\epsilon = 0$, the independent portfolio-flow shock vanishes, so the equity–debt exposure matrix need not have full rank. Individual positions in equity and corporate debt may therefore be indeterminate at the benchmark. The benchmark nevertheless determines each investor's total risk exposure uniquely as $\sigma_{j,0} = \sigma$. Because the aggregate equilibrium conditions depend on portfolios only through these exposures, the perturbation below can be conducted directly in exposure space without selecting a benchmark decomposition into equity and corporate-debt positions.

□

A.3.2 Proof of Proposition 3

Proof. The aggregate state vector is $X = (x, \overline{\alpha}_p)$, where $x = (x_h, x_\ell)$ collects the wealth shares of active investors.

We write the portfolio share of passive investors as $\overline{\alpha}_p = 1 + \hat{\alpha}_p \epsilon$, where $\hat{\alpha}_p$ follows a process independent of ϵ:

$$d\hat{\alpha}_p = \kappa_p(\hat{\alpha} - \hat{\alpha}_p)dt + \hat{\sigma}_p dZ_t.$$

This implies $\overline{\alpha} = 1 + \hat{\alpha}\epsilon$ and that the diffusion of $\overline{\alpha}_p$ is of order $O(\epsilon)$. The main text focuses on $\hat{\alpha} = 0$, so the unconditional mean of $\overline{\alpha}_p$ is one.

We consider the following expansion for the consumption-wealth ratio of investor j and the price-dividend ratio of the aggregate claim:

$$\begin{aligned} c_j(x, \overline{\alpha}_p; \epsilon) &= c_{j,0}(x, \hat{\alpha}_p) + c_{j,1}(x, \hat{\alpha}_p)\epsilon + O(\epsilon^2), \\ p_Y(x, \overline{\alpha}_p; \epsilon) &= p_{Y,0}(x, \hat{\alpha}_p) + p_{Y,1}(x, \hat{\alpha}_p)\epsilon + O(\epsilon^2). \end{aligned}$$

The terms $c_{j,0}(x, \hat{\alpha}_p)$ and $p_{Y,0}(x, \hat{\alpha}_p)$ are independent of $(x, \hat{\alpha}_p)$ and are characterized in Lemma 1. We focus on the first-order correction terms $c_{j,1}(x, \hat{\alpha}_p)$ and $p_{Y,1}(x, \hat{\alpha}_p)$.

Diffusion for consumption-wealth ratios and price-dividend ratio. Given the specification of $\overline{\alpha}_p$, we have $\mu_{\overline{\alpha}_p} = O(\epsilon)$ and $\sigma_{\overline{\alpha}_p} = O(\epsilon)$. Similarly, $\mu_{x_j} = O(\epsilon)$ and $\sigma_{x_j} = O(\epsilon)$ for $j \in \{h, \ell\}$. Hence, the drift and diffusion of the state variables are first-order in ϵ, that is, $\mu_X = O(\epsilon)$ and $\sigma_X = O(\epsilon)$.

We characterize next the derivatives of the consumption-wealth ratios and the valuation ratio with respect to X. Given the consumption-wealth ratios and the price-dividend ratio for the aggregate claim are independent of $(x, \hat{\alpha}_p)$ at $\epsilon = 0$, we have that $c_{j,x_j} = O(\epsilon)$ and $p_{Y,x_j} = O(\epsilon)$, as these derivatives are zero at order zero. Moreover, we guess-and-verify that the derivatives of the consumption-wealth ratios and the price-dividend ratio with respect to $\overline{\alpha}_p$ are first-order in ϵ:

$$c_{j,\overline{\alpha}_p} = O(\epsilon); \qquad p_{Y,\overline{\alpha}_p} = O(\epsilon).$$

From the expansion for the consumption-wealth ratios, we can express the derivatives with respect to $\overline{\alpha}_p$ in terms of the derivatives of the first-order correction terms with respect to $\hat{\alpha}_p$:

$$c_{j,\overline{\alpha}_p} = \frac{\partial c_{j,1}}{\partial \hat{\alpha}_p} + O(\epsilon),$$

using the fact that $c_{j,0}$ is independent of $\hat{\alpha}_p$.

Hence, $c_{j,\overline{\alpha}_p} = O(\epsilon)$ implies $\frac{\partial c_{j,1}}{\partial \hat{\alpha}_p} = 0$, so $c_{j,1}$ is independent of $\hat{\alpha}_p$. A similar argument shows that $p_{Y,\overline{\alpha}_p} = O(\epsilon)$ implies that $p_{Y,1}$ is independent of $\hat{\alpha}_p$.

We can then write the diffusion terms for the consumption-wealth ratios and the price-dividend ratio are second-order terms in ϵ:

$$\sigma_{c_j} = \underbrace{\frac{c_{j,X}}{c_j}}_{O(\epsilon)} \times \underbrace{\sigma_X}_{O(\epsilon)} = O(\epsilon^2), \qquad \sigma_{p_Y} = \underbrace{\frac{p_{Y,X}}{p_Y}}_{O(\epsilon)} \times \underbrace{\sigma_X}_{O(\epsilon)} = O(\epsilon^2),$$

This implies that $\sigma_{R_Y,1} = 0_{1\times d}$.

Risk exposure and market price of risk. The risk exposure for the passive investor is given by

$$\sigma_{p,1} = \hat{\alpha}_p \sigma.$$

The first-order correction to the risk exposure for an active investor is given by

$$\sigma_{j,1} = \frac{\eta_1}{\gamma} - \frac{\eta_0}{\gamma}(\hat{\gamma}_j + \hat{\nu}_j),$$

where $\hat{\nu}_j$ is the Lagrange multiplier on the leverage constraint for investor j.

Expanding the market clearing condition for risky assets, we obtain

$$\sum_{j\in\mathcal{J}} x_j \sigma_{j,1} = 0$$

using the fact that $\sigma_{R_Y,1} = 0$.

The first-order correction to the market price of risk is given by

$$\eta_1 = \gamma\sigma \left[\sum_{j\in\{h,\ell\}} \frac{x_j}{x_a}(\hat{\gamma}_j + \hat{\nu}_j) - \frac{x_p}{x_a}\hat{\alpha}_p \right],$$

where $x_a = x_h + x_\ell = 1 - x_p$ is the wealth share of active investors.

The risk exposure of the high-risk-aversion active investor is

$$\sigma_{h,1} = -\left[\frac{x_\ell}{x_a}(\hat{\gamma}_h - \hat{\gamma}_\ell - \hat{\nu}_\ell) + \frac{x_p}{x_a}\hat{\alpha}_p \right] \sigma.$$

and the risk exposure of the low-risk-aversion active investor is

$$\sigma_{\ell,1} = \left[\frac{x_h}{x_a}(\hat{\gamma}_h - \hat{\gamma}_\ell - \hat{\nu}_\ell) - \frac{x_p}{x_a}\hat{\alpha}_p \right] \sigma.$$

Lagrange multiplier on the leverage constraint. The leverage constraint for investor j is

$$\|\sigma_j\| \leq \overline{\sigma}.$$

Using $\sigma_j = \sigma + \sigma_{j,1}\epsilon + O(\epsilon^2)$ and $\overline{\sigma} = \|\sigma\|(1 + \hat{\sigma}\epsilon)$, the first-order expansion of the constraint is

$$\sqrt{(\sigma + \sigma_{j,1}\epsilon)(\sigma + \sigma_{j,1}\epsilon)^\top} = \|\sigma\| \left(1 + \frac{\sigma_{j,1}\sigma^\top}{\sigma\sigma^\top}\epsilon\right) + O(\epsilon^2).$$

Therefore, the first-order constraint is

$$\frac{\sigma_{j,1}\sigma^\top}{\sigma\sigma^\top} \leq \hat{\sigma}.$$

We assume that investor h is unconstrained and investor ℓ may be constrained, so $\hat{\nu}_h = 0$ and $\hat{\nu}_\ell \geq 0$. For investor ℓ, using the expression for $\sigma_{\ell,1}$ above, the first-order leverage constraint is

$$\frac{x_h}{x_a}(\hat{\gamma}_h - \hat{\gamma}_\ell - \hat{\nu}_\ell) - \frac{x_p}{x_a}\hat{\alpha}_p \leq \hat{\sigma}.$$

If the constraint binds, this condition holds with equality, which implies

$$\hat{\nu}_\ell = \hat{\gamma}_h - \hat{\gamma}_\ell - \frac{x_p}{x_h}\hat{\alpha}_p - \frac{x_a}{x_h}\hat{\sigma}.$$

If the constraint is slack, then $\hat{\nu}_\ell = 0$ and

$$\frac{x_h}{x_a}(\hat{\gamma}_h - \hat{\gamma}_\ell) - \frac{x_p}{x_a}\hat{\alpha}_p < \hat{\sigma},$$

or, equivalently,

$$\hat{\gamma}_h - \hat{\gamma}_\ell - \frac{x_p}{x_h}\hat{\alpha}_p - \frac{x_a}{x_h}\hat{\sigma} < 0.$$

Hence, the first-order Lagrange multiplier is

$$\hat{\nu}_\ell = \max\left\{\hat{\gamma}_h - \hat{\gamma}_\ell - \frac{x_p}{x_h}\hat{\alpha}_p - \frac{x_a}{x_h}\hat{\sigma}, 0\right\}.$$

When the constraint binds, substituting the expression for $\hat{\nu}_\ell$ into the market price of risk yields

$$\eta_1 = \gamma\sigma\left[\hat{\gamma}_h - \frac{x_\ell}{x_h}\hat{\sigma} - \frac{x_p}{x_h}\hat{\alpha}_p\right].$$

When the constraint is slack, $\hat{\nu}_\ell = 0$, and the market price of risk is

$$\eta_1 = \gamma\sigma\left[\sum_{j\in\{h,\ell\}}\frac{x_j}{x_a}\hat{\gamma}_j - \frac{x_p}{x_a}\hat{\alpha}_p\right].$$

Equivalently, both cases can be summarized as

$$\eta_1 = \gamma\sigma \left[\sum_{j\in\{h,\ell\}} \frac{x_j}{x_a}(\hat{\gamma}_j + \hat{\nu}_j) - \frac{x_p}{x_a}\hat{\alpha}_p \right],$$

with $\hat{\nu}_h = 0$ and $\hat{\nu}_\ell$ given above.

Risk premium. The first-order risk premium is

$$\pi_{Y,1} = \eta_1 \sigma^\top = \gamma \|\sigma\|^2 \left[\sum_{j\in\{h,\ell\}} \frac{x_j}{x_a}(\hat{\gamma}_j + \hat{\nu}_j) - \frac{x_p}{x_a}\hat{\alpha}_p \right].$$

It increases with the average risk aversion of active investors and, when the leverage constraint binds, with the shadow value of that constraint.

The passive portfolio perturbation changes passive investors' exposure by $x_{p,t}\hat{\alpha}_{p,t}\epsilon\sigma$ in the aggregate exposure-clearing condition. Writing the corresponding normalized flow as $f_t = x_{p,t}\hat{\alpha}_{p,t}$, the response of the risk premium is

$$\frac{\partial \hat{\pi}_{Y,t}}{\partial f_t} = \frac{1}{x_{p,t}} \frac{\partial \pi_{Y,1}}{\partial \hat{\alpha}_p} = -\frac{1}{\zeta_{r,t}},$$

Here, $\zeta_{r,t} = x_{a,t}/(\gamma\|\sigma\|^2)$ when the constraint is slack and $\zeta_{r,t} = x_{h,t}/(\gamma\|\sigma\|^2)$ when it binds.

Drift for consumption-wealth ratios and price-dividend ratio. Itô's lemma implies that the drifts are

$$\mu_{c_j} = \frac{c_{j,X}}{c_j}\mu_X + \frac{1}{2}\sum_{k=1}^{d} \sigma_{X,k}^\top \frac{c_{j,XX}}{c_j}\sigma_{X,k},$$

$$\mu_{p_Y} = \frac{p_{Y,X}}{p_Y}\mu_X + \frac{1}{2}\sum_{k=1}^{d} \sigma_{X,k}^\top \frac{p_{Y,XX}}{p_Y}\sigma_{X,k},$$

where $\sigma_{X,k}$ is the k-th column of σ_X.

Hence, $\mu_{c_j} = O(\epsilon^2)$ and $\mu_{p_Y} = O(\epsilon^2)$, since the gradients are $O(\epsilon)$ and both μ_X and σ_X are $O(\epsilon)$.

Interest rate. The first-order correction to the interest rate is

$$r_1 = (1-\psi^{-1})\frac{\gamma\|\sigma\|^2}{2}\sum_{j\in\mathcal{J}} x_j\hat{\gamma}_j + (1-\psi^{-1})\gamma\|\sigma\|^2 \sum_{j\in\mathcal{J}} x_j \frac{\sigma_{j,1}\sigma^\top}{\|\sigma\|^2} - \pi_{Y,1},$$

where we used the fact that $\xi_j = O(\epsilon^2)$, as $\mu_{c_j} = O(\epsilon^2)$ and $\sigma_{c_j} = O(\epsilon^2)$.

Using $\sum_{j\in\mathcal{J}} x_j \sigma_{j,1} = 0_{1\times d}$, we obtain

$$r_1 = -\gamma\|\sigma\|^2\left[\sum_{j\in\{h,\ell\}}\frac{x_j}{x_a}(\hat{\gamma}_j+\hat{\nu}_j) - \sum_{j\in\mathcal{J}} x_j\hat{\gamma}_j - \frac{x_p}{x_a}\hat{\alpha}_p + \frac{1+\psi^{-1}}{2}\sum_{j\in\mathcal{J}} x_j\hat{\gamma}_j\right].$$

Notice that we can write the expression above as

$$r_1 = -\gamma\|\sigma\|^2\frac{1+\psi^{-1}}{2}\sum_{j\in\mathcal{J}} x_j\hat{\gamma}_j - \left[\pi_{Y,1} - \pi^{fb}_{Y,1}\right],$$

where

$$\pi^{fb}_{Y,1} = \gamma\|\sigma\|^2\sum_{j\in\mathcal{J}} x_j\hat{\gamma}_j.$$

The first term captures the effect of heterogeneity in risk aversion in the frictionless benchmark (no passive investors or leverage constraints). The second term captures the impact of frictions through the wedge between the equilibrium risk premium and its frictionless counterpart.

This implies that the response of the interest rate to portfolio flows is given by

$$\frac{\partial\hat{r}_t}{\partial f_t} = -\frac{\partial\hat{\pi}_{Y,t}}{\partial f_t} = \frac{1}{\zeta_{r,t}}.$$

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

$$c_{j,1} = (1-\psi)\left[r_1 + \eta_1\,\sigma^\top + \eta_0\sigma^\top_{j,1} - \gamma\sigma\sigma^\top_{j,1} - \frac{\gamma}{2}\|\sigma\|^2\hat{\gamma}_j\right].$$

Substituting for r_1 and $\pi_{Y,1}$, the first-order correction to the consumption-wealth ratio is

$$c_{j,1} = \frac{\gamma\|\sigma\|^2}{2}(1-\psi)\left[\left(1-\psi^{-1}\right)\sum_{k\in\mathcal{J}} x_k\hat{\gamma}_k - \hat{\gamma}_j\right],$$

which is independent of $\hat{\alpha}_p$ as conjectured.

Price-dividend ratio of the aggregate claim. The pricing condition implies

$$-\frac{1}{(p^*_Y)^2}p_{Y,1} = r_1 + \pi_{Y,1}.$$

Hence, the first-order correction to the price-dividend ratio is

$$\frac{p_{Y,1}}{p^*_Y} = -p^*_Y(1-\psi^{-1})\frac{\gamma\|\sigma\|^2}{2}\sum_{j\in\mathcal{J}} x_j\hat{\gamma}_j,$$

which is independent of $\hat{\alpha}_p$ as conjectured. If $\psi > 1$, an increase in average risk aversion reduces the price-dividend ratio of the aggregate claim.

This implies that the price impact of the aggregate claim is zero up to first order:

$$\varepsilon_{Y,t}^{-1} = \frac{\partial \hat{p}_{Y,t}}{\partial f_t} = \frac{1}{x_{p,t} p_Y^*} \frac{\partial p_{Y,1}}{\partial \hat{\alpha}_p} = 0.$$

State dynamics. The diffusion of the wealth share of investor j, $j \in \{h, \ell\}$, is given by

$$\begin{aligned}\sigma_{x_j,1} &= x_j(\sigma_{j,1} - \sigma_{R_Y,1}) \\ &= x_j \left[\sum_{k \in \{h,\ell\}} \frac{x_k}{x_a}(\hat{\gamma}_k + \hat{\nu}_k) - \frac{x_p}{x_a}\hat{\alpha}_p - (\hat{\gamma}_j + \hat{\nu}_j) \right] \sigma.\end{aligned}$$

The drift of the wealth share of investor j, $j \in \{h, \ell\}$, is given by

$$\mu_{x_j,1} = x_j \left[(\gamma - 1)\sigma\sigma_{j,1}^\top + \frac{\gamma \|\sigma\|^2}{2}(\psi - 1) \left(\sum_{k \in \mathcal{J}} x_k \hat{\gamma}_k - \hat{\gamma}_j \right) \right] + \hat{\kappa}(\omega_j - x_j),$$

Finally, the law of motion of $\hat{\alpha}_p$ is

$$d\hat{\alpha}_p = \kappa_p(\hat{\alpha} - \hat{\alpha}_p)dt + \hat{\sigma}_p dZ_t,$$

so the drift and diffusion of $\overline{\alpha}_p$ are $\mu_{\overline{\alpha}_p} = \kappa_p(\hat{\alpha} - \hat{\alpha}_p)\epsilon$ and $\sigma_{\overline{\alpha}_p} = \hat{\sigma}_p \epsilon$.

□

A.3.3 Second-order correction: no dynamics in passive portfolio share

We next derive the second-order correction. To keep the algebra transparent, we proceed in two steps. First, we consider the case in which the portfolio share of passive investors is fixed, so

$$\kappa_p = 0, \qquad \hat{\sigma}_p = 0_{1 \times d}.$$

Equivalently, $\hat{\alpha}_p$ is a fixed parameter rather than a stochastic state variable. This removes all drift and diffusion terms associated with $\overline{\alpha}_p$ and leaves the endogenous wealth shares $x = (x_h, x_\ell)$ as the only state variables that contribute to the second-order drift and diffusion of equilibrium objects. After deriving this benchmark second-order system, we then allow for non-zero drift and diffusion in the passive investors' portfolio share.

Step 1: Drift and diffusion of c_j and y_Y.

The expansion of $\sigma_{c_j,t}$ in ϵ is given by

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

This follows from $c_{j,x} = O(\epsilon)$ and $\sigma_x = O(\epsilon)$, so their product is second order.

At second order, it is convenient to work with the dividend yield of the aggregate claim,

$$y_Y \equiv \frac{1}{p_Y}.$$

This is only a change of variables. After solving for the second-order correction to y_Y, we recover the correction to p_Y from

$$p_Y = \frac{1}{y_Y}.$$

If

$$y_Y = y_{Y,0} + y_{Y,1}\epsilon + y_{Y,2}\epsilon^2 + O(\epsilon^3),$$

then

$$p_Y = p_{Y,0} - p_{Y,0}^2 y_{Y,1}\epsilon + \left(p_{Y,0}^3 y_{Y,1}^2 - p_{Y,0}^2 y_{Y,2}\right)\epsilon^2 + O(\epsilon^3).$$

We now derive the dynamics of y_Y directly:

$$\begin{aligned}\sigma_{y_Y}(X,\epsilon) &= \frac{y_{Y,X}(X,\epsilon)}{y_Y(X,\epsilon)}\sigma_X(X,\epsilon) \\ &= \left(1-\psi^{-1}\right)\frac{\gamma\|\sigma\|^2}{2y_{Y,0}}\sum_{k\in\{h,\ell\}}\left(\hat{\gamma}_k - \hat{\gamma}_p\right)x_k\sigma_{k,1}(x,\hat{\alpha}_p)\epsilon^2 + O(\epsilon^3).\end{aligned}$$

The same scaling applies to y_Y, implying that its diffusion starts at order ϵ^2.

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

$$\begin{aligned}\mu_{c_j}(X,\epsilon) &= \frac{c_{j,X}(X,\epsilon)}{c_j(X,\epsilon)}\mu_X(X,\epsilon) + \sum_{k=1}^{d}\frac{1}{2}\sigma_{X,k}^{\top}(X,\epsilon)\frac{c_{j,XX}(X,\epsilon)}{c_j(X,\epsilon)}\sigma_{X,k}(X,\epsilon) \\ &= \frac{c_{j,1,x}(x,\hat{\alpha}_p)}{c_{j,0}(x,\hat{\alpha}_p)}\mu_{x,1}(x,\hat{\alpha}_p)\epsilon^2 + O(\epsilon^3) \\ &= -(\psi-1)\frac{\left(1-\psi^{-1}\right)\gamma\|\sigma\|^2}{2c_{j,0}}\sum_{k\in\{h,\ell\}}\left(\hat{\gamma}_k - \hat{\gamma}_p\right)\mu_{x_k,1}(x,\hat{\alpha}_p)\epsilon^2 + O(\epsilon^3).\end{aligned}$$

The drift of y_Y, t is given by

$$\begin{aligned}\mu_{y_Y}(X,\epsilon) &= \frac{y_{Y,X}(X,\epsilon)}{y_Y(X,\epsilon)}\mu_X(X,\epsilon) + \sum_{k=1}^{d}\frac{1}{2}\sigma_{X,k}^{\top}(X,\epsilon)\frac{y_{Y,XX}(X,\epsilon)}{y_Y(X,\epsilon)}\sigma_{X,k}(X,\epsilon) \\ &= \frac{y_{Y,1,x}(x,\hat{\alpha}_p)}{y_{Y,0}(x,\hat{\alpha}_p)}\mu_{x,1}(x,\hat{\alpha}_p)\epsilon^2 + O(\epsilon^3) \\ &= \left(1-\psi^{-1}\right)\frac{\gamma\|\sigma\|^2}{2y_{Y,0}}\sum_{k\in\{h,\ell\}}\left(\hat{\gamma}_k - \hat{\gamma}_p\right)\mu_{x_k,1}(x,\hat{\alpha}_p)\epsilon^2 + O(\epsilon^3).\end{aligned}$$

Step 2: Risk exposures of investors.

The risk exposure of an active investor is

$$\sigma_{j,t} = \frac{1}{1+\nu_{j,t}}\left[\frac{\eta_t}{\gamma_j} + \frac{1-\gamma_j^{-1}}{\psi-1}\,\sigma_{c_j,t}\right].$$

It is useful to first define the desired exposure before imposing the leverage multiplier:

$$\widetilde{\sigma}_{j,t} \equiv \frac{\eta_t}{\gamma_j} + \frac{1-\gamma_j^{-1}}{\psi-1}\,\sigma_{c_j,t}.$$

We write

$$\begin{aligned}
\eta_t &= \eta_0 + \eta_1\epsilon + \eta_2\epsilon^2 + O(\epsilon^3),\\
\nu_{j,t} &= \hat{\nu}_{j,1}\epsilon + \hat{\nu}_{j,2}\epsilon^2 + O(\epsilon^3),\\
\sigma_{c_j,t} &= \sigma_{c_j,2}\epsilon^2 + O(\epsilon^3),
\end{aligned}$$

where $\hat{\nu}_{j,1} = \hat{\nu}_j$ is the first-order multiplier characterized above. Since $\gamma_j = \gamma(1+\hat{\gamma}_j\epsilon)$, we have

$$\gamma_j^{-1} = \gamma^{-1}\left(1 - \hat{\gamma}_j\epsilon + \hat{\gamma}_j^2\epsilon^2\right) + O(\epsilon^3).$$

Therefore,

$$\widetilde{\sigma}_{j,t} = \widetilde{\sigma}_{j,0} + \widetilde{\sigma}_{j,1}\epsilon + \widetilde{\sigma}_{j,2}\epsilon^2 + O(\epsilon^3),$$

with

$$\begin{aligned}
\widetilde{\sigma}_{j,0} &= \frac{\eta_0}{\gamma} = \sigma,\\
\widetilde{\sigma}_{j,1} &= \frac{\eta_1}{\gamma} - \sigma\hat{\gamma}_j,\\
\widetilde{\sigma}_{j,2} &= \frac{\eta_2}{\gamma} - \frac{\eta_1}{\gamma}\hat{\gamma}_j + \sigma\hat{\gamma}_j^2 + \frac{1-\gamma^{-1}}{\psi-1}\sigma_{c_j,2}.
\end{aligned}$$

Expanding $\sigma_{j,t} = \widetilde{\sigma}_{j,t}/(1+\nu_{j,t})$ gives

$$\begin{aligned}
\sigma_{j,1} &= \widetilde{\sigma}_{j,1} - \hat{\nu}_{j,1}\sigma,\\
\sigma_{j,2} &= \widetilde{\sigma}_{j,2} - \hat{\nu}_{j,1}\widetilde{\sigma}_{j,1} + \left(\hat{\nu}_{j,1}^2 - \hat{\nu}_{j,2}\right)\sigma.
\end{aligned}$$

Equivalently,

$$\sigma_{j,2} = \frac{\eta_2}{\gamma} - \frac{\eta_1}{\gamma}\hat{\gamma}_j + \sigma\hat{\gamma}_j^2 + \frac{1-\gamma^{-1}}{\psi-1}\sigma_{c_j,2} - \hat{\nu}_{j,1}\left(\frac{\eta_1}{\gamma} - \sigma\hat{\gamma}_j\right) + \left(\hat{\nu}_{j,1}^2 - \hat{\nu}_{j,2}\right)\sigma. \tag{A.2}$$

For an unconstrained investor, $\hat{\nu}_{j,1} = \hat{\nu}_{j,2} = 0$, and this reduces to

$$\sigma_{j,2} = \frac{\eta_2}{\gamma} - \frac{\eta_1}{\gamma}\hat{\gamma}_j + \sigma\hat{\gamma}_j^2 + \frac{1-\gamma^{-1}}{\psi - 1}\sigma_{c_j,2}. \tag{A.3}$$

For a constrained investor, the multiplier is determined by the second-order expansion of the leverage constraint. The leverage constraint is

$$\|\sigma_j\| \leq \|\sigma\|(1 + \hat{\sigma}\epsilon).$$

Using

$$\sigma_j = \sigma + \sigma_{j,1}\epsilon + \sigma_{j,2}\epsilon^2 + O(\epsilon^3),$$

we obtain

$$\|\sigma_j\| = \|\sigma\|\left[1 + \frac{\sigma_{j,1}\sigma^\top}{\sigma\sigma^\top}\epsilon + \left(\frac{\sigma_{j,2}\sigma^\top}{\sigma\sigma^\top} + \frac{1}{2}\left(\frac{\sigma_{j,1}\sigma_{j,1}^\top}{\sigma\sigma^\top} - \left(\frac{\sigma_{j,1}\sigma^\top}{\sigma\sigma^\top}\right)^2\right)\right)\epsilon^2\right] + O(\epsilon^3).$$

At first order, $\sigma_{j,1}$ is proportional to σ. Therefore, the curvature term in the second line is zero. If the constraint binds through second order and $\overline{\sigma} = \|\sigma\|(1 + \hat{\sigma}\epsilon)$ has no second-order correction, then

$$\frac{\sigma_{j,2}\sigma^\top}{\sigma\sigma^\top} = 0. \tag{A.4}$$

Combining (A.2) and (A.4) yields

$$\hat{\nu}_{j,2} = \frac{\widetilde{\sigma}_{j,2}\sigma^\top}{\sigma\sigma^\top} - \hat{\nu}_{j,1}\frac{\widetilde{\sigma}_{j,1}\sigma^\top}{\sigma\sigma^\top} + \hat{\nu}_{j,1}^2.$$

Substituting for $\widetilde{\sigma}_{j,1}$ and $\widetilde{\sigma}_{j,2}$ gives

$$\hat{\nu}_{j,2} = \frac{\eta_2\sigma^\top}{\gamma\sigma\sigma^\top} - \hat{\gamma}_j\frac{\eta_1\sigma^\top}{\gamma\sigma\sigma^\top} + \hat{\gamma}_j^2 + \frac{1-\gamma^{-1}}{\psi-1}\frac{\sigma_{c_j,2}\sigma^\top}{\sigma\sigma^\top} - \hat{\nu}_{j,1}\left(\frac{\eta_1\sigma^\top}{\gamma\sigma\sigma^\top} - \hat{\gamma}_j\right) + \hat{\nu}_{j,1}^2. \tag{A.5}$$

In the maintained two-active-investor case, investor h is unconstrained and investor ℓ may be constrained. Thus $\hat{\nu}_{h,1} = \hat{\nu}_{h,2} = 0$. If investor ℓ is constrained, then $\hat{\nu}_{\ell,1} = \hat{\nu}_\ell$ and $\hat{\nu}_{\ell,2}$ is given by (A.5) with $j = \ell$. If investor ℓ is unconstrained, then $\hat{\nu}_{\ell,1} = \hat{\nu}_{\ell,2} = 0$ and (A.3) applies to both active investors.

Step 3: Market price of risk.

Step 4: Risk premium.

The risk premium on the aggregate claim is given by

$$\pi_{Y,t} = \sigma_{R_Y,t}\eta_t^\top.$$

Using that $\sigma_{R_Y,1} = 0$, expanding the risk premium to second order gives

$$\pi_{Y,2} = \sigma_{R_Y,2}\eta_0^\top + \sigma\eta_2^\top = -\sigma_{y_Y,2}\eta_0^\top + \sigma\eta_2^\top.$$

where we used $\sigma_{R_Y,t} = \sigma - \sigma_{y_Y,2}\epsilon^2 + O(\epsilon^3)$.

Step 5: Interest rate.

$$r_t = \rho + \psi^{-1}\left(\mu_{P_Y,t} + \xi_t\right) + (1-\psi^{-1})\sum_{j\in\mathcal{J}} x_{j,t}\frac{\gamma_j}{2}\|\sigma_{j,t}\|^2 - \pi_{Y,t},$$

where $\mu_{P_Y,t}$ is the drift of the aggregate claim value.

Since $P_Y = p_Y Y$, it satisfies

$$\mu_{P_Y,t} = \mu + \mu_{p_Y,t} + \sigma\sigma_{p_Y,t}^\top.$$

At second order, working with the dividend yield implies

$$\mu_{P_Y,2} = -\mu_{y_Y,2} - \sigma\sigma_{y_Y,2}^\top,$$

up to terms of order higher than ϵ^2, since $\sigma_{y_Y,t} = O(\epsilon^2)$ and $\sigma_{p_Y,2} = -\sigma_{y_Y,2}$. Moreover, the intertemporal-hedging term satisfies

$$\xi_2 = \sum_{j\in\mathcal{J}} x_j\left[\mu_{c_j,2} + (1-\gamma)\sigma_{c_j,2}\sigma^\top\right],$$

where terms involving $\|\sigma_{c_j,t}\|^2$ are fourth order and therefore do not enter at order ϵ^2.

Expanding the risk-compensation term gives

$$\left[\sum_{j\in\mathcal{J}} x_j\frac{\gamma_j}{2}\|\sigma_j\|^2\right]_2 = \gamma\sum_{j\in\mathcal{J}} x_j\left[\sigma_{j,2}\sigma^\top + \frac{1}{2}\sigma_{j,1}\sigma_{j,1}^\top + \hat{\gamma}_j\sigma_{j,1}\sigma^\top\right].$$

Therefore, the second-order correction to the interest rate is

$$\begin{aligned} r_2 = {} & -\psi^{-1}\left(\mu_{y_Y,2} + \sigma\sigma_{y_Y,2}^\top\right) + \psi^{-1}\sum_{j\in\mathcal{J}} x_j\left[\mu_{c_j,2} + (1-\gamma)\sigma_{c_j,2}\sigma^\top\right] \\ & + (1-\psi^{-1})\gamma\sum_{j\in\mathcal{J}} x_j\left[\sigma_{j,2}\sigma^\top + \frac{1}{2}\sigma_{j,1}\sigma_{j,1}^\top + \hat{\gamma}_j\sigma_{j,1}\sigma^\top\right] - \pi_{Y,2}. \end{aligned} \tag{A.6}$$

Using market clearing for the risky claim,

$$\sum_{j\in\mathcal{J}} x_j\sigma_{j,2} = \sigma_{R_Y,2} = -\sigma_{y_Y,2},$$

we can also write

$$r_2 = -\psi^{-1}\left(\mu_{y_Y,2} + \sigma\sigma_{y_Y,2}^\top\right) + \psi^{-1}\sum_{j\in\mathcal{J}} x_j\left[\mu_{c_j,2} + (1-\gamma)\sigma_{c_j,2}\sigma^\top\right]$$
$$+ (1-\psi^{-1})\gamma\left[-\sigma_{y_Y,2}\sigma^\top + \sum_{j\in\mathcal{J}} x_j\left(\frac{1}{2}\sigma_{j,1}\sigma_{j,1}^\top + \hat{\gamma}_j\sigma_{j,1}\sigma^\top\right)\right] - \pi_{Y,2}. \tag{A.7}$$

This expression separates the second-order interest-rate correction into the aggregate-claim drift term, the intertemporal-hedging term, the second-order change in average risk compensation, and the second-order risk premium.

Step 6: Dividend yield.

The dividend yield is given by

$$y_Y = r + \eta(\sigma + \sigma_{p_Y})^\top - \left(\mu + \mu_{p_Y} + \sigma\sigma_{p_Y}^\top\right),$$

where σ_{p_Y} and μ_{p_Y} denote the diffusion and drift of the price-dividend ratio. Since $y_Y = 1/p_Y$, we have

$$\sigma_{p_Y,2} = -\sigma_{y_Y,2}, \qquad \mu_{p_Y,2} = -\mu_{y_Y,2},$$

up to terms of order higher than ϵ^2. Expanding the dividend yield to second order gives

$$\begin{aligned} y_{Y,2} &= r_2 + \sigma\eta_2^\top - \eta_0\sigma_{y_Y,2}^\top + \mu_{y_Y,2} + \sigma\sigma_{y_Y,2}^\top \\ &= r_2 + \pi_{Y,2} + \mu_{y_Y,2} + \sigma\sigma_{y_Y,2}^\top, \end{aligned}$$

where the second equality uses the expression for $\pi_{Y,2}$ from Step 4. Substituting the expression for r_2 from (A.7), we obtain

$$y_{Y,2} = (1-\psi^{-1})\mu_{y_Y,2} + (1-\psi^{-1})\sigma\sigma_{y_Y,2}^\top + \psi^{-1}\sum_{j\in\mathcal{J}} x_j\left[\mu_{c_j,2} + (1-\gamma)\sigma_{c_j,2}\sigma^\top\right]$$
$$+ (1-\psi^{-1})\gamma\left[-\sigma_{y_Y,2}\sigma^\top + \sum_{j\in\mathcal{J}} x_j\left(\frac{1}{2}\sigma_{j,1}\sigma_{j,1}^\top + \hat{\gamma}_j\sigma_{j,1}\sigma^\top\right)\right].$$

From Step 1, the first-order corrections imply

$$\mu_{c_j,2} = (1-\psi)\mu_{y_Y,2}, \qquad \sigma_{c_j,2} = (1-\psi)\sigma_{y_Y,2}.$$

Using $\sum_{j\in\mathcal{J}} x_j = 1$, the terms involving $\mu_{y_Y,2}$ cancel. The terms involving $\sigma_{y_Y,2}\sigma^\top$ also cancel. Therefore,

$$y_{Y,2} = (1-\psi^{-1})\gamma\sum_{j\in\mathcal{J}} x_j\left(\frac{1}{2}\sigma_{j,1}\sigma_{j,1}^\top + \hat{\gamma}_j\sigma_{j,1}\sigma^\top\right). \tag{A.8}$$

Equivalently,

$$y_{Y,2} = (1-\psi^{-1})\gamma\|\sigma\|^2\sum_{j\in\mathcal{J}} x_j\left(\frac{1}{2}m_{j,1}^2 + \hat{\gamma}_j m_{j,1}\right), \tag{A.9}$$

where $\sigma_{j,1} = m_{j,1}\sigma$. This expression shows that the second-order correction to the dividend yield depends only on first-order risk reallocations and preference heterogeneity.

The first-order exposure coefficients are affine in $\hat{\alpha}_p$. The following expressions apply locally within a fixed constraint regime; at the boundary between regimes, the solution is only piecewise differentiable. Write

$$m_{j,1} = \bar{m}_j - \ell_j \hat{\alpha}_p.$$

For passive investors,

$$\bar{m}_p = 0, \qquad \ell_p = -1,$$

so that $m_{p,1} = \hat{\alpha}_p$. When the leverage constraint is slack, $\hat{\nu}_\ell = 0$ and

$$\begin{aligned} \bar{m}_h &= \frac{x_\ell}{x_a}(\hat{\gamma}_\ell - \hat{\gamma}_h), & \ell_h &= \frac{x_p}{x_a}, \\ \bar{m}_\ell &= \frac{x_h}{x_a}(\hat{\gamma}_h - \hat{\gamma}_\ell), & \ell_\ell &= \frac{x_p}{x_a}. \end{aligned}$$

When the leverage constraint binds, investor ℓ remains at the exposure cap while investor h absorbs the marginal flow, so

$$\begin{aligned} \bar{m}_h &= -\frac{x_\ell}{x_h}\hat{\sigma}, & \ell_h &= \frac{x_p}{x_h}, \\ \bar{m}_\ell &= \hat{\sigma}, & \ell_\ell &= 0. \end{aligned}$$

These definitions satisfy the first-order market-clearing condition

$$\sum_{j \in \mathcal{J}} x_j m_{j,1} = 0.$$

Substituting $m_{j,1} = \bar{m}_j - \ell_j \hat{\alpha}_p$ into (A.9), the solution for $y_{Y,2}$ is quadratic in $\hat{\alpha}_p$:

$$y_{Y,2}(x, \hat{\alpha}_p) = a_Y^0(x) + b_Y^0(x)\hat{\alpha}_p + d_Y^0(x)\hat{\alpha}_p^2,$$

where

$$\begin{aligned} a_Y^0(x) &= (1 - \psi^{-1})\gamma \|\sigma\|^2 \sum_{j \in \mathcal{J}} x_j \left(\frac{1}{2}\bar{m}_j^2 + \hat{\gamma}_j \bar{m}_j \right), \\ b_Y^0(x) &= -(1 - \psi^{-1})\gamma \|\sigma\|^2 \sum_{j \in \mathcal{J}} x_j \ell_j \left(\bar{m}_j + \hat{\gamma}_j \right), \\ d_Y^0(x) &= \frac{1}{2}(1 - \psi^{-1})\gamma \|\sigma\|^2 \sum_{j \in \mathcal{J}} x_j \ell_j^2. \end{aligned}$$

Using the values of ℓ_j, the quadratic coefficient is

$$d_Y^0(x) = \frac{1}{2}(1-\psi^{-1})\gamma\|\sigma\|^2 \begin{cases} \dfrac{x_p}{x_a}, & \text{if the leverage constraint is slack,} \\ \dfrac{x_p(1-x_\ell)}{x_h}, & \text{if the leverage constraint binds.} \end{cases}$$

The second-order price-dividend-ratio correction follows from the change of variables above:

$$p_{Y,2} = p_{Y,0}^3 y_{Y,1}^2 - p_{Y,0}^2 y_{Y,2}.$$

Step 7: Consumption-wealth ratio.
The consumption-wealth ratio is given by

$$c_{j,t} = \psi\rho + (1-\psi)\left[r_t + \eta_t\,\sigma_{j,t}^\top - \frac{\gamma_j}{2}\|\sigma_{j,t}\|^2\right] + \mu_{c_j,t} + (1-\gamma_j)\sigma_{c_j,t}\sigma_{j,t}^\top + \frac{\psi-\gamma_j}{1-\psi}\frac{\|\sigma_{c_j,t}\|^2}{2}.$$

We expand each term to second order. Using

$$\begin{aligned} r_t &= r_0 + r_1\epsilon + r_2\epsilon^2 + O(\epsilon^3), \\ \eta_t &= \eta_0 + \eta_1\epsilon + \eta_2\epsilon^2 + O(\epsilon^3), \\ \sigma_{j,t} &= \sigma + \sigma_{j,1}\epsilon + \sigma_{j,2}\epsilon^2 + O(\epsilon^3), \end{aligned}$$

and

$$\begin{aligned} \gamma_j &= \gamma(1+\hat{\gamma}_j\epsilon), \\ \sigma_{c_j,t} &= \sigma_{c_j,2}\epsilon^2 + O(\epsilon^3), \\ \mu_{c_j,t} &= \mu_{c_j,2}\epsilon^2 + O(\epsilon^3), \end{aligned}$$

we collect second-order terms.

The first-order term is

$$c_{j,1} = (1-\psi)\left[r_1 + \eta_1\sigma^\top - \frac{\gamma}{2}\|\sigma\|^2\hat{\gamma}_j\right],$$

which matches the previous result.

Before simplifying, the second-order term is

$$\begin{aligned} c_{j,2} = (1-\psi)\Big[&r_2 + \eta_2\sigma^\top + \eta_1\sigma_{j,1}^\top + \eta_0\sigma_{j,2}^\top - \gamma\sigma\sigma_{j,2}^\top - \frac{\gamma}{2}\sigma_{j,1}\sigma_{j,1}^\top - \gamma\hat{\gamma}_j\sigma\sigma_{j,1}^\top\Big] \\ &+ \mu_{c_j,2} + (1-\gamma)\sigma_{c_j,2}\sigma^\top. \end{aligned}$$

Using $\eta_0 = \gamma\sigma$, the terms involving $\sigma_{j,2}$ cancel. Therefore,

$$\begin{aligned} c_{j,2} = (1-\psi)\Big[&r_2 + \eta_2\sigma^\top + \eta_1\sigma_{j,1}^\top - \frac{\gamma}{2}\sigma_{j,1}\sigma_{j,1}^\top - \gamma\hat{\gamma}_j\sigma\sigma_{j,1}^\top\Big] \\ &+ \mu_{c_j,2} + (1-\gamma)\sigma_{c_j,2}\sigma^\top. \end{aligned}$$

Substituting the expressions for r_2, η_2, and $\sigma_{j,2}$ derived in Steps 2–5 yields a closed-form expression for $c_{j,2}$ in terms of first-order reallocations and heterogeneity.

A.3.4 Second-order correction: time-varying passive portfolio share

We next derive the second-order correction in the case the passive portfolio share is time-varying. Now, $\overline{\alpha}_p$ becomes a state variable. The dynamics of $\overline{\alpha}_p = 1 + \hat{\alpha}_p \epsilon$ is given by

$$\mu_{\overline{\alpha}_p,t} = -\kappa_p \hat{\alpha}_p \epsilon,$$
$$\sigma_{\overline{\alpha}_p,t} = \hat{\sigma}_p \epsilon.$$

We consider the following expansion of $c_j(x, \overline{\alpha}_p; \epsilon)$:

$$c_j(x, \overline{\alpha}_p; \epsilon) = c_{j,0} + c_{j,1}(x, \hat{\alpha}_p)\epsilon + c_{j,2}(x, \hat{\alpha}_p)\epsilon^2 + O(\epsilon^3).$$

Because $\hat{\alpha}_p = (\overline{\alpha}_p - 1)/\epsilon$, derivatives with respect to the physical state $\overline{\alpha}_p$ require a change of scale. For a generic object $f(x, \overline{\alpha}_p; \epsilon) = f_0 + f_1(x, \hat{\alpha}_p)\epsilon + f_2(x, \hat{\alpha}_p)\epsilon^2 + O(\epsilon^3)$, we have

$$f_{\overline{\alpha}_p} = f_{1,\hat{\alpha}_p} + \epsilon f_{2,\hat{\alpha}_p} + O(\epsilon^2),$$
$$f_{\overline{\alpha}_p \overline{\alpha}_p} = \epsilon^{-1} f_{1,\hat{\alpha}_p \hat{\alpha}_p} + f_{2,\hat{\alpha}_p \hat{\alpha}_p} + O(\epsilon).$$

In our setting, $c_{j,1}$ and $y_{Y,1}$ are independent of $\hat{\alpha}_p$, so

$$f_{1,\hat{\alpha}_p} = 0, \qquad f_{1,\hat{\alpha}_p \hat{\alpha}_p} = 0,$$

and therefore

$$f_{\overline{\alpha}_p} = \epsilon f_{2,\hat{\alpha}_p} + O(\epsilon^2), \qquad f_{\overline{\alpha}_p \overline{\alpha}_p} = f_{2,\hat{\alpha}_p \hat{\alpha}_p} + O(\epsilon).$$

Similarly, cross-derivatives satisfy

$$f_{x\overline{\alpha}_p} = f_{1,x\hat{\alpha}_p} + \epsilon f_{2,x\hat{\alpha}_p} + O(\epsilon^2) = \epsilon f_{2,x\hat{\alpha}_p} + O(\epsilon^2),$$

since f_1 does not depend on $\hat{\alpha}_p$.

Step 1: Drift and diffusion for c_j and y_Y.

We avoid repeating the derivation for the case in which $\hat{\alpha}_p$ is fixed. For any second-order object z_2, write

$$z_2 = z_2^0 + \Delta z_2,$$

where z_2^0 denotes the second-order correction derived in the fixed-$\hat{\alpha}_p$ economy, and Δz_2 collects only the additional terms generated by the drift and diffusion of $\overline{\alpha}_p$. By construction, Δz_2 vanishes when $\kappa_p = 0$ and $\hat{\sigma}_p = 0_{1\times d}$.

Because $c_{j,1}$ and $y_{Y,1}$ are independent of $\hat{\alpha}_p$, the diffusion terms coming from the time variation

in $\overline{\alpha}_p$ are

$$\Delta\sigma_{c_j,2} = \frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\hat{\sigma}_p,$$
$$\Delta\sigma_{y_Y,2} = \frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}}\hat{\sigma}_p.$$

Therefore,

$$\sigma_{c_j,2} = \sigma^0_{c_j,2} + \frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\hat{\sigma}_p,$$
$$\sigma_{y_Y,2} = \sigma^0_{y_Y,2} + \frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}}\hat{\sigma}_p.$$

Similarly, the new drift terms generated by the time variation in $\overline{\alpha}_p$ are

$$\Delta\mu_{c_j,2} = -\frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\kappa_p\hat{\alpha}_p + \frac{1}{2}\frac{c_{j,2,\hat{\alpha}_p\hat{\alpha}_p}}{c_{j,0}}\hat{\sigma}_p\hat{\sigma}_p^\top,$$
$$\Delta\mu_{y_Y,2} = -\frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}}\kappa_p\hat{\alpha}_p + \frac{1}{2}\frac{y_{Y,2,\hat{\alpha}_p\hat{\alpha}_p}}{y_{Y,0}}\hat{\sigma}_p\hat{\sigma}_p^\top.$$

Thus the only new second-order terms are proportional to κ_p or $\hat{\sigma}_p$.

Step 2 (time-varying α_p): Risk exposures of investors.

Write

$$\sigma_{j,2} = \sigma^0_{j,2} + \Delta\sigma_{j,2},$$

where $\sigma^0_{j,2}$ is the second-order exposure from the fixed-$\hat{\alpha}_p$ economy and $\Delta\sigma_{j,2}$ collects only the new terms generated by the time variation in $\overline{\alpha}_p$.

The desired exposure remains

$$\widetilde{\sigma}_{j,t} \equiv \frac{\eta_t}{\gamma_j} + \frac{1-\gamma_j^{-1}}{\psi-1}\,\sigma_{c_j,t}.$$

Relative to the fixed-$\hat{\alpha}_p$ economy, the second-order desired exposure changes for two reasons: the market price of risk changes by $\Delta\eta_2$, and the consumption-wealth ratio has an additional diffusion induced by $\hat{\sigma}_p$. Thus,

$$\Delta\widetilde{\sigma}_{j,2} = \frac{\Delta\eta_2}{\gamma} + \frac{1-\gamma^{-1}}{\psi-1}\Delta\sigma_{c_j,2}.$$

Using Steps 1 and 2,

$$\Delta\widetilde{\sigma}_{j,2} = \frac{\Delta\eta_2}{\gamma} + \frac{1-\gamma^{-1}}{\psi-1}\frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\hat{\sigma}_p.$$

For an unconstrained investor, the multiplier is zero, so the incremental second-order exposure is simply

$$\Delta\sigma_{j,2} = \Delta\widetilde{\sigma}_{j,2}.$$

Therefore,

$$\Delta\sigma_{j,2} = \frac{\Delta\eta_2}{\gamma} + \frac{1-\gamma^{-1}}{\psi-1}\frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\hat{\sigma}_p.$$

For a constrained investor, the second-order multiplier adjusts to keep the constraint binding. Writing

$$\hat{\nu}_{j,2} = \hat{\nu}^0_{j,2} + \Delta\hat{\nu}_{j,2},$$

the incremental exposure is

$$\Delta\sigma_{j,2} = \Delta\widetilde{\sigma}_{j,2} - \Delta\hat{\nu}_{j,2}\sigma.$$

The second-order constraint condition requires the component of $\Delta\sigma_{j,2}$ in the direction of σ to vanish:

$$\Delta\sigma_{j,2}\sigma^\top = 0.$$

Hence

$$\Delta\hat{\nu}_{j,2} = \frac{\Delta\widetilde{\sigma}_{j,2}\sigma^\top}{\sigma\sigma^\top},$$

and therefore

$$\Delta\sigma_{j,2} = \Delta\widetilde{\sigma}_{j,2} - \frac{\Delta\widetilde{\sigma}_{j,2}\sigma^\top}{\sigma\sigma^\top}\sigma.$$

Thus, for constrained investors, only the component of the new desired exposure orthogonal to σ affects actual risk exposure; the component parallel to σ is absorbed by the second-order multiplier.

Combining the results, the time-varying passive portfolio share affects second-order risk exposures through the new market-price-of-risk term $\Delta\eta_2$ and the new consumption-wealth diffusion term proportional to $\hat{\sigma}_p$.

Step 3 (time-varying α_p): Market price of risk.

Write

$$\eta_2 = \eta_2^0 + \Delta\eta_2,$$

where η_2^0 is the second-order market price of risk from the fixed-$\hat{\alpha}_p$ economy and $\Delta\eta_2$ collects only the new terms generated by the time variation in $\overline{\alpha}_p$.

Recall that

$$\eta_t = \gamma_{a,t}\left[\chi_{a,t}\sigma_{R_Y,t} - h_{a,t}\right].$$

Relative to the fixed-$\hat{\alpha}_p$ economy, there are three new second-order effects. First, the aggregate claim has an additional diffusion induced by $\hat{\sigma}_p$. Second, active investors have additional hedging demand because their consumption-wealth ratios load on $\hat{\sigma}_p$. Third, if the leverage constraint binds for investor ℓ, the second-order multiplier changes, which changes effective aggregate risk aversion.

From Step 1,

$$\Delta\sigma_{y_Y,2} = \frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}}\hat{\sigma}_p,$$

and the additional hedging demand is

$$\Delta h_{a,2} = \frac{1-\gamma^{-1}}{\psi-1}\mathbb{E}_a[\Delta\sigma_{c_j,2}] = \frac{1-\gamma^{-1}}{\psi-1}\mathbb{E}_a\left[\frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\right]\hat{\sigma}_p.$$

The new contribution to effective aggregate risk aversion comes from the second-order multi-

plier. Using the expansion for $\gamma_{a,t}$ from the fixed-$\hat{\alpha}_p$ case,

$$\Delta\Gamma_{a,2} = \mathbb{E}_a[\Delta\hat{\nu}_{j,2}].$$

Therefore, the incremental market price of risk is

$$\Delta\eta_2 = \gamma\left[\mathbb{E}_a[\Delta\hat{\nu}_{j,2}]\sigma - \Delta\sigma_{y_Y,2} - \Delta h_{a,2}\right]. \tag{A.10}$$

If no leverage constraint binds, then $\Delta\hat{\nu}_{j,2} = 0$ for all active investors and (A.10) reduces to

$$\Delta\eta_2 = -\gamma\left[\frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}} + \frac{1-\gamma^{-1}}{\psi-1}\mathbb{E}_a\left[\frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\right]\right]\hat{\sigma}_p.$$

If the leverage constraint binds for investor ℓ, then $\Delta\hat{\nu}_{h,2} = 0$ and $\Delta\hat{\nu}_{\ell,2}$ is pinned down by the incremental second-order constraint. Let

$$H_j \equiv \frac{1-\gamma^{-1}}{\psi-1}\frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}.$$

The incremental desired exposure of investor ℓ is

$$\Delta\widetilde{\sigma}_{\ell,2} = \frac{\Delta\eta_2}{\gamma} + H_\ell\hat{\sigma}_p.$$

The binding constraint requires the component of $\Delta\sigma_{\ell,2}$ in the direction of σ to be zero. Thus,

$$\Delta\hat{\nu}_{\ell,2} = \frac{\Delta\widetilde{\sigma}_{\ell,2}\sigma^\top}{\sigma\sigma^\top} = \frac{\Delta\eta_2\sigma^\top}{\gamma\sigma\sigma^\top} + H_\ell\frac{\hat{\sigma}_p\sigma^\top}{\sigma\sigma^\top}.$$

$$\begin{aligned}\Delta\eta_2 &= \gamma\left[\frac{x_\ell}{x_a}\Delta\hat{\nu}_{\ell,2}\sigma - \frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}}\hat{\sigma}_p - \frac{1-\gamma^{-1}}{\psi-1}\mathbb{E}_a\left[\frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\right]\hat{\sigma}_p\right], \\ \Delta\hat{\nu}_{\ell,2} &= \frac{\Delta\eta_2\sigma^\top}{\gamma\sigma\sigma^\top} + H_\ell\frac{\hat{\sigma}_p\sigma^\top}{\sigma\sigma^\top}.\end{aligned} \tag{A.11}$$

Equations (A.11) jointly determine the new market-price-of-risk component and the new second-order multiplier when investor ℓ is constrained. The key difference relative to the slack case is that part of the new risk is absorbed by the multiplier, which changes effective aggregate risk aversion.

This system can be solved explicitly. Let

$$C_p \equiv \frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}} + \frac{1-\gamma^{-1}}{\psi-1}\mathbb{E}_a\left[\frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\right], \qquad \varrho_p \equiv \frac{\hat{\sigma}_p\sigma^\top}{\sigma\sigma^\top}.$$

Then (A.11) becomes

$$\Delta\eta_2 = \gamma\left[\frac{x_\ell}{x_a}\Delta\hat{\nu}_{\ell,2}\sigma - C_p\hat{\sigma}_p\right],$$
$$\Delta\hat{\nu}_{\ell,2} = \frac{\Delta\eta_2\sigma^\top}{\gamma\sigma\sigma^\top} + H_\ell\varrho_p.$$

Projecting the first equation on σ and substituting into the second equation gives

$$\Delta\hat{\nu}_{\ell,2} = \frac{x_\ell}{x_a}\Delta\hat{\nu}_{\ell,2} - C_p\varrho_p + H_\ell\varrho_p.$$

Since $1 - x_\ell/x_a = x_h/x_a$, we obtain

$$\Delta\hat{\nu}_{\ell,2} = \frac{x_a}{x_h}(H_\ell - C_p)\varrho_p.$$

Substituting back into the first equation yields

$$\Delta\eta_2 = -\gamma C_p\hat{\sigma}_p + \gamma\frac{x_\ell}{x_h}(H_\ell - C_p)\varrho_p\sigma.$$

Equivalently, using

$$\mathbb{E}_a[H_j] = \frac{x_h}{x_a}H_h + \frac{x_\ell}{x_a}H_\ell, \qquad C_p = \frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}} + \mathbb{E}_a[H_j],$$

we can write

$$H_\ell - C_p = \frac{x_h}{x_a}(H_\ell - H_h) - \frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}}.$$

Combining the expressions above, we obtain

$$\Delta\eta_2 = -\gamma\left[\frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}} + \mathbb{E}_a[H_j]\right]\hat{\sigma}_p - \gamma\frac{x_\ell}{x_h}\left[\frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}} + \frac{x_h}{x_a}(H_h - H_\ell)\right]\frac{\hat{\sigma}_p\sigma^\top}{\sigma\sigma^\top}\sigma.$$

Thus the multiplier offsets the component of the new desired exposure in the direction of σ, while the constrained investor's incremental exposure retains the orthogonal component.

Step 4 (time-varying α_p): Risk premium.

The risk premium on the aggregate claim is given by

$$\pi_{Y,t} = \sigma_{R_Y,t}\eta_t^\top.$$

We write

$$\pi_{Y,2} = \pi^0_{Y,2} + \Delta\pi_{Y,2},$$

where $\pi^0_{Y,2}$ is the second-order risk premium in the fixed-$\hat{\alpha}_p$ economy.

Since

$$\sigma_{R_Y,t} = \sigma - \sigma_{y_Y,2}\epsilon^2 + O(\epsilon^3),$$

the second-order risk premium is

$$\pi_{Y,2} = -\sigma_{y_Y,2}\eta_0^\top + \sigma\eta_2^\top.$$

Therefore, the incremental risk premium induced by time variation in $\overline{\alpha}_p$ is

$$\Delta\pi_{Y,2} = -\Delta\sigma_{y_Y,2}\eta_0^\top + \sigma\Delta\eta_2^\top. \tag{A.12}$$

Using $\eta_0 = \gamma\sigma$ and Step 1,

$$\Delta\pi_{Y,2} = -\gamma\frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}}\hat{\sigma}_p\sigma^\top + \sigma\Delta\eta_2^\top.$$

If no leverage constraint binds, substituting the expression for $\Delta\eta_2$ from Step 3 gives

$$\Delta\pi_{Y,2} = -\gamma\left[\frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}}\hat{\sigma}_p\sigma^\top + \left(\frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}} + \frac{1-\gamma^{-1}}{\psi-1}\mathbb{E}_a\left[\frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\right]\right)\sigma\hat{\sigma}_p^\top\right].$$

When the leverage constraint binds for investor ℓ, the same formula (A.12) applies, but $\Delta\eta_2$ is jointly determined with $\Delta\hat{\nu}_{\ell,2}$ by (A.11).

Step 5 (time-varying α_p): Interest rate.

We write the second-order interest rate as

$$r_2 = r_2^0 + \Delta r_2,$$

where r_2^0 is the fixed-$\hat{\alpha}_p$ expression in (A.6), and Δr_2 collects only the new terms generated by the dynamics of $\overline{\alpha}_p$.

From (A.6), the incremental correction is

$$\begin{aligned}\Delta r_2 = &-\psi^{-1}\left(\Delta\mu_{y_Y,2} + \sigma\Delta\sigma_{y_Y,2}^\top\right)\\ &+\psi^{-1}\sum_{j\in\mathcal{J}}x_j\left[\Delta\mu_{c_j,2} + (1-\gamma)\Delta\sigma_{c_j,2}\sigma^\top\right]\\ &+(1-\psi^{-1})\gamma\sum_{j\in\mathcal{J}}x_j\Delta\sigma_{j,2}\sigma^\top - \Delta\pi_{Y,2}.\end{aligned} \tag{A.13}$$

All first-order terms in $\sigma_{j,1}$ and $\hat{\gamma}_j$ are already included in r_2^0 and therefore do not appear in Δr_2.

Using Step 1, the new drift and diffusion terms are

$$\begin{aligned}\Delta\sigma_{c_j,2} &= \frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\hat{\sigma}_p,\\ \Delta\sigma_{y_Y,2} &= \frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}}\hat{\sigma}_p,\\ \Delta\mu_{c_j,2} &= -\frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\kappa_p\hat{\alpha}_p + \frac{1}{2}\frac{c_{j,2,\hat{\alpha}_p\hat{\alpha}_p}}{c_{j,0}}\hat{\sigma}_p\hat{\sigma}_p^\top,\\ \Delta\mu_{y_Y,2} &= -\frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}}\kappa_p\hat{\alpha}_p + \frac{1}{2}\frac{y_{Y,2,\hat{\alpha}_p\hat{\alpha}_p}}{y_{Y,0}}\hat{\sigma}_p\hat{\sigma}_p^\top.\end{aligned}$$

Moreover, from Step 4,

$$\Delta\pi_{Y,2} = -\Delta\sigma_{y_Y,2}\eta_0^\top + \sigma\Delta\eta_2^\top.$$

Substituting this expression into (A.13) gives

$$\begin{aligned}\Delta r_2 = &-\psi^{-1}\Delta\mu_{y_Y,2} - \psi^{-1}\sigma\Delta\sigma_{y_Y,2}^\top \\ &+ \psi^{-1}\sum_{j\in\mathcal{J}} x_j\left[\Delta\mu_{c_j,2} + (1-\gamma)\Delta\sigma_{c_j,2}\sigma^\top\right] \\ &+ (1-\psi^{-1})\gamma\sum_{j\in\mathcal{J}} x_j\Delta\sigma_{j,2}\sigma^\top + \Delta\sigma_{y_Y,2}\eta_0^\top - \sigma\Delta\eta_2^\top.\end{aligned}$$

Since $\eta_0 = \gamma\sigma$, the terms involving $\Delta\sigma_{y_Y,2}$ can also be combined as

$$\left(\gamma - \psi^{-1}\right)\sigma\Delta\sigma_{y_Y,2}^\top.$$

Thus, equivalently,

$$\begin{aligned}\Delta r_2 = &-\psi^{-1}\Delta\mu_{y_Y,2} + \psi^{-1}\sum_{j\in\mathcal{J}} x_j\Delta\mu_{c_j,2} \\ &+ \psi^{-1}(1-\gamma)\sum_{j\in\mathcal{J}} x_j\Delta\sigma_{c_j,2}\sigma^\top + (1-\psi^{-1})\gamma\sum_{j\in\mathcal{J}} x_j\Delta\sigma_{j,2}\sigma^\top \\ &+ \left(\gamma - \psi^{-1}\right)\sigma\Delta\sigma_{y_Y,2}^\top - \sigma\Delta\eta_2^\top. \qquad (A.14)\end{aligned}$$

This expression simplifies substantially because the baseline economy is homogeneous. In particular, $c_{j,0} = y_{Y,0}$ for all j, and the aggregation identity $\Delta y_{Y,2} = \sum_{j\in\mathcal{J}} x_j\Delta c_{j,2}$ implies

$$\sum_{j\in\mathcal{J}} x_j\Delta\mu_{c_j,2} = \Delta\mu_{y_Y,2}, \qquad \sum_{j\in\mathcal{J}} x_j\Delta\sigma_{c_j,2} = \Delta\sigma_{y_Y,2}.$$

The drift terms in (A.14) therefore cancel. Moreover, the consumption-diffusion terms combine with the aggregate-claim diffusion term as

$$\psi^{-1}(1-\gamma)\Delta\sigma_{y_Y,2}\sigma^\top + \left(\gamma - \psi^{-1}\right)\sigma\Delta\sigma_{y_Y,2}^\top = (1-\psi^{-1})\gamma\Delta\sigma_{y_Y,2}\sigma^\top.$$

Using market clearing for the risky claim,

$$\sum_{j\in\mathcal{J}} x_j\Delta\sigma_{j,2} = -\Delta\sigma_{y_Y,2},$$

the remaining exposure terms cancel as well. Hence the incremental interest-rate correction reduces to

$$\Delta r_2 = -\sigma\Delta\eta_2^\top. \qquad (A.15)$$

Thus, after aggregation and market clearing, the only effect of time-varying passive demand on the second-order interest rate operates through the induced change in the market price of risk. When

no leverage constraint binds, $\Delta\eta_2$ is given directly by Step 3. When the leverage constraint binds for investor ℓ, $\Delta\eta_2$ and $\Delta\hat{v}_{\ell,2}$ are jointly determined by (A.11).

Step 6 (time-varying α_p): Dividend yield.

We write the second-order dividend yield as

$$y_{Y,2} = y^0_{Y,2} + \Delta y_{Y,2},$$

where $y^0_{Y,2}$ is the fixed-$\hat{\alpha}_p$ correction in (A.8), and $\Delta y_{Y,2}$ collects only the new terms generated by the dynamics of $\overline{\alpha}_p$.

From the pricing identity,

$$y_{Y,2} = r_2 + \pi_{Y,2} + \mu_{y_Y,2} + \sigma\sigma^\top_{y_Y,2},$$

so the incremental correction satisfies

$$\Delta y_{Y,2} = \Delta r_2 + \Delta\pi_{Y,2} + \Delta\mu_{y_Y,2} + \sigma\Delta\sigma^\top_{y_Y,2}.$$

Using (A.15) and (A.12), we obtain

$$\begin{aligned}\Delta y_{Y,2} &= -\sigma\Delta\eta_2^\top + \left(-\Delta\sigma_{y_Y,2}\eta_0^\top + \sigma\Delta\eta_2^\top\right) + \Delta\mu_{y_Y,2} + \sigma\Delta\sigma^\top_{y_Y,2} \\ &= \Delta\mu_{y_Y,2} + (1-\gamma)\Delta\sigma_{y_Y,2}\sigma^\top,\end{aligned}$$

where the terms involving $\Delta\eta_2$ cancel and we use $\eta_0 = \gamma\sigma$.

Using Step 1, this becomes

$$\Delta y_{Y,2} = -\frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}}\kappa_p\hat{\alpha}_p + \frac{1}{2}\frac{y_{Y,2,\hat{\alpha}_p\hat{\alpha}_p}}{y_{Y,0}}\hat{\sigma}_p\hat{\sigma}_p^\top + (1-\gamma)\frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}}\hat{\sigma}_p\sigma^\top.$$

Equivalently, since $y_{Y,2} = y^0_{Y,2} + \Delta y_{Y,2}$, the incremental dividend-yield correction solves

$$\Delta y_{Y,2} = \frac{(y^0_{Y,2} + \Delta y_{Y,2})_{\hat{\alpha}_p}}{y_{Y,0}}\left[-\kappa_p\hat{\alpha}_p + (1-\gamma)\hat{\sigma}_p\sigma^\top\right] + \frac{1}{2}\frac{(y^0_{Y,2} + \Delta y_{Y,2})_{\hat{\alpha}_p\hat{\alpha}_p}}{y_{Y,0}}\hat{\sigma}_p\hat{\sigma}_p^\top. \tag{A.16}$$

This is the time-varying-α_p counterpart of (A.8): the new correction is the drift and risk-adjusted diffusion of the total second-order dividend yield induced by the state variable $\overline{\alpha}_p$.

The second-order price-dividend-ratio correction follows from

$$p_{Y,2} = p^3_{Y,0}y^2_{Y,1} - p^2_{Y,0}y_{Y,2},$$

so the incremental term is

$$\Delta p_{Y,2} = -p^2_{Y,0}\Delta y_{Y,2}.$$

Step 7 (time-varying α_p): Solving for coefficients of dividend yield.

We can solve the coefficients of $\Delta y_{Y,2}$ directly from (A.16). Write the fixed-$\hat{\alpha}_p$ correction as

$$y^0_{Y,2}(x,\hat{\alpha}_p) = a^0_Y(x) + b^0_Y(x)\hat{\alpha}_p + d^0_Y(x)\hat{\alpha}_p^2,$$

and conjecture

$$\Delta y_{Y,2}(x, \hat{\alpha}_p) = a_Y(x) + b_Y(x)\hat{\alpha}_p + d_Y(x)\hat{\alpha}_p^2.$$

Let

$$b_Y^T(x) \equiv b_Y^0(x) + b_Y(x), \qquad d_Y^T(x) \equiv d_Y^0(x) + d_Y(x)$$

denote the linear and quadratic coefficients of the total second-order dividend-yield correction $y_{Y,2}^0 + \Delta y_{Y,2}$. Then

$$(y_{Y,2}^0 + \Delta y_{Y,2})_{\hat{\alpha}_p} = b_Y^T + 2d_Y^T\hat{\alpha}_p, \qquad (y_{Y,2}^0 + \Delta y_{Y,2})_{\hat{\alpha}_p\hat{\alpha}_p} = 2d_Y^T.$$

For compactness, define

$$M_p \equiv (1-\gamma)\hat{\sigma}_p\sigma^\top, \qquad V_p \equiv \hat{\sigma}_p\hat{\sigma}_p^\top.$$

Then (A.16) can be written as

$$\Delta y_{Y,2} = \frac{1}{y_{Y,0}}(b_Y^T + 2d_Y^T\hat{\alpha}_p)\left(M_p - \kappa_p\hat{\alpha}_p\right) + \frac{d_Y^T}{y_{Y,0}}V_p. \tag{A.17}$$

We first match the coefficient on $\hat{\alpha}_p^2$. The only term in (A.17) that generates a quadratic coefficient is the product of $2d_Y^T\hat{\alpha}_p$ with $-\kappa_p\hat{\alpha}_p$. Therefore,

$$d_Y = -\frac{2\kappa_p}{y_{Y,0}}d_Y^T = -\frac{2\kappa_p}{y_{Y,0}}(d_Y^0 + d_Y).$$

Solving for d_Y yields

$$d_Y = -\frac{\frac{2\kappa_p}{y_{Y,0}}}{1 + \frac{2\kappa_p}{y_{Y,0}}}\, d_Y^0. \tag{A.18}$$

Equivalently, the quadratic coefficient of the total second-order correction is

$$d_Y^T = \frac{1}{1 + \frac{2\kappa_p}{y_{Y,0}}}\, d_Y^0. \tag{A.19}$$

Thus, mean reversion in passive demand attenuates the quadratic dependence of the dividend yield on $\hat{\alpha}_p$. When $\kappa_p = 0$, we recover $d_Y = 0$ and $d_Y^T = d_Y^0$, as in the fixed-$\hat{\alpha}_p$ economy.

We can similarly solve for the coefficient on $\hat{\alpha}_p$. Matching the linear coefficient in (A.17) gives

$$b_Y = \frac{1}{y_{Y,0}}\left(2d_Y^T M_p - \kappa_p b_Y^T\right) = \frac{1}{y_{Y,0}}\left(2d_Y^T M_p - \kappa_p(b_Y^0 + b_Y)\right).$$

Solving for b_Y yields

$$b_Y = \frac{2d_Y^T M_p - \kappa_p b_Y^0}{y_{Y,0} + \kappa_p}. \tag{A.20}$$

Substituting (A.19), this can also be written as

$$b_Y = \frac{\frac{2M_p}{1+\frac{2\kappa_p}{y_{Y,0}}} d_Y^0 - \kappa_p b_Y^0}{y_{Y,0} + \kappa_p}.$$

Equivalently, the linear coefficient of the total second-order correction is

$$b_Y^T = b_Y^0 + b_Y = \frac{y_{Y,0} b_Y^0 + 2 d_Y^T M_p}{y_{Y,0} + \kappa_p}. \tag{A.21}$$

Finally, we solve for the constant coefficient. Matching the constant term in (A.17) gives

$$a_Y = \frac{1}{y_{Y,0}} \left(b_Y^T M_p + d_Y^T V_p \right).$$

Using (A.21) and (A.19), this becomes

$$a_Y = \frac{1}{y_{Y,0}} \left[\frac{y_{Y,0} b_Y^0 + 2 d_Y^T M_p}{y_{Y,0} + \kappa_p} M_p + d_Y^T V_p \right], \tag{A.22}$$

where

$$d_Y^T = \frac{1}{1 + \frac{2\kappa_p}{y_{Y,0}}} d_Y^0.$$

Equivalently, substituting out d_Y^T,

$$a_Y = \frac{1}{y_{Y,0}} \left[\frac{y_{Y,0} b_Y^0 + \frac{2M_p}{1+\frac{2\kappa_p}{y_{Y,0}}} d_Y^0}{y_{Y,0} + \kappa_p} M_p + \frac{1}{1 + \frac{2\kappa_p}{y_{Y,0}}} d_Y^0 V_p \right].$$

Therefore, the incremental dividend-yield correction is

$$\Delta y_{Y,2}(x, \hat{\alpha}_p) = a_Y(x) + b_Y(x)\hat{\alpha}_p + d_Y(x)\hat{\alpha}_p^2,$$

with d_Y, b_Y, and a_Y given by (A.18), (A.20), and (A.22).

Step 8 (time-varying α_p): Consumption-wealth ratio.
We write the second-order consumption-wealth ratio as

$$c_{j,2} = c_{j,2}^0 + \Delta c_{j,2},$$

where $c_{j,2}^0$ is the fixed-$\hat{\alpha}_p$ correction from the previous subsection and $\Delta c_{j,2}$ collects only the new terms generated by the dynamics of $\overline{\alpha}_p$.

The consumption-wealth ratio satisfies

$$c_{j,t} = \psi\rho + (1-\psi)\left[r_t + \eta_t\,\sigma_{j,t}^\top - \frac{\gamma_j}{2}\|\sigma_{j,t}\|^2\right] + \mu_{c_j,t} + (1-\gamma_j)\sigma_{c_j,t}\sigma_{j,t}^\top + \frac{\psi-\gamma_j}{1-\psi}\frac{\|\sigma_{c_j,t}\|^2}{2}.$$

The last term is fourth order and does not enter the second-order expansion.

From the fixed-$\hat{\alpha}_p$ derivation, the second-order term can be written as

$$\begin{aligned} c_{j,2} = (1-\psi)\Big[r_2 + \eta_2\sigma^\top + \eta_1\sigma_{j,1}^\top - \frac{\gamma}{2}\sigma_{j,1}\sigma_{j,1}^\top - \gamma\hat{\gamma}_j\sigma\sigma_{j,1}^\top\Big] \\ + \mu_{c_j,2} + (1-\gamma)\sigma_{c_j,2}\sigma^\top. \end{aligned}$$

Here the terms involving $\sigma_{j,2}$ cancel because $\eta_0 = \gamma\sigma$. Therefore, relative to the fixed-$\hat{\alpha}_p$ economy,

$$\Delta c_{j,2} = (1-\psi)\left[\Delta r_2 + \Delta\eta_2\sigma^\top\right] + \Delta\mu_{c_j,2} + (1-\gamma)\Delta\sigma_{c_j,2}\sigma^\top.$$

All first-order terms involving η_1, $\sigma_{j,1}$, and $\hat{\gamma}_j$ are already included in $c_{j,2}^0$ and therefore do not appear in $\Delta c_{j,2}$.

Using Step 1, this becomes

$$\begin{aligned} \Delta c_{j,2} = (1-\psi)\left[\Delta r_2 + \Delta\eta_2\sigma^\top\right] \\ - \frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\kappa_p\hat{\alpha}_p + \frac{1}{2}\frac{c_{j,2,\hat{\alpha}_p\hat{\alpha}_p}}{c_{j,0}}\hat{\sigma}_p\hat{\sigma}_p^\top \\ + (1-\gamma)\frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\hat{\sigma}_p\sigma^\top. \end{aligned} \tag{A.23}$$

This equation is the time-varying-α_p counterpart of the fixed-$\hat{\alpha}_p$ expression for $c_{j,2}$. It shows that the new correction to the consumption-wealth ratio is driven by the changes in the interest rate and market price of risk, together with the direct drift and diffusion terms induced by the dynamics of passive demand.

Equations (A.15), (A.10), and (A.23) form the closed system for the incremental second-order corrections when the passive portfolio share is time-varying.

Step 9 (time-varying α_p): Solving for coefficients of consumption-wealth ratios.

We now solve for the coefficients of the incremental correction to the consumption-wealth ratios. The simplification in Step 5 implies

$$\Delta r_2 + \Delta\eta_2\sigma^\top = 0.$$

Using (A.23), the incremental correction to $c_{j,2}$ satisfies

$$\Delta c_{j,2} = \frac{(c_{j,2}^0 + \Delta c_{j,2})_{\hat{\alpha}_p}}{c_{j,0}}\left[-\kappa_p\hat{\alpha}_p + (1-\gamma)\hat{\sigma}_p\sigma^\top\right] + \frac{1}{2}\frac{(c_{j,2}^0 + \Delta c_{j,2})_{\hat{\alpha}_p\hat{\alpha}_p}}{c_{j,0}}\hat{\sigma}_p\hat{\sigma}_p^\top. \tag{A.24}$$

This is the same scalar coefficient problem as for the aggregate dividend yield, but applied investor by investor.

Write the fixed-$\hat{\alpha}_p$ correction as

$$c^0_{j,2}(x, \hat{\alpha}_p) = a^0_j(x) + b^0_j(x)\hat{\alpha}_p + d^0_j(x)\hat{\alpha}^2_p,$$

and conjecture

$$\Delta c_{j,2}(x, \hat{\alpha}_p) = a_j(x) + b_j(x)\hat{\alpha}_p + d_j(x)\hat{\alpha}^2_p.$$

Let

$$b^T_j(x) \equiv b^0_j(x) + b_j(x), \qquad d^T_j(x) \equiv d^0_j(x) + d_j(x)$$

denote the linear and quadratic coefficients of the total second-order correction $c^0_{j,2} + \Delta c_{j,2}$. As before, define

$$M_p \equiv (1-\gamma)\hat{\sigma}_p\sigma^\top, \qquad V_p \equiv \hat{\sigma}_p\hat{\sigma}^\top_p.$$

Then (A.24) can be written as

$$\Delta c_{j,2} = \frac{1}{c_{j,0}}(b^T_j + 2d^T_j\hat{\alpha}_p)(M_p - \kappa_p\hat{\alpha}_p) + \frac{d^T_j}{c_{j,0}}V_p.$$

Matching the coefficient on $\hat{\alpha}^2_p$ gives

$$d_j = -\frac{2\kappa_p}{c_{j,0}}d^T_j = -\frac{2\kappa_p}{c_{j,0}}(d^0_j + d_j).$$

Hence

$$d_j = -\frac{\frac{2\kappa_p}{c_{j,0}}}{1 + \frac{2\kappa_p}{c_{j,0}}}\, d^0_j, \tag{A.25}$$

and the quadratic coefficient of the total correction is

$$d^T_j = \frac{1}{1 + \frac{2\kappa_p}{c_{j,0}}}\, d^0_j. \tag{A.26}$$

Matching the coefficient on $\hat{\alpha}_p$ gives

$$b_j = \frac{1}{c_{j,0}}\left(2d^T_j M_p - \kappa_p b^T_j\right) = \frac{1}{c_{j,0}}\left(2d^T_j M_p - \kappa_p(b^0_j + b_j)\right).$$

Solving for b_j yields

$$b_j = \frac{2d^T_j M_p - \kappa_p b^0_j}{c_{j,0} + \kappa_p}. \tag{A.27}$$

Equivalently, the linear coefficient of the total correction is

$$b^T_j = b^0_j + b_j = \frac{c_{j,0}b^0_j + 2d^T_j M_p}{c_{j,0} + \kappa_p}. \tag{A.28}$$

Finally, matching the constant term gives

$$a_j = \frac{1}{c_{j,0}} \left(b_j^T M_p + d_j^T V_p \right).$$

Using (A.28) and (A.26), this becomes

$$a_j = \frac{1}{c_{j,0}} \left[\frac{c_{j,0} b_j^0 + 2 d_j^T M_p}{c_{j,0} + \kappa_p} M_p + d_j^T V_p \right]. \tag{A.29}$$

Therefore,

$$\Delta c_{j,2}(x, \hat{\alpha}_p) = a_j(x) + b_j(x)\hat{\alpha}_p + d_j(x)\hat{\alpha}_p^2, \tag{A.30}$$

with d_j, b_j, and a_j given by (A.25), (A.27), and (A.29). Because $c_{j,0} = y_{Y,0}$ in the homogeneous benchmark economy and $y_{Y,2} = \sum_{j \in \mathcal{J}} x_j c_{j,2}$, aggregating (A.30) across investors recovers the coefficient formulas for $\Delta y_{Y,2}$ derived above.

B Wealthy Hand-to-Mouth Consumers

This appendix introduces wealthy hand-to-mouth investors into the dynamic model of Section 3. We first establish an exact neutrality result: investors who hold a fixed fraction of the aggregate claim and consume its payouts leave equilibrium prices and the normalized allocations of all other investors unchanged. We then use the two-period model of Section 2 to characterize departures from this neutrality benchmark.

B.1 Neutrality in the dynamic model

Consider the economy of Section 3 and introduce an additional investor sector, indexed by m. Let its population mass be $\omega_m \in (0, 1)$, and scale the population masses of the original types proportionally so that they sum to $1 - \omega_m$. The sector holds a constant fraction $x_m \in [0, 1)$ of the outstanding equity and corporate-debt claims:

$$\omega_m Q_{m,D,t} = \omega_m Q_{m,B,t} = x_m. \tag{B.1}$$

Thus, the sector holds the two risky claims in aggregate-market proportions, holds no risk-free bonds, and has aggregate wealth

$$\omega_m W_{m,t} = x_m P_{Y,t}.$$

The sector consumes the payouts on these claims:

$$\omega_m C_{m,t} = x_m (D_t + B_t) = x_m Y_t. \tag{B.2}$$

Its consumption therefore responds to current income but not to changes in the market value of its wealth.

The rule in Equations (B.1)–(B.2) is internally consistent. Indeed, the sector's budget constraint

implies

$$\begin{aligned} d(\omega_m W_{m,t}) &= x_m\big(dP_{D,t} + D_t dt\big) + x_m\big(dP_{B,t} + B_t dt\big) - \omega_m C_{m,t} dt \\ &= x_m dP_{Y,t}. \end{aligned} \tag{B.3}$$

This is precisely the change in wealth generated by maintaining the constant positions in Equation (B.1). We assume that demographic turnover preserves these positions and that the estate transfers among the remaining investor types are scaled proportionally, so their conditional wealth distribution evolves as in the baseline economy.

The following proposition establishes the neutrality result.

Proposition 8 (Wealthy hand-to-mouth neutrality)**.** *Consider any equilibrium of the dynamic economy in Section 3. Introduce the wealthy hand-to-mouth sector described by Equations (B.1)–(B.2), and scale the aggregate wealth, consumption, and asset positions of every original investor sector by* $1 - x_m$*. Then the original price processes, consumption–wealth ratios, portfolio shares, and conditional wealth distribution constitute an equilibrium of the extended economy. In particular, the valuation ratios* $(p_{Y,t}, p_{D,t})$*, the interest rate* r_t*, the market price of risk* η_t*, and all risk premia are unchanged for any path of the passive portfolio share* $\overline{\alpha}_{p,t}$*.*

Proof. Let $(C_{j,t}, W_{j,t}, Q_{j,D,t}, Q_{j,B,t}, A_{j,t})_{j\in\mathcal{J}}$ denote an equilibrium allocation in the original economy. Let $\omega_j^m = (1-\omega_m)\omega_j$ denote the population mass of original type j in the extended economy, and define

$$s_m \equiv \frac{1 - x_m}{1 - \omega_m}.$$

For each original investor type, define the allocation in the extended economy by

$$\begin{aligned} C_{j,t}^m &= s_m C_{j,t}, & W_{j,t}^m &= s_m W_{j,t}, \\ Q_{j,k,t}^m &= s_m Q_{j,k,t}, & A_{j,t}^m &= s_m A_{j,t}, \qquad k \in \{D, B\}. \end{aligned}$$

Recursive preferences are homothetic, so scaling consumption and wealth by the same factor leaves each original investor's consumption–wealth ratio and portfolio shares unchanged. Given the original price processes, the scaled allocation therefore satisfies the same optimality conditions and type-specific portfolio restrictions as the original allocation. Moreover, $\omega_j^m s_m = (1 - x_m)\omega_j$, so every original sector's aggregate allocation is scaled by $1 - x_m$.

Goods-market clearing follows from

$$\begin{aligned} \sum_{j\in\mathcal{J}} \omega_j^m C_{j,t}^m + \omega_m C_{m,t} &= (1 - x_m) Y_t + x_m Y_t \\ &= Y_t. \end{aligned}$$

For each risky claim $k \in \{D, B\}$, market clearing follows from

$$\begin{aligned} \sum_{j\in\mathcal{J}} \omega_j^m Q_{j,k,t}^m + \omega_m Q_{m,k,t} &= (1 - x_m) + x_m \\ &= 1. \end{aligned}$$

The risk-free-bond market continues to clear because the original positions are scaled proportionally and the hand-to-mouth sector holds no risk-free bonds:

$$\sum_{j\in\mathcal{J}} \omega_j^m A_{j,t}^m = (1 - x_m) \sum_{j\in\mathcal{J}} \omega_j A_{j,t} = 0.$$

Finally, Equation (B.3) verifies the hand-to-mouth sector's budget constraint. All equilibrium conditions are therefore satisfied at the original prices. □

Proposition 8 is an exact aggregation result. The hand-to-mouth sector removes the same fraction of aggregate consumption, wealth, and risky-claim supply from the original investor sectors. These effects cancel in goods and asset markets, even though the sector's consumption does not respond to changes in wealth or expected returns. Thus, optimal consumption adjustment by every investor is not necessary for any of the equilibrium results in Sections 3 and 4.

To compare economies with and without the hand-to-mouth sector, hold fixed the passive wealth share among the remaining investors and the process for $\overline{\alpha}_{p,t}$. Proposition 8 then implies that a given change in $\overline{\alpha}_{p,t}$ has exactly the same effect on prices, interest rates, and risk premia in both economies.[1]

The neutrality result parallels aggregation results in heterogeneous-agent macroeconomics. Werning (2015) shows that incomplete markets can leave aggregate consumption dynamics unchanged under benchmark restrictions on household exposures. Bilbiie (2008) shows that the aggregate-demand effects of limited asset-market participation depend on how income, particularly profits, is distributed between asset holders and hand-to-mouth households. The condition imposed here has the same logic: the hand-to-mouth sector's consumption exposure is exactly aligned with its ownership of aggregate payouts.

B.2 Departures from neutrality in the two-period model

The exact result relies on the alignment between the hand-to-mouth sector's consumption, wealth, and asset ownership. The two-period model allows us to relax this alignment and characterize when hand-to-mouth behavior amplifies price impact.

Environment. The economy is that of Section 2, with a third investor type m of mass ω_m added to the active (a) and passive (p) types. As before, initial wealth is $W_j = (P + Y)Q_{j,-1} > 0$, with $\sum_{j\in\{a,p,m\}} \omega_j Q_{j,-1} = 1$, and the budget constraints are unchanged. Type m follows a rule on both margins:

$$C_m = \overline{C}_m, \qquad \alpha_m = \overline{\alpha}_m \in [0, 1],$$

where $\overline{C}_m \in (0, W_m)$ is exogenous, so first-period consumption does not respond to $(p, r, \overline{\alpha}_p)$. The fixed portfolio share $\overline{\alpha}_m$ is applied to savings, so $Q_m = \overline{\alpha}_m(W_m - \overline{C}_m)/P$. Second-period consumption is the residual, $C'_m = Y'Q_m + B_m$. We interpret $\overline{C}_m$ as consumption out of current income by a household with positive asset holdings, in the spirit of the wealthy hand-to-mouth households in Kaplan and Violante (2014). The rule nests the case $\overline{C}_m = YQ_{m,-1}$, in which the

[1] The corresponding dollar flow relative to total market capitalization is smaller by the factor $1 - x_m$, because the hand-to-mouth sector does not participate in the reallocation. Normalizing flows by the wealth of the participating sectors removes this mechanical float effect.

household consumes exactly its current asset income. Active and passive investors are exactly as in Section 2, and market clearing adds type m to each condition.

Since initial wealth is proportional to $P + Y$ for every type, wealth shares are determined by initial holdings. The hand-to-mouth sector is therefore summarized by two constants,

$$x_m \equiv \frac{\omega_m W_m}{P+Y} = \omega_m Q_{m,-1}, \qquad \vartheta_m \equiv \frac{\omega_m \overline{C}_m}{Y},$$

which are its shares of aggregate wealth and aggregate consumption. The inequality $\vartheta_m > x_m$ holds if and only if $\overline{C}_m/W_m > Y/(P+Y)$, so the hand-to-mouth sector then consumes a larger fraction of its wealth than the economy as a whole.

We impose Assumption 1 on the optimizing types, with wealth weights normalized within that group. Letting $\tilde{x}_j \equiv \omega_j W_j/(\omega_a W_a + \omega_p W_p)$ for $j \in \{a, p\}$, we assume

$$\tilde{x}_a\, \overline{c}_a(r,\pi) + \tilde{x}_p\, \overline{c}_p(r,\pi) = \overline{c}(r+\pi). \tag{B.4}$$

The goods market. The following lemma generalizes Equation (3).

Lemma B.1. *Under Assumption 1 applied to the optimizing types, goods-market clearing is equivalent to*

$$(1 - x_m)\, \overline{c}(\mu - p) \;=\; \frac{1-\vartheta_m}{1+e^p}. \tag{B.5}$$

Proof. Goods-market clearing for the optimizing types is

$$\omega_a C_a + \omega_p C_p = Y - \omega_m \overline{C}_m = (1-\vartheta_m)Y.$$

Equation (B.4) implies

$$\omega_a C_a + \omega_p C_p = \overline{c}(r+\pi)(\omega_a W_a + \omega_p W_p).$$

Summing wealth across all types gives

$$\sum_j \omega_j W_j = (P+Y) \sum_j \omega_j Q_{j,-1} = P + Y.$$

It follows that $\omega_a W_a + \omega_p W_p = (1 - x_m)(P+Y)$. Substituting $r + \pi = \mu - p$ from Equation (1) and using $P + Y = (1 + e^p)Y$ yields Equation (B.5). □

Setting $x_m = \vartheta_m = 0$ recovers Equation (3). The separation established in Section 2 is preserved: the goods market determines the price-dividend ratio, while the risky-asset market determines the risk premium. The hand-to-mouth sector removes a fraction x_m of the wealth behind the consumption-wealth schedule and a fraction ϑ_m of output from the resource side.

Hand-to-mouth neutrality. An important benchmark arises when wealthy hand-to-mouth households consume the dividends on their asset holdings.

Corollary 1 (Dividends-consumption neutrality)**.** *Suppose*

$$\overline{C}_m = YQ_{m,-1}. \tag{B.6}$$

Then $\vartheta_m = x_m$, and the goods-market condition in Equation (B.5) reduces exactly to the baseline condition in Equation (3). Consequently, the hand-to-mouth sector leaves unchanged the equilibrium response of the price-dividend ratio, and hence aggregate market elasticity, for any shock to the consumption-wealth schedule.

Proof. Equation (B.6) implies

$$\vartheta_m = \frac{\omega_m \overline{C}_m}{Y} = \omega_m Q_{m,-1} = x_m.$$

Substituting $\vartheta_m = x_m$ into Equation (B.5) gives

$$(1 - x_m)\overline{c}(\mu - p) = \frac{1 - x_m}{1 + e^p}.$$

Dividing by $1 - x_m$ recovers Equation (3). Because the equation determining p is identical to the baseline equation, its response to any shifter of the consumption-wealth schedule is also identical. □

Corollary 1 is an aggregation result. Although hand-to-mouth households do not adjust their consumption in response to prices, they remove the same fraction of aggregate wealth from the price-sensitive consumption schedule as they remove from aggregate consumption on the resource side. The two effects cancel, so heterogeneity in consumption behavior alone does not amplify price movements. The hand-to-mouth sector matters for aggregate market elasticity only when its consumption share differs from its wealth share.

This result parallels neutrality results in heterogeneous-agent macroeconomics. Werning (2015) shows that incomplete markets can leave the aggregate sensitivity of consumption to interest rates unchanged under benchmark restrictions on household exposures. Bilbiie (2008) shows in a two-agent New Keynesian model that the aggregate-demand effects of limited asset-market participation depend on the distribution of income, particularly profits, between asset holders and hand-to-mouth households. The condition $\vartheta_m = x_m$ plays the analogous role here: the presence of hand-to-mouth households is not sufficient for amplification when their exposure to aggregate resources is aligned with their asset holdings.

This neutrality concerns the price-dividend ratio and aggregate market elasticity. The equilibrium movements in the interest rate and the risk premium can still differ from the baseline because the hand-to-mouth sector changes the demand elasticities entering risky-asset market clearing.

The linearized equilibrium. We linearize around the reference equilibrium (p^*, r^*) as in Section 2. The local slopes of the dividend-yield and consumption-wealth schedules are

$$d_p \equiv \frac{e^{p^*}}{(1 + e^{p^*})^2} > 0, \qquad \chi_p \equiv -\overline{c}\,'(\mu - p^*) > 0.$$

The hand-to-mouth demand for the risky asset is $Q_m = \overline{\alpha}_m(W_m - \overline{C}_m)/P$ with $W_m = (P+Y)Q_{m,-1}$, so

$$\zeta_{m,p} \equiv -\frac{\partial \log Q_m}{\partial p} = 1 - \frac{P\,Q_{m,-1}}{W_m - \overline{C}_m} = \frac{YQ_{m,-1} - \overline{C}_m}{W_m - \overline{C}_m}, \qquad \zeta_{m,r} = 0, \qquad f_m = 0.$$

The cross-price elasticity is zero, while the own-price elasticity depends on saving out of current asset income relative to total saving. A household that consumes exactly its asset income has $\zeta_{m,p} = 0$. None of the results below depends on the value of $\zeta_{m,p}$. With quantity shares $s_j \equiv \omega_j Q_j^*$, the linearized risky-asset market-clearing condition generalizes Equation (5) to

$$s_a q_a + s_p q_p + s_m q_m = 0, \qquad q_j = -\zeta_{j,p}\hat{p} - \zeta_{j,r}\hat{r} + f_j. \tag{B.7}$$

Define the aggregate flow as $f \equiv \sum_{j\in\{a,p,m\}} s_j f_j$. We allow a shock to shift the optimizers' consumption-wealth ratio according to $\hat{c} = \chi_0 + \chi_p \hat{p}$. The direct shifter is $\chi_0 = 0$ for the portfolio-flow shock in the benchmark model and $\chi_0 = \chi_\mu \hat{\mu} = -\chi_p \hat{\mu}$ for the fundamental shock. Linearizing Equation (B.5) gives

$$(1 - x_m)\left(\chi_0 + \chi_p \hat{p}\right) = -(1 - \vartheta_m) d_p \hat{p}. \tag{B.8}$$

Results. The first result shows that the infinite-elasticity benchmark of Proposition 1 survives.

Proposition 9 (Neutrality)**.** *Suppose the aggregate cross-price elasticity*

$$\zeta_r \equiv s_a \zeta_{a,r} + s_p \zeta_{p,r} + s_m \zeta_{m,r}$$

is nonzero. Then a portfolio-flow shock has no price impact in the presence of a wealthy hand-to-mouth sector:

$$\hat{p} = 0, \qquad \hat{r} = \frac{f}{\zeta_r}, \qquad \hat{\pi} = -\frac{f}{\zeta_r}.$$

Proof. With $\chi_0 = 0$, Equation (B.8) becomes

$$\left[(1 - x_m)\chi_p + (1 - \vartheta_m) d_p\right]\hat{p} = 0.$$

The term in brackets is strictly positive, so $\hat{p} = 0$. Substituting $\hat{p} = 0$ into Equation (B.7) gives $\hat{r} = f/\zeta_r$. The identity $\hat{\pi} = -\hat{p} - \hat{r}$ then implies $\hat{\pi} = -f/\zeta_r$. □

Proposition 9 shows that the infinite-elasticity result does not require every household to make an optimal consumption-saving decision. It requires the portfolio-flow shock to leave unchanged the part of aggregate consumption that responds to prices. Hand-to-mouth consumption satisfies this condition because it does not respond to the flow or to equilibrium prices. Relative to Section 2, hand-to-mouth asset holdings absorb a share s_m of the market while contributing no cross-price elasticity. The aggregate cross-price elasticity is therefore smaller, and the interest rate must move by more to absorb a given flow. Hand-to-mouth behavior does not generate price impact by itself, but it can amplify price impact generated by a shift in consumption-saving behavior.

Proposition 10 (Amplification)**.** *Consider a shock that shifts the optimizers' consumption-wealth ratio by $\chi_0 \neq 0$. The price response is*

$$\hat{p} \;=\; -\frac{(1 - x_m)\,\chi_0}{(1 - x_m)\chi_p + (1 - \vartheta_m) d_p} \;=\; \mathcal{A}(x_m, \vartheta_m)\;\hat{p}_0, \qquad \hat{p}_0 \equiv -\frac{\chi_0}{\chi_p + d_p},$$

where $\hat{p}_0$ is the response in the economy without hand-to-mouth consumers and

$$\mathcal{A}(x_m, \vartheta_m) \equiv \frac{(1-x_m)(\chi_p + d_p)}{(1-x_m)\chi_p + (1-\vartheta_m)d_p}.$$

Proof. Immediate from Equation (B.8); the benchmark corresponds to $x_m = \vartheta_m = 0$. □

Corollary 2. *(i) $\mathcal{A} > 1$ if and only if $\vartheta_m > x_m$, that is, if and only if the hand-to-mouth sector's consumption-wealth ratio exceeds the economy-wide average; $\mathcal{A} = 1$ when $\vartheta_m = x_m$. (ii) $\mathcal{A}$ is increasing in ϑ_m and decreasing in x_m. (iii) For any $x_m \in [0, 1)$,*

$$\sup_{\vartheta_m \leq 1} \mathcal{A}(x_m, \vartheta_m) = 1 + \frac{d_p}{\chi_p},$$

independent of x_m. Under the micro-founded demand system of Section 4, where $\chi_p = (\psi - 1)\, d_p$, the bound is $\psi/(\psi - 1)$.

Proof. (i) $\mathcal{A} > 1$ is equivalent to $(1 - x_m)d_p > (1 - \vartheta_m)d_p$, that is, $\vartheta_m > x_m$; the second statement follows from the definitions of ϑ_m and x_m. (ii) With $D \equiv (1 - x_m)\chi_p + (1 - \vartheta_m)d_p$, direct differentiation gives $\partial\mathcal{A}/\partial\vartheta_m = (1 - x_m)(\chi_p + d_p)d_p/D^2 > 0$ and $\partial\mathcal{A}/\partial x_m = -(\chi_p + d_p)(1 - \vartheta_m)d_p/D^2 < 0$. (iii) Set $\vartheta_m = 1$. □

The mechanism follows directly from goods-market clearing. The hand-to-mouth sector removes wealth from the price-sensitive consumption schedule in proportion to x_m and output from the resource side in proportion to ϑ_m. These effects cancel when $\vartheta_m = x_m$. When $\vartheta_m > x_m$, the resource side contracts by more, so a given shift in saving incentives generates a larger price response. This amplification result concerns departures from the neutrality benchmark; it does not alter the exact dynamic neutrality result established above. The hand-to-mouth and passive sectors may overlap empirically, but modeling them separately isolates the consumption margin from the portfolio-flow margin.

C Production and Monetary Policy

This appendix embeds the baseline model in a production economy with sticky prices and monetary policy. We first show that the flexible-price real allocation is equivalent to the baseline endowment economy. We then show that an inflation-targeting central bank can implement this allocation by tracking its natural rate of interest.

C.1 The flexible-price economy

Firms. Final output is a CES aggregate of differentiated goods produced by monopolistically competitive firms $i \in [0, 1]$. Firm i produces

$$\widetilde{Y}_{i,t} = A_t L_{i,t}^{1-\alpha},$$

Productivity follows

$$\frac{dA_t}{A_t} = \mu\, dt + \sigma\, dZ_t, \qquad \sigma \in \mathbb{R}^{1\times d}.$$

The productivity process therefore has the same fundamental exposure as aggregate endowment in the baseline model.

Demand for good i is

$$\widetilde{Y}_{i,t} = \left(\frac{P_{i,t}}{P_t}\right)^{-\varepsilon_g} \widetilde{Y}_t, \tag{C.1}$$

where P_t is the aggregate price level and $\varepsilon_g > 1$ is the elasticity of substitution across goods. Under flexible prices, firm i chooses its price to maximize real profits,

$$\frac{P_{i,t}}{P_t}\widetilde{Y}_{i,t} - \frac{\mathcal{W}_t}{P_t}\left(\frac{\widetilde{Y}_{i,t}}{A_t}\right)^{\frac{1}{1-\alpha}},$$

subject to Equation (C.1). In a symmetric equilibrium, the resulting labor-demand condition is

$$\frac{\mathcal{W}_t}{P_t} = (1-\varepsilon_g^{-1})(1-\alpha)A_t L_t^{-\alpha}. \tag{C.2}$$

Households. Following Bigio, Silva, and Zilberman (2025), households have Epstein–Zin preferences over a GHH composite of consumption and labor. Type $j \in \mathcal{J}$ chooses gross consumption $\widetilde{C}_{j,t}$, labor $L_{j,t}$, and its feasible portfolio $\alpha_{j,t} \in \Omega_{j,t}$ to solve

$$V_{j,t} = \max_{\widetilde{C}_{j,t}, L_{j,t}, \alpha_{j,t}} \mathbb{E}_t\left[\int_t^\infty f_j\left(\widetilde{C}_{j,s} - \xi A_s L_{j,s}, V_{j,s}\right) ds\right],$$

subject to

$$\frac{dW_{j,t}}{W_{j,t}} = \left[i_t - \Pi_t + \pi_t^\top \alpha_{j,t} - \frac{\widetilde{C}_{j,t}}{W_{j,t}} + \frac{\mathcal{W}_t}{P_t}\frac{L_{j,t}}{W_{j,t}}\right] dt + \sigma_{j,t} dZ_t.$$

Here, i_t is the nominal interest rate, Π_t is inflation, and π_t denotes the vector of excess returns on equity and corporate debt. The feasible portfolio sets $\Omega_{j,t}$ are the same as in Section 3.

The intratemporal first-order condition for labor is

$$\frac{\mathcal{W}_t}{P_t} = \xi A_t. \tag{C.3}$$

Define net consumption by

$$C_{j,t} \equiv \widetilde{C}_{j,t} - \xi A_t L_{j,t}.$$

Using Equation (C.3), the household budget constraint becomes

$$\frac{dW_{j,t}}{W_{j,t}} = \left[r_t + \pi_t^\top \alpha_{j,t} - \frac{C_{j,t}}{W_{j,t}}\right] dt + \sigma_{j,t} dZ_t, \qquad r_t \equiv i_t - \Pi_t.$$

Thus, expressed in net consumption, the household problem is exactly the problem in Section 3.

Labor and goods markets. Combining Equations (C.2) and (C.3) gives constant aggregate labor,

$$L_t = \left[\frac{(1-\varepsilon_g^{-1})(1-\alpha)}{\xi}\right]^{1/\alpha} \equiv \overline{L}.$$

Goods-market clearing in net units is

$$\sum_{j\in\mathcal{J}} \omega_j C_{j,t} = Y_t,$$

where net output is

$$\begin{aligned} Y_t &\equiv A_t \overline{L}^{1-\alpha} - \xi A_t \overline{L} \\ &= \left[\alpha + \frac{1-\alpha}{\varepsilon_g}\right] A_t \overline{L}^{1-\alpha}. \end{aligned}$$

The second equality uses the equilibrium labor condition. Consequently,

$$\frac{dY_t}{Y_t} = \mu\, dt + \sigma\, dZ_t. \tag{C.4}$$

Financial equilibrium and the natural rate. Firms distribute net output through the equity and corporate-debt claims defined in Section 3. The aggregate value of these claims is $P_{Y,t}$, and households can also trade the risk-free asset. Because the household budget constraints in net-consumption units and the process for net output coincide with those in the baseline model, all normalized real equilibrium objects solve the same system as in Sections 3 and 4. In particular, the flexible-price price-dividend ratio, risk premium, and aggregate market elasticity are identical to their baseline counterparts.

Let r_t^n denote the equilibrium risk-free rate in this flexible-price economy. This is the natural rate of interest. It incorporates the endogenous adjustment to portfolio-flow shocks characterized in the main text.

C.2 Sticky prices and monetary policy

Price setting. Suppose firm i chooses its inflation rate $\Pi_{i,t}$ subject to

$$\frac{dP_{i,t}}{P_{i,t}} = \Pi_{i,t} dt.$$

Changing prices entails the quadratic Rotemberg cost $\frac{\varphi}{2}\Pi_{i,t}^2 \widetilde{Y}_t$. The cost is rebated lump-sum to households and therefore does not absorb aggregate resources.

Let Λ_t denote the real stochastic discount factor defined in Section 3. Taking aggregate prices

and quantities as given, firm i solves

$$\mathcal{V}_{i,t}^{F} = \max_{\{\Pi_{i,s}\}_{s\geq t}} \mathbb{E}_t \int_t^\infty \frac{\Lambda_s}{\Lambda_t}\left[\frac{P_{i,s}}{P_s}\widetilde{Y}_{i,s} - \frac{\mathcal{W}_s}{P_s}\left(\frac{\widetilde{Y}_{i,s}}{A_s}\right)^{\frac{1}{1-\alpha}} - \frac{\varphi}{2}\Pi_{i,s}^2\widetilde{Y}_s\right]ds,$$

subject to Equation (C.1) and the law of motion for $P_{i,t}$.

To derive the pricing condition, let $u_{i,t} \equiv \log P_{i,t}$ and let $\mathcal{K}_{i,t}$ denote the costate associated with $u_{i,t}$. The first-order condition for $\Pi_{i,t}$ is

$$\mathcal{K}_{i,t} = \varphi\Pi_{i,t}\widetilde{Y}_t. \tag{C.5}$$

Define firm i's real marginal cost as

$$mc_{i,t} \equiv \frac{\mathcal{W}_t/P_t}{(1-\alpha)A_t L_{i,t}^{-\alpha}}.$$

Differentiating operating profits with respect to $u_{i,t}$ and then imposing symmetry gives

$$\left.\frac{\partial\Pi_{i,t}}{\partial u_{i,t}}\right|_{P_{i,t}=P_t} = \varepsilon_g\widetilde{Y}_t\left(mc_t - mc^*\right), \qquad mc^* \equiv \frac{\varepsilon_g - 1}{\varepsilon_g}. \tag{C.6}$$

The stochastic costate equation is

$$0 = \left.\frac{\partial\Pi_{i,t}}{\partial u_{i,t}}\right|_{P_{i,t}=P_t} dt + \frac{\mathbb{E}_t\left[d\left(\Lambda_t\mathcal{K}_{i,t}\right)\right]}{\Lambda_t}. \tag{C.7}$$

Integrating Equation (C.7) forward, imposing the transversality condition, and using Equations (C.5) and (C.6) gives

$$\varphi\Pi_t\widetilde{Y}_t = \mathbb{E}_t\left[\int_t^\infty \frac{\Lambda_s}{\Lambda_t}\varepsilon_g\widetilde{Y}_s\left(mc_s - mc^*\right)ds\right]. \tag{C.8}$$

Equation (C.8) is the nonlinear forward-looking Phillips curve in this economy. Current inflation reflects the discounted path of deviations of real marginal cost from its flexible-price level. In particular, $mc_t = mc^*$ at all dates implies $\Pi_t = 0$.

Monetary policy. Aggregate inflation is

$$\Pi_t \equiv \frac{\dot{P}_t}{P_t}.$$

The central bank sets the nominal interest rate according to

$$i_t = r_t^n + \phi_\Pi\Pi_t, \qquad \phi_\Pi > 1. \tag{C.9}$$

The following proposition establishes the implementation result used in the main text.

Proposition 11 (Implementation of the flexible-price allocation)**.** *Suppose all firms begin with the same price. Under the policy rule in Equation* (C.9)*, the flexible-price allocation with zero inflation is an equilibrium of the sticky-price economy. In this equilibrium,* $i_t = r_t^n$*, and aggregate market elasticity is identical to that in the baseline model.*

Proof. At the flexible-price allocation, Equations (C.2) and (C.3) imply that real marginal cost equals mc^*. The firms' pricing condition is therefore satisfied at $\Pi_t = 0$, and Rotemberg adjustment costs vanish. With zero inflation, the real and nominal risk-free rates coincide. Equation (C.9) then implies $i_t = r_t^n$, so households face the same real interest rate as in the flexible-price economy. All remaining household, firm, financial-market, and goods-market conditions consequently coincide with those of the flexible-price allocation. The process for net output remains Equation (C.4), so the price-dividend ratio and its response to portfolio flows are unchanged. □

Proposition 11 establishes the existence of a zero-inflation equilibrium in which an inflation-targeting central bank implements the interest-rate adjustment characterized in the baseline model; it does not require a claim of global uniqueness.

D Belief Distortions from Portfolio Flows

This appendix studies a belief distortion through which non-fundamental portfolio flows affect investors' expectations of future cash-flow growth. The distortion can be interpreted as investors partially confusing portfolio-flow shocks with fundamental news, or more broadly as reduced-form mislearning from flow-induced price movements. We first introduce the mechanism in the simple model of Section 2 and then embed it in the dynamic model.

D.1 Extension of the simple model

We allow a portfolio-flow shock to shift investors' perceived expected dividend growth without changing actual cash flows or aggregate resources. This belief distortion makes aggregate market elasticity finite because the flow now shifts perceived saving incentives.

Perceived cash-flow news. Let $\hat{\mu}$ denote the true change in expected dividend growth and let f denote the portfolio flow defined in Section 2. Investors behave as if expected dividend growth had changed by

$$\widetilde{\mu} = \hat{\mu} + \theta f, \tag{D.1}$$

where $\theta \geq 0$ measures the degree of confusion. When $\theta = 0$, investors correctly distinguish portfolio flows from cash-flow news, and the baseline model is recovered. When $\theta > 0$, a flow into the risky asset is partly interpreted as positive news about its future payouts.

The distortion in Equation (D.1) applies to investors' portfolio and consumption-saving decisions. Actual current dividends and aggregate resources remain unchanged. Thus, the linearized demand system becomes

$$q_j = -\zeta_{j,p}\hat{p} - \zeta_{j,r}\hat{r} + \zeta_{j,\mu}\widetilde{\mu} + f_j,$$

where

$$\zeta_{j,\mu} \equiv \frac{\partial \log \overline{Q}_j}{\partial \mu}.$$

The change in the average consumption-wealth ratio is

$$\hat{c} = \chi_\mu \widetilde{\mu} + \chi_p \hat{p}, \qquad \chi_\mu = -\chi_p. \tag{D.2}$$

Equilibrium price impact. Goods-market clearing combines Equation (D.2) with the linearized dividend-yield schedule:

$$\chi_\mu \widetilde{\mu} + \chi_p \hat{p} = -d_p \hat{p}. \tag{D.3}$$

Define

$$\lambda_\mu \equiv -\frac{\chi_\mu}{\chi_p + d_p} = \frac{\chi_p}{\chi_p + d_p}.$$

The coefficient λ_μ is the response of the price-dividend ratio to perceived cash-flow news derived in Equation (8).

Proposition 12 (Price impact under flow confusion)**.** *Suppose investors' beliefs satisfy Equation* (D.1)*. In response to a pure portfolio-flow shock, $\hat{\mu} = 0$, the equilibrium price response is*

$$\hat{p} = \lambda_\mu \theta f. \tag{D.4}$$

Consequently, inverse aggregate market elasticity is

$$\mathcal{E}_M^{-1} = \lambda_\mu \theta = \frac{\chi_p}{\chi_p + d_p}\theta. \tag{D.5}$$

Proof. Substituting $\widetilde{\mu} = \theta f$ into Equation (D.3) gives Equation (D.4). Equation (D.5) then follows from the definition of aggregate market elasticity in Equation (7). □

The result isolates the relevant economic channel. Confusion does not make the price response depend on the own-price elasticities of individual demand curves. Instead, it gives the portfolio flow a perceived fundamental component, which shifts the consumption-wealth schedule and therefore the price-dividend ratio selected by goods-market clearing. The individual demand elasticities determine the accompanying adjustment in the interest rate and the risk premium. In particular, defining

$$\zeta_p \equiv \sum_j s_j \zeta_{j,p}, \qquad \zeta_r \equiv \sum_j s_j \zeta_{j,r}, \qquad \zeta_\mu \equiv \sum_j s_j \zeta_{j,\mu},$$

risky-asset market clearing implies

$$\hat{r} = \frac{1 + \zeta_\mu \theta - \zeta_p \lambda_\mu \theta}{\zeta_r} f,$$

provided $\zeta_r \neq 0$.

The strength of the consumption-saving response disciplines the magnitude of this mechanism.

Under the micro-founded relation $\chi_p = (\psi - 1)d_p$ used in the dynamic model,

$$\mathcal{E}_M^{-1} = \left(1 - \frac{1}{\psi}\right)\theta.$$

Thus, confusion has no price effect under unit EIS, because perceived cash-flow news does not shift the consumption-wealth ratio in that case.

Feedback from prices to perceived news. Equation (D.1) captures the initial misclassification of a flow shock. A natural extension allows the resulting price movement itself to reinforce investors' inference about fundamentals:

$$\widetilde{\mu} = \hat{\mu} + \theta f + \kappa_b \hat{p}, \tag{D.6}$$

where $\kappa_b \geq 0$ measures the strength of the feedback. Combining Equations (D.3) and (D.6) gives

$$\hat{p} = \frac{\lambda_\mu}{1 - \kappa_b \lambda_\mu} \left(\hat{\mu} + \theta f\right),$$

provided $\kappa_b \lambda_\mu < 1$. For a pure flow shock,

$$\mathcal{E}_M^{-1} = \frac{\lambda_\mu \theta}{1 - \kappa_b \lambda_\mu}.$$

The direct confusion in Proposition 12 initiates the price response, while inference from prices amplifies it through a geometric feedback multiplier. If perceived news depended only on $\hat{p}$, without the direct term θf, the baseline solution $\hat{p} = 0$ would remain an equilibrium whenever $\kappa_b \lambda_\mu < 1$. This is why confusion about the source of the initial innovation, rather than price extrapolation alone, is necessary to generate price impact locally around the benchmark. The quantitative model below includes the direct-confusion channel but not this additional price-feedback mechanism.

D.2 Extension of the dynamic model

We next introduce the same form of confusion into the dynamic model of Section 3. To keep the extension tractable, all investors share the same distorted beliefs. The objective processes for aggregate endowment, dividends, and the passive portfolio share remain unchanged.

A belief state. Recall that $Z_t = (Z_{1,t}, Z_{2,t})$, where $Z_{1,t}$ is the fundamental innovation and $Z_{2,t}$ is the non-fundamental portfolio-flow innovation. Define the non-fundamental innovation to the passive portfolio share as

$$dF_t \equiv \sigma_{\alpha_p,2}\, dZ_{2,t}.$$

We introduce a belief state m_t that follows

$$dm_t = -\kappa_m m_t\, dt + \theta\, dF_t,$$

where $\kappa_m > 0$ governs the persistence of distorted beliefs and $\theta \geq 0$ measures the extent to which investors interpret a non-fundamental flow innovation as cash-flow news. A positive realization of

$dZ_{2,t}$ raises perceived expected cash-flow growth, even though it does not change current endowment or dividends. Subsequent observations gradually undo this inference at rate κ_m.

This formulation captures confusion about news regarding future payouts rather than confusion about observed current payouts. It therefore preserves the objective cash-flow processes

$$\frac{dY_t}{Y_t} = \mu\, dt + \sigma_1\, dZ_{1,t}, \qquad \frac{dD_t}{D_t} = \mu\, dt + \phi\sigma_1\, dZ_{1,t}.$$

The belief state affects decisions but does not alter aggregate resources.

Subjective beliefs. Let $\mathbb{P}$ denote the objective probability measure and let $\widetilde{\mathbb{P}}$ denote investors' common subjective probability measure. Define the belief distortion

$$h_t \equiv \left(\frac{m_t}{\sigma_1}, 0\right),$$

and the subjective Brownian motion

$$d\widetilde{Z}_t = dZ_t - h_t^\top\, dt. \tag{D.7}$$

Under the usual integrability condition, $\widetilde{Z}_t$ is a Brownian motion under $\widetilde{\mathbb{P}}$. Rewriting the endowment process under investors' beliefs gives

$$\frac{dY_t}{Y_t} = (\mu + m_t)\, dt + \sigma_1\, d\widetilde{Z}_{1,t}. \tag{D.8}$$

Similarly, investors perceive expected dividend growth to be $\mu + \phi m_t$:

$$\frac{dD_t}{D_t} = (\mu + \phi m_t)\, dt + \phi\sigma_1\, d\widetilde{Z}_{1,t}. \tag{D.9}$$

Thus, a flow innovation is interpreted as news about expected aggregate growth, and the cash-flow leverage of equity magnifies its perceived effect on dividends. Beliefs about the residual corporate-debt payout $B_t = Y_t - D_t$ are inherited from Equations (D.8) and (D.9).

Consider an asset with objective return dynamics

$$dR_{k,t} = \mu_{R_k,t}\, dt + \sigma_{R_k,t}\, dZ_t.$$

Under investors' beliefs, the same observed return process is written as

$$dR_{k,t} = \left(\mu_{R_k,t} + \sigma_{R_k,t} h_t^\top\right) dt + \sigma_{R_k,t}\, d\widetilde{Z}_t.$$

Consequently, the vector of perceived risk premia is

$$\widetilde{\pi}_t = \pi_t + \Sigma_t h_t^\top = \Sigma_t \widetilde{\eta}_t^\top, \qquad \widetilde{\eta}_t \equiv \eta_t + h_t.$$

The vector η_t continues to describe the vector of objective market prices of risk, while $\widetilde{\eta}_t$ describes the compensation for risk perceived by investors.

Investors' problems. Investor j evaluates utility under the subjective measure and solves

$$V_{j,t} = \max_{C_{j,t},\alpha_{j,t}} \widetilde{\mathbb{E}}_t \left[\int_t^\infty f_j(C_{j,s}, V_{j,s})\, ds \right]$$

subject to the same type-specific investment set $\Omega_{j,t}$ as in the baseline model. Under the subjective measure, the investor perceives the wealth dynamics as

$$\frac{dW_{j,t}}{W_{j,t}} = \left[r_t + \widetilde{\pi}_t^\top \alpha_{j,t} - \frac{C_{j,t}}{W_{j,t}} \right] dt + \sigma_{j,t}\, d\widetilde{Z}_t,$$

where $\sigma_{j,t} = \alpha_{j,t}^\top \Sigma_t$. Passive investors continue to follow their prescribed portfolio rule, but their consumption decision reflects the same subjective beliefs as those of active investors.

The optimal risk exposure of an active investor becomes

$$\sigma_{j,t} = \frac{1}{1+\nu_{j,t}} \left[\frac{\widetilde{\eta}_t}{\gamma_j} + \frac{1-\gamma_j^{-1}}{\psi - 1} \sigma_{c_j,t} \right], \qquad j \in \{h, \ell\}. \tag{D.10}$$

The change of measure affects conditional drifts but not diffusions, so $\sigma_{c_j,t}$ is also the diffusion of the consumption-wealth ratio under investors' beliefs. The corresponding HJB condition is

$$c_{j,t} = \psi\rho + (1-\psi) \left[r_t + \widetilde{\eta}_t \sigma_{j,t}^\top - \frac{\gamma_j}{2} \|\sigma_{j,t}\|^2 \right] + \widetilde{\xi}_{j,t}.$$

The continuation-value term is

$$\widetilde{\xi}_{j,t} = \widetilde{\mu}_{c_j,t} + (1-\gamma_j)\sigma_{c_j,t}\sigma_{j,t}^\top + \frac{\psi - \gamma_j}{1-\psi} \frac{\|\sigma_{c_j,t}\|^2}{2}, \qquad \widetilde{\mu}_{c_j,t} = \mu_{c_j,t} + \sigma_{c_j,t} h_t^\top.$$

Equilibrium. The state vector expands to

$$\widetilde{X}_t = \left(x_{h,t}, x_{\ell,t}, \overline{\alpha}_{p,t}, m_t \right).$$

An equilibrium consists of prices and allocations such that investors optimize under $\widetilde{\mathbb{P}}$, while all goods and asset markets clear under the objective resource constraints. The market-clearing conditions are unchanged:

$$\sum_{j\in\mathcal{J}} x_{j,t} c_{j,t} = \frac{1}{p_{Y,t}}, \qquad \sum_{j\in\mathcal{J}} x_{j,t}\sigma_{j,t} = \sigma_{R_Y,t}. \tag{D.11}$$

Actual wealth shares evolve under $\mathbb{P}$ according to realized returns and consumption, not according to the returns investors expected under $\widetilde{\mathbb{P}}$. Hence, the objective law of motion remains

$$\begin{aligned} \frac{dx_{j,t}}{x_{j,t}} = &\left[(\eta_t - \sigma_{R_Y,t})(\sigma_{j,t} - \sigma_{R_Y,t})^\top - \left(c_{j,t} - \frac{1}{p_{Y,t}} \right) + \kappa \frac{\omega_j - x_{j,t}}{x_{j,t}} \right] dt \\ &+ (\sigma_{j,t} - \sigma_{R_Y,t})\, dZ_t, \qquad j \in \{h, \ell\}. \end{aligned} \tag{D.12}$$

Belief distortions affect realized wealth redistribution through the portfolios and consumption choices they induce, but they do not mechanically alter realized cash flows.

Setting $\theta = 0$ and $m_0 = 0$ recovers the baseline model. For $\theta > 0$, a non-fundamental portfolio-flow innovation changes both passive demand and investors' perceived investment opportunities. The latter channel shifts consumption-saving decisions and can therefore generate additional price impact through goods-market clearing.

D.2.1 The equilibrium system

We now collect the equations needed to compute the behavioral equilibrium. The central distinction is that optimal decisions use subjective conditional expectations, while market clearing, realized wealth accumulation, and the valuation equations are expressed under the objective probability measure.

Objective and subjective state generators. Write the objective law of motion of the expanded state vector as

$$d\widetilde{X}_t = \mu_{\widetilde{X},t}\, dt + \sigma_{\widetilde{X},t}\, dZ_t. \tag{D.13}$$

Under investors' beliefs, the same process satisfies

$$d\widetilde{X}_t = \widetilde{\mu}_{\widetilde{X},t}\, dt + \sigma_{\widetilde{X},t}\, d\widetilde{Z}_t, \qquad \widetilde{\mu}_{\widetilde{X},t} = \mu_{\widetilde{X},t} + \sigma_{\widetilde{X},t} h_t^\top.$$

Hence, for any twice continuously differentiable function $g(\widetilde{X}_t)$, the objective and subjective infinitesimal generators are

$$\begin{aligned}
\mathcal{L}g &= g_X \mu_{\widetilde{X}} + \frac{1}{2} \operatorname{tr}\left(\sigma_{\widetilde{X}} \sigma_{\widetilde{X}}^\top g_{XX}\right), \\
\widetilde{\mathcal{L}}g &= \mathcal{L}g + g_X \sigma_{\widetilde{X}} h^\top.
\end{aligned}$$

The diffusion is identical under the two measures. Only the drift changes.

For a positive function g, define its objective drift and diffusion by

$$\frac{dg_t}{g_t} = \mu_{g,t}\, dt + \sigma_{g,t}\, dZ_t.$$

Itô's lemma gives

$$\mu_{g,t} = \frac{g_X}{g} \mu_{\widetilde{X},t} + \frac{1}{2g} \operatorname{tr}\left(\sigma_{\widetilde{X},t} \sigma_{\widetilde{X},t}^\top g_{XX}\right), \tag{D.14}$$

$$\sigma_{g,t} = \frac{g_X}{g} \sigma_{\widetilde{X},t}, \tag{D.15}$$

$$\widetilde{\mu}_{g,t} = \mu_{g,t} + \sigma_{g,t} h_t^\top. \tag{D.16}$$

Equations (D.14)–(D.16) provide the derivatives entering the subjective HJB equations and the objective valuation equations.

Market price of risk. Let $x_{a,t} = x_{h,t} + x_{\ell,t} = 1 - x_{p,t}$ and define

$$\chi_{a,t} \equiv \frac{1 - x_{p,t}\overline{\alpha}_{p,t}}{x_{a,t}}.$$

Risk-market clearing requires

$$\sum_{j\in\{h,\ell\}} \frac{x_{j,t}}{x_{a,t}}\sigma_{j,t} = \chi_{a,t}\sigma_{R_Y,t}. \tag{D.17}$$

Define average effective risk aversion and average hedging demand as

$$\gamma_{a,t} \equiv \left[\sum_{j\in\{h,\ell\}} \frac{x_{j,t}}{x_{a,t}} \frac{1}{\gamma_j(1+\nu_{j,t})}\right]^{-1},$$
$$\mathcal{H}_{a,t} \equiv \sum_{j\in\{h,\ell\}} \frac{x_{j,t}}{x_{a,t}} \frac{1-\gamma_j^{-1}}{\psi - 1} \frac{\sigma_{c_j,t}}{1+\nu_{j,t}}.$$

Substituting Equation (D.10) into Equation (D.17) gives the perceived market price of risk:

$$\widetilde{\eta}_t = \gamma_{a,t}\left[\chi_{a,t}\sigma_{R_Y,t} - \mathcal{H}_{a,t}\right].$$

The objective market price of risk is therefore

$$\eta_t = \gamma_{a,t}\left[\chi_{a,t}\sigma_{R_Y,t} - \mathcal{H}_{a,t}\right] - h_t. \tag{D.18}$$

Relative to the baseline model, risk-market clearing determines the compensation investors perceive. Optimism about future cash flows lowers the objective compensation required to induce them to hold the outstanding risk.

Interest rate. Define

$$\widetilde{\xi}_t \equiv \sum_{j\in\mathcal{J}} x_{j,t}\widetilde{\xi}_{j,t}, \qquad \mathcal{R}_t \equiv \sum_{j\in\mathcal{J}} x_{j,t}\frac{\gamma_j}{2}\|\sigma_{j,t}\|^2,$$

and let

$$b_t \equiv \sigma_{R_Y,t}h_t^\top = \widetilde{\pi}_{Y,t} - \pi_{Y,t}.$$

Wealth-weighting the subjective HJB equations and imposing goods-market clearing gives

$$\frac{1}{p_{Y,t}} = \psi\rho + (1-\psi)\left[r_t + \pi_{Y,t} + b_t - \mathcal{R}_t\right] + \widetilde{\xi}_t. \tag{D.19}$$

The objective pricing identity for the aggregate endowment claim is

$$\frac{1}{p_{Y,t}} = r_t + \pi_{Y,t} - \mu_{P_Y,t}, \tag{D.20}$$

where $\mu_{P_Y,t}$ is the objective expected capital-gain rate. Combining Equations (D.19) and (D.20) yields

$$r_t = \rho + \psi^{-1}\left(\mu_{P_Y,t} + \widetilde{\xi}_t\right) + (1-\psi^{-1})\mathcal{R}_t - \pi_{Y,t} - (1-\psi^{-1})b_t. \tag{D.21}$$

The final term is new. When investors interpret a flow innovation as favorable cash-flow news, their perceived return on aggregate wealth exceeds its objective expected return, shifting their desired consumption-saving choice.

The objective pricing kernel. The preceding equations can also be obtained directly from the stochastic discount factor. Let L_t denote the likelihood-ratio process associated with investors' beliefs. Given Equation (D.7), it satisfies

$$\frac{dL_t}{L_t} = h_t\, dZ_t.$$

Let $\widetilde{\Lambda}_t$ denote the SDF under investors' subjective measure:

$$\frac{d\widetilde{\Lambda}_t}{\widetilde{\Lambda}_t} = -r_t\, dt - \widetilde{\eta}_t\, d\widetilde{Z}_t.$$

The objective state-price density is $\Lambda_t = L_t\widetilde{\Lambda}_t$. Itô's product rule gives

$$\frac{d\Lambda_t}{\Lambda_t} = -r_t\, dt - (\widetilde{\eta}_t - h_t)\, dZ_t = -r_t\, dt - \eta_t\, dZ_t.$$

This verifies that the objective valuation equations use $\eta_t = \widetilde{\eta}_t - h_t$, not the perceived market price of risk $\widetilde{\eta}_t$.

Valuation equations. For $k \in \{Y, D\}$, let $\phi_Y = 1$ and $\phi_D = \phi$. Write $p_{k,t} = p_k(\widetilde{X}_t)$ and compute its objective drift and diffusion from Equations (D.14) and (D.15). Return volatility is

$$\sigma_{R_k,t} = \phi_k\sigma + \sigma_{p_k,t}.$$

The objective valuation equation is

$$\frac{1}{p_{k,t}} = r_t + \eta_t\left(\phi_k\sigma + \sigma_{p_k,t}\right)^\top - \left(\mu + \mu_{p_k,t} + \phi_k\sigma\sigma_{p_k,t}^\top\right).$$

The same equation results under the subjective measure. Under that representation, expected cash-flow growth rises by $\phi_k\sigma h_t^\top$, the subjective drift of the valuation ratio rises by $\sigma_{p_k,t}h_t^\top$, and perceived expected returns rise by $\sigma_{R_k,t}h_t^\top$. Because $\sigma_{R_k,t} = \phi_k\sigma + \sigma_{p_k,t}$, these three belief terms cancel exactly in the pricing identity. Beliefs nevertheless affect prices through equilibrium consumption, portfolios, the interest rate, and the objective market price of risk.

State dynamics and closure. Under the objective measure, the passive portfolio share and belief state satisfy

$$d\overline{\alpha}_{p,t} = \kappa_p \left(\overline{\alpha} - \overline{\alpha}_{p,t}\right) dt + \sigma_{\alpha_p} dZ_t,$$
$$dm_t = -\kappa_m m_t \, dt + \theta \sigma_{\alpha_p,2} \, dZ_{2,t}.$$

The objective wealth-share dynamics are given by Equation (D.12). Together, these equations determine $\mu_{\widetilde{X},t}$ and $\sigma_{\widetilde{X},t}$.

Empirical discipline. The belief loading can be disciplined using the response of dividend-growth expectations to non-fundamental price movements estimated by Chaudhry (2025). Along a pure portfolio-flow innovation, our specification implies

$$\mathcal{I}_D^F \equiv \frac{\partial \log(P_{D,t}/D_t)}{\partial dF_t},$$

where $\mathcal{I}_D^F$ is the response of the log price-dividend ratio per unit of the passive-share innovation dF_t. This differs from the headline price-impact coefficient, whose flow is additionally scaled by the passive sector's lagged wealth share. The model counterpart of Chaudhry's estimate is therefore

$$\alpha_{\text{model}} \equiv \frac{\partial(\mu + \phi m_t)/\partial dF_t}{\partial \log(P_{D,t}/D_t)/\partial dF_t} = \frac{\phi\theta}{\mathcal{I}_D^F},$$

Chaudhry's preferred average-treatment-effect estimate implies a target of $\alpha_{\text{model}} = 0.12$, while the flow-induced-trading local-treatment estimate of 0.206 provides a higher-response alternative. We calibrate θ in the economy with homogeneous risk aversion and slack leverage constraints, which isolates the belief-distortion mechanism. The value $\theta = 0.08$ implies $\alpha_{\text{model}} \approx 0.121$, closely matching Chaudhry's preferred estimate of 0.12. We then hold θ fixed when introducing preference heterogeneity and the leverage constraint. This calibration matches the response to a flow-induced price movement without making the belief state respond directly to the equity price. Doing so would introduce an additional contemporaneous fixed point between beliefs and equity valuation.

The computational problem therefore consists of three HJB equations and one aggregate-valuation equation for

$$\left\{c_p(\widetilde{X}), c_h(\widetilde{X}), c_\ell(\widetilde{X}), p_Y(\widetilde{X})\right\},$$

followed by the valuation equation for $p_D(\widetilde{X})$. The number of unknown functions is unchanged relative to the baseline model, but each function depends on the additional state m. At every state, Equations (D.10), (D.18), (D.21), and (D.11) determine portfolios, the objective market price of risk, the interest rate, and the aggregate valuation ratio. Equations (D.13)–(D.16) supply the objective derivatives used in simulation and valuation and the subjective derivatives used in the HJB equations.

When $\theta = 0$ and $m = 0$, we have $h = 0$, $\widetilde{\eta} = \eta$, $\widetilde{\mu}_{\widetilde{X}} = \mu_{\widetilde{X}}$, and $b = 0$. Every equation then reduces to its baseline counterpart. This provides a direct code-level check of the behavioral implementation.

E Closed-Form Perturbation Results

This appendix specializes the general perturbation system in Appendix A.3 to obtain the closed-form results used in Section 4. We first derive the solution without preference heterogeneity or leverage constraints and then show that the endogenous response of the wealth distribution affects price impact only at a higher order.

E.1 Passive demand without preference heterogeneity or leverage constraints

We now specialize the general second-order system to the case used in Propositions 4 and 5. Suppose $\hat{\gamma}_j = 0$ for every $j \in \mathcal{J}$, the leverage constraint is slack, and the mean passive portfolio share equals its benchmark value. The first-order exposure coefficients are

$$m_{p,1} = \hat{\alpha}_p, \qquad m_{j,1} = -\frac{x_p}{x_a}\hat{\alpha}_p, \quad j \in \{h, \ell\},$$

where $x_a = 1 - x_p$. Therefore,

$$\sum_{j\in\mathcal{J}} x_j m_{j,1}^2 = \frac{x_p}{x_a}\hat{\alpha}_p^2.$$

The fixed-share second-order correction to the payout yield is therefore

$$y_{Y,2}^0 = (1-\psi^{-1})\frac{\gamma\|\sigma\|^2}{2}\sum_{j\in\mathcal{J}} x_j m_{j,1}^2 = d_Y^0\hat{\alpha}_p^2,$$

where

$$d_Y^0 \equiv \frac{1}{2}(1-\psi^{-1})\gamma\|\sigma\|^2\frac{x_p}{x_a}.$$

Define

$$M_p \equiv (1-\gamma)\hat{\sigma}_p\sigma^\top, \qquad V_p \equiv \hat{\sigma}_p\hat{\sigma}_p^\top.$$

The general coefficient equations in Appendix A.3.4 imply

$$y_{Y,2}^T = a_Y^T + b_Y^T\hat{\alpha}_p + d_Y^T\hat{\alpha}_p^2,$$

where

$$\begin{aligned} d_Y^T &= \frac{d_p}{d_p + 2\kappa_p}d_Y^0, \\ b_Y^T &= \frac{2d_Y^T M_p}{d_p + \kappa_p}, \\ a_Y^T &= \frac{b_Y^T M_p + d_Y^T V_p}{d_p}. \end{aligned}$$

Only the slope of $y^T_{Y,2}$ with respect to $\hat{\alpha}_p$ matters for price impact:

$$\frac{\partial y^T_{Y,2}}{\partial \hat{\alpha}_p} = 2d^T_Y \left(\hat{\alpha}_p + \frac{M_p}{d_p + \kappa_p} \right).$$

The variance term V_p affects only the constant coefficient a^T_Y and therefore does not enter the derivative with respect to $\hat{\alpha}_p$.

A normalized passive flow is $f_t = x_p \hat{\alpha}_{p,t} \epsilon$. Since $y_Y \equiv p_Y^{-1}$ and $y_{Y,0} = d_p$, its price impact is

$$\varepsilon^{-1}_{Y,t} = -\frac{1}{d_p} \frac{\partial y^T_{Y,2}}{\partial \hat{\alpha}_p} \frac{\epsilon}{x_p} + O(\epsilon^2).$$

Using $\overline{\alpha}_{p,t} = 1 + \hat{\alpha}_{p,t}\epsilon$ and $\sigma_{\alpha_p} = \hat{\sigma}_p \epsilon$ gives

$$\varepsilon^{-1}_{Y,t} = \frac{1 - \psi^{-1}}{d_p + 2\kappa_p} \frac{\gamma \|\sigma\|^2}{x_a} \left[1 - \overline{\alpha}_{p,t} + \frac{\gamma - 1}{d_p + \kappa_p} \sigma_{\alpha_p} \sigma^\top \right] + O(\epsilon^2),$$

which proves Proposition 4.

The diffusion of the valuation ratio then implies

$$\sigma_{R_Y,t} = \sigma + x_{p,t} \varepsilon^{-1}_{Y,t} \sigma_{\alpha_p} + O(\epsilon^3).$$

Substituting this expression into the active investors' optimal exposure and imposing aggregate exposure clearing yields

$$\eta_t = \gamma \chi_{a,t} \sigma + (1 - H_{a,t}) \gamma x_{p,t} \varepsilon^{-1}_{Y,t} \sigma_{\alpha_p} + O(\epsilon^3),$$

where

$$\chi_{a,t} \equiv \frac{1 - x_{p,t} \overline{\alpha}_{p,t}}{x_{a,t}}, \qquad H_{a,t} \equiv \frac{1 - \gamma^{-1}}{\psi - 1} \left(\frac{\psi}{x_{a,t}} - 1 \right).$$

Consequently,

$$\pi_{Y,t} = \eta_t \sigma^\top + \gamma x_{p,t} \varepsilon^{-1}_{Y,t} \sigma_{\alpha_p} \sigma^\top + O(\epsilon^3).$$

Finally, aggregating the consumption-wealth conditions and using goods-market clearing gives

$$r_t = r^* + (1 - \psi^{-1}) \frac{\gamma \|\sigma\|^2}{2} \frac{x_{p,t}}{x_{a,t}} (\overline{\alpha}_{p,t} - 1)^2 - (\eta_t - \eta^*) \sigma^\top + O(\epsilon^3).$$

These expressions establish Proposition 5.

E.2 The role of the wealth distribution

The aggregate market elasticity computed above is a partial derivative: it changes the passive portfolio share while holding the wealth distribution fixed. A portfolio-flow shock, however, also changes wealth shares because passive and active investors have different exposures to the risky claim. We now show that, in the special case without preference heterogeneity or leverage

constraints, this wealth-distribution channel is one order smaller than the direct portfolio-flow channel.

Let x_p denote the wealth share of passive investors and $x_a = 1 - x_p$ the wealth share of active investors. Because passive investors hold the portfolio share $\overline{\alpha}_p$, their wealth-share diffusion is

$$\sigma_{x_p,t} = x_{p,t}\left(\sigma_{p,t} - \sigma_{R_Y,t}\right) = x_{p,t}\left(\overline{\alpha}_{p,t} - 1\right)\sigma_{R_Y,t}.$$

Using $\overline{\alpha}_{p,t} = 1 + \hat{\alpha}_{p,t}\epsilon$ and $\sigma_{R_Y,t} = \sigma + O(\epsilon^2)$, this gives

$$\sigma_{x_p,t} = x_{p,t}\hat{\alpha}_{p,t}\sigma\epsilon + O(\epsilon^3).$$

Thus, the wealth-distribution state moves in response to the same aggregate return shock that redistributes wealth between passive and active investors.

The payout yield of the aggregate claim satisfies

$$y_Y(x_p, \hat{\alpha}_p; \epsilon) = y_{Y,0} + y_{Y,2}^T(x_p, \hat{\alpha}_p)\epsilon^2 + O(\epsilon^3).$$

Therefore, allowing both x_p and $\hat{\alpha}_p$ to move, its diffusion is

$$\sigma_{y_Y,t} = \frac{G_p}{y_{Y,0}}\hat{\sigma}_p\epsilon^2 O(\epsilon^3), \qquad G_p \equiv y_{Y,2,\hat{\alpha}_p}^T.$$

The leading term is the direct portfolio-flow channel studied above. Because $\sigma_{x_p,t} = O(\epsilon)$, endogenous wealth redistribution enters only at the next order. Consequently, the diffusion of the aggregate claim's valuation ratio is

$$\sigma_{p_Y,t} = -\frac{G_p}{y_{Y,0}}\hat{\sigma}_p\epsilon^2 + O(\epsilon^3).$$

Using the aggregate market elasticity relation

$$\varepsilon_{Y,t}^{-1} = -\frac{G_p}{y_{Y,0}}\frac{\epsilon}{x_p} + O(\epsilon^2),$$

we can write

$$\sigma_{p_Y,t} = x_p\varepsilon_{Y,t}^{-1}\sigma_{\alpha_p} + O(\epsilon^3).$$

Thus, to the order at which aggregate market elasticity is computed, the equilibrium price response equals the direct effect obtained by holding the wealth distribution fixed. The endogenous movement in the wealth distribution contributes only at the next order.

F Proofs of Propositions 6 and 7

F.1 Proof of Proposition 6

Proof. We first derive price impact for a fixed passive portfolio share, holding the wealth distribution fixed. Write the first-order risk exposure of investor j as

$$\sigma_{j,1} = m_{j,1}\sigma.$$

The passive investor's exposure is

$$m_{p,1} = \hat{\alpha}_p. \tag{F.1}$$

The second-order correction to the aggregate payout yield that depends on these exposures is

$$y^0_{Y,2} = (1-\psi^{-1})\gamma\|\sigma\|^2 \sum_{j\in\mathcal{J}} x_j \left(\frac{1}{2}m^2_{j,1} + \hat{\gamma}_j m_{j,1}\right). \tag{F.2}$$

Slack leverage constraint. When the leverage constraint is slack, all active investors are unconstrained, so $x_c = 0$ and $x_u = x_a = x_h + x_\ell$. Define the wealth-weighted average perturbation in risk aversion among unconstrained active investors by

$$\mathbb{E}^u[\hat{\gamma}_j] \equiv \sum_{j\in\{h,\ell\}} \frac{x_j}{x_u}\hat{\gamma}_j.$$

The first-order solution gives

$$m_{j,1} = \mathbb{E}^u[\hat{\gamma}_j] - \frac{x_p}{x_u}\hat{\alpha}_p - \hat{\gamma}_j, \qquad j \in \{h,\ell\}. \tag{F.3}$$

Differentiating Equation (F.2) with respect to $\hat{\alpha}_p$ and using Equations (F.1) and (F.3) yields

$$\frac{\partial y^0_{Y,2}}{\partial \hat{\alpha}_p} = (1-\psi^{-1})\gamma\|\sigma\|^2 x_p \left[\frac{\hat{\alpha}_p}{x_u} + \hat{\gamma}_p - \mathbb{E}^u[\hat{\gamma}_j]\right]. \tag{F.4}$$

Since $\gamma_j = \gamma(1+\hat{\gamma}_j\epsilon)$, define

$$\overline{\alpha}^{\text{opt}}_p \equiv 1 - x_u\frac{\gamma_p - \mathbb{E}^u[\gamma_j]}{\gamma}.$$

Equation (F.4) can then be written as

$$\frac{\partial y^0_{Y,2}}{\partial \hat{\alpha}_p} = (1-\psi^{-1})\gamma\|\sigma\|^2\frac{x_p}{x_u}\frac{\overline{\alpha}_p - \overline{\alpha}^{\text{opt}}_p}{\epsilon}. \tag{F.5}$$

Binding leverage constraint. When the constraint of type ℓ binds, $x_c = x_\ell$, $x_u = x_h$, and $\mathbb{E}^u[\gamma_j] = \gamma_h$. The constrained investor's first-order exposure is $m_{\ell,1} = \hat{\sigma}$, while market clearing

implies

$$m_{h,1} = -\frac{x_p\hat{\alpha}_p + x_c\hat{\sigma}}{x_u}.$$

Differentiating Equation (F.2) now gives

$$\frac{\partial y^0_{Y,2}}{\partial \hat{\alpha}_p} = (1-\psi^{-1})\gamma\|\sigma\|^2 x_p\left[\frac{1-x_c}{x_u}\hat{\alpha}_p + \frac{x_c}{x_u}\hat{\sigma} + \hat{\gamma}_p - \hat{\gamma}_h\right]. \tag{F.6}$$

In this regime, define

$$\overline{\alpha}_p^{\text{opt}} \equiv 1 - \frac{x_u}{1-x_c}\frac{\gamma_p - \mathbb{E}^u[\gamma_j]}{\gamma} - \frac{x_c}{1-x_c}\left(\frac{\overline{\sigma}}{\|\sigma\|} - 1\right).$$

Using $\overline{\sigma} = \|\sigma\|(1+\hat{\sigma}\epsilon)$, Equation (F.6) becomes

$$\frac{\partial y^0_{Y,2}}{\partial \hat{\alpha}_p} = (1-\psi^{-1})\gamma\|\sigma\|^2 x_p \frac{1-x_c}{x_u}\frac{\overline{\alpha}_p - \overline{\alpha}_p^{\text{opt}}}{\epsilon}. \tag{F.7}$$

Equation (F.7) also nests Equation (F.5) when $x_c = 0$ and $x_u = x_a$.

Because $f_t = x_p\hat{\alpha}_p\epsilon$ and $y_Y = p_Y^{-1}$ with $y_{Y,0} = d_p$, fixed-share price impact is

$$\begin{aligned}(\varepsilon^0_{Y,t})^{-1} &= -\frac{1}{d_p}\frac{\partial y^0_{Y,2}}{\partial\hat{\alpha}_p}\frac{\epsilon}{x_p} + O(\epsilon^2)\\ &= (1-\psi^{-1})\frac{\gamma\|\sigma\|^2}{d_p}\frac{1-x_c}{x_u}\left(\overline{\alpha}_p^{\text{opt}} - \overline{\alpha}_p\right) + O(\epsilon^2).\end{aligned} \tag{F.8}$$

We next introduce the mean-reverting passive portfolio share. Holding the wealth distribution fixed, write the fixed-share payout-yield correction as

$$y^0_{Y,2} = a^0_Y + b^0_Y\hat{\alpha}_p + d^0_Y\hat{\alpha}_p^2.$$

Equation (F.8) implies

$$\begin{aligned}\frac{2d^0_Y}{x_p d_p} &= (1-\psi^{-1})\frac{\gamma\|\sigma\|^2}{d_p}\frac{1-x_c}{x_u},\\ -\frac{b^0_Y}{x_p d_p}\epsilon &= \frac{2d^0_Y}{x_p d_p}\left(\overline{\alpha}_p^{\text{opt}} - 1\right).\end{aligned} \tag{F.9}$$

The general second-order coefficient equations in Appendix A.3.4 apply pointwise in the wealth

distribution and give

$$d_Y^T = \frac{d_p}{d_p + 2\kappa_p} d_Y^0,$$
$$b_Y^T = \frac{d_p}{d_p + \kappa_p} b_Y^0 + \frac{2d_Y^T}{d_p + \kappa_p} M_p, \tag{F.10}$$

where

$$M_p \equiv \kappa_p \hat{\alpha} + (1-\gamma)\hat{\sigma}_p \sigma^\top.$$

The constant coefficient does not enter price impact. Therefore,

$$\varepsilon_{Y,t}^{-1} = -\frac{1}{d_p}\left(2d_Y^T \hat{\alpha}_{p,t} + b_Y^T\right)\frac{\epsilon}{x_{p,t}} + O(\epsilon^2).$$

Substituting Equations (F.9) and (F.10), and using

$$\overline{\alpha}_{p,t} = 1 + \hat{\alpha}_{p,t}\epsilon, \qquad \overline{\alpha} = 1 + \hat{\alpha}\epsilon, \qquad \sigma_{\alpha_p} = \hat{\sigma}_p \epsilon,$$

gives

$$\varepsilon_{Y,t}^{-1} = (1-\psi^{-1})\frac{\gamma\|\sigma\|^2}{d_p + 2\kappa_p}\frac{1-x_{c,t}}{x_{u,t}}\left[\overline{\alpha}_{p,t}^{\text{opt}} - \overline{\alpha}_{p,t} + \frac{\kappa_p(\overline{\alpha}_{p,t}^{\text{opt}} - \overline{\alpha})}{d_p + \kappa_p} + \frac{(\gamma-1)\sigma_{\alpha_p}\sigma^\top}{d_p + \kappa_p}\right] + O(\epsilon^2),$$

which is Equation (25). □

F.2 Proof of Proposition 7

Proof. The equity valuation ratio satisfies

$$\frac{1}{p_{D,t}} = r_t + \eta_t(\phi\sigma + \sigma_{p_D,t})^\top - \left(\mu + \mu_{p_D,t} + \phi\sigma\sigma_{p_D,t}^\top\right). \tag{F.11}$$

At the benchmark, the valuation ratio is constant and therefore

$$d_D \equiv \frac{1}{p_D^*} = r^* + \phi\eta^*\sigma^\top - \mu.$$

Let $f_t = x_{p,t}(\overline{\alpha}_{p,t} - 1)$ denote the normalized passive flow. Proposition 3 implies

$$\frac{\partial r_t}{\partial f_t} = \zeta_{r,t}^{-1}, \qquad \frac{\partial(\eta_t\sigma^\top)}{\partial f_t} = -\zeta_{r,t}^{-1}.$$

It follows that the required return on equity responds according to

$$\frac{\partial}{\partial f_t}\left(r_t + \phi\eta_t\sigma^\top\right) = -(\phi - 1)\zeta_{r,t}^{-1}. \tag{F.12}$$

To compute the valuation response, write the component of the first-order equity valuation ratio that depends on passive demand as

$$p_{D,t} = p_D^* + b_D(x_t)\hat{\alpha}_{p,t}\epsilon + \text{terms independent of } \hat{\alpha}_{p,t} + O(\epsilon^2). \tag{F.13}$$

The passive portfolio share satisfies

$$d\hat{\alpha}_{p,t} = \kappa_p(\hat{\alpha} - \hat{\alpha}_{p,t})dt + \hat{\sigma}_p dZ_t.$$

Consequently, the component of $\mu_{p_D,t}$ proportional to $\hat{\alpha}_{p,t}\epsilon$ is

$$-\kappa_p d_D b_D(x_t)\hat{\alpha}_{p,t}\epsilon. \tag{F.14}$$

The diffusion $\hat{\sigma}_p$ contributes to terms independent of the current value of $\hat{\alpha}_{p,t}$ and therefore does not affect the slope $b_D(x_t)$.

Expanding the left side of Equation (F.11) gives

$$\frac{1}{p_{D,t}} = d_D - d_D^2 b_D(x_t)\hat{\alpha}_{p,t}\epsilon + \text{terms independent of } \hat{\alpha}_{p,t} + O(\epsilon^2).$$

Using $f_t = x_{p,t}\hat{\alpha}_{p,t}\epsilon$, Equations (F.12) and (F.14) imply that matching the coefficient on $\hat{\alpha}_{p,t}\epsilon$ yields

$$-d_D^2 b_D(x_t) = -(\phi - 1)x_{p,t}\zeta_{r,t}^{-1} + \kappa_p d_D b_D(x_t).$$

Therefore,

$$b_D(x_t) = \frac{(\phi - 1)x_{p,t}\zeta_{r,t}^{-1}}{d_D(d_D + \kappa_p)}. \tag{F.15}$$

Finally, Equations (F.13) and (F.15) give

$$\begin{aligned}
\varepsilon_{D,t}^{-1} &= \frac{1}{x_{p,t}} \frac{\partial \log p_{D,t}}{\partial \overline{\alpha}_{p,t}} \\
&= \frac{d_D b_D(x_t)}{x_{p,t}} + O(\epsilon) \\
&= \frac{\phi - 1}{d_D + \kappa_p}\zeta_{r,t}^{-1} + O(\epsilon),
\end{aligned}$$

which proves the result. □

G Details of Section 6

G.1 Flow of Funds aggregation

This appendix describes the procedure used in Section 6.1 to aggregate Flow of Funds (FoF) sectors into the three ultimate sectors of the model: passive, high-risk-aversion active, and low-risk-aversion active. We follow the aggregation in Koijen et al. (2024), drop sectors with no or minimal equity holdings, and route pass-through vehicles to their ultimate end investors. ETFs

(L.124) and closed-end funds (L.123) are aggregated onto the household balance sheet (L.101); mutual-fund holdings (L.122) are split across end investors using the FoF's breakdown, with the household portion kept separate. Defined-contribution pensions (L.118c, L.119c, L.120c) are routed to households. Hedge funds are separated out using Table B.101.f for domestic funds and the Enhanced Financial Accounts for foreign funds, and the corresponding holdings are subtracted from the household (L.101) and rest-of-the-world (L.133) balance sheets. The remaining insurance and pension sectors (L.115, L.116, L.118b, L.119b, L.120b) are aggregated into a single long-horizon sector. The three ultimate sectors correspond, respectively, to *passive* (households and non-profits net of hedge funds, government, monetary authority, and money-market funds), *high-risk-aversion active* (insurance, pensions, and rest-of-the-world net of hedge funds), and *low-risk-aversion active* (hedge funds, broker-dealers, and other short-horizon intermediaries). The leverage constraint applies to the low-risk-aversion active sector but binds only occasionally.

G.2 Constructing quarterly equity flows

This appendix describes how we construct the quarterly aggregate equity flow series plotted in Figure A.1 and used to discipline the passive-demand process in Section 6.1. The construction follows Appendix D of Gabaix and Koijen (2023) and replicates their Figure D.7.

The raw inputs are sectoral dollar values in equity from the U.S. Flow of Funds (FoF) at quarterly frequency.

For each FoF sector j and quarter t, let $V_{j,t}$ be the dollar value of equity holdings at the end of the quarter. We first infer sector-specific betas β_j from the time-series regressions

$$\log(V_{j,t}/V_{j,t-1}) = a_j + \beta_j \mathrm{Mkt}_t + \epsilon_{j,t}.$$

where Mkt_t is the CRSP market return factor. We then impute the "active" dollar holdings change as

$$\Delta F^E_{j,t} = V_{j,t} - V_{j,t-1}(1 + \beta_j \mathrm{Mkt}_t).$$

Note that this adjustment removes mechanical revaluation effects. Let $\Delta F^{\mathrm{Firm}}_t$ denote the corresponding dollar net issuance of equity by U.S. firms (positive for net issuance, negative for net repurchases), and let E_{t-1} denote the aggregate U.S. equity market capitalization at the end of quarter $t-1$. We aggregate FoF sectors into the same set of investor groups used in our Flow of Funds aggregation (Appendix G.1), and treat the corporate sector as an additional "investor" on the issuance side.

We scale each dollar flow by lagged aggregate market capitalization to obtain a unitless rebalancing measure, $\Delta F^E_{j,t}/E_{t-1}$, that is comparable across sectors and over time. The aggregate gross flow is one half of the sum of absolute scaled flows,

$$y^{\mathrm{Gross}}_t = \frac{\sum_j \left|\Delta F^E_{j,t}\right| + \left|\Delta F^{\mathrm{Firm}}_t\right|}{2\,E_{t-1}},$$

where the division by two avoids double counting: each dollar of equity bought by one sector is matched by a dollar sold by another, so the total absolute trade exceeds the equity actually reallocated by a factor of two. We refer to y^{Gross}_t as the *gross flow*. This is the blue line in Figure A.1 and provides the empirical target for quarterly gross-flow volatility in the model.

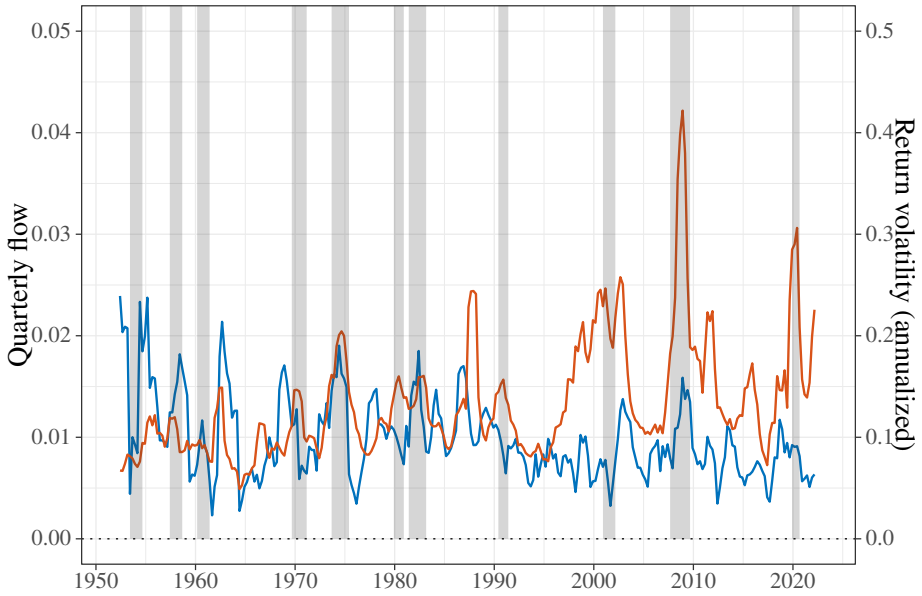


Figure A.1. Quarterly flows and return volatility.
The figure plots quarterly aggregate equity flows in blue and realized return volatility in red. Quarterly flows are scaled by lagged aggregate market capitalization. Both series are shown as four-quarter centered moving averages over 1952.Q2–2022.Q1; the calibration uses the 1993.Q1–2018.Q4 subsample described below. Source: Flow of Funds and CRSP.

Net firm issuance also reallocates equity risk: when firms issue (repurchase) shares, equity risk is created (redeemed) outside the existing investor base. We measure this component separately as the *absolute net flow*,

$$y_t^{\text{AbsNet}} = \frac{\left|\Delta F_t^{\text{Firm}}\right|}{E_{t-1}}.$$

The difference $y_t^{\text{Gross}} - y_t^{\text{AbsNet}}$ measures the equity risk that is reallocated across investor sectors at date t, as opposed to created or destroyed by corporate actions.

We use the FoF sample from 1993.Q1 to 2018.Q4, the period over which the underlying sectoral series are most reliable. Following Gabaix and Koijen (2023), we merge the mutual-fund and ETF sectors to mitigate the volatility of ETF flows in the early years of the sample, and winsorize the cross section of scaled flows at the 5th and 95th percentiles over 1993.Q1–2006.Q4 to limit the influence of outliers in the early part of the sample. For consistency with our model, where the passive sector is exogenous and other sectors clear the market, we treat the corporate issuance term symmetrically with sectoral flows when computing y_t^{Gross}.

To target the volatility of passive demand, we use y_t^{Gross} as a model-consistent measure of quarterly equity-risk reallocation. Concretely, we calibrate $\|\sigma_{\alpha_p}\|$ so that the standard deviation of simulated quarterly gross flows in the model matches the standard deviation of y_t^{Gross} in the data. Normalizing flows by market capitalization makes them comparable across time and between the data and model. For the correlation parameter ν_p, we use the contemporaneous correlation between annual changes in the passive portfolio share implied by our FoF aggregation and annual real CRSP dividend growth. We measure the annual portfolio share at year-end and use the 1993–2018 sample. The estimated correlation is 0.08.

We discipline the persistence of passive demand using the household-level evidence of Brunnermeier and Nagel (2008). They construct the change in a household's risky-asset share that would arise under complete portfolio inertia and estimate that approximately 0.743 of this passive change remains over a five-year interval and 0.754 remains over a two-year interval. In our continuous-time specification, a deviation of the passive portfolio share from its long-run mean survives for Δ years at rate $\exp(-\kappa_p \Delta)$. The two estimates therefore imply

$$-\frac{\log(0.743)}{5} = 0.059 \qquad \text{and} \qquad -\frac{\log(0.754)}{2} = 0.141.$$

We accordingly view $\kappa_p \in [0.06, 0.14]$ as the range consistent with their evidence on household portfolio inertia.

G.3 Numerical methodology

We solve the rational and behavioral economies using deep neural networks to approximate the equilibrium functions, as in Duarte et al. (2024) and Gopalakrishna and Wu (2026). The solution has two stages. The first stage solves jointly for consumption, portfolios, the interest rate, and the valuation of the aggregate endowment claim. Conditional on this equilibrium, the second stage prices the dividend claim. The same procedure applies to all four economies considered in Section 6; preference heterogeneity and the leverage constraint change the equilibrium conditions imposed at each state, while belief distortions add one state variable.

G.3.1 Aggregate equilibrium

In the rational economy, the state vector is

$$X = (x_\ell, x_h, \overline{\alpha}_p), \qquad x_p = 1 - x_\ell - x_h.$$

We approximate one positive consumption–wealth ratio for each investor type:

$$c_j(X) = c_j^0 \exp\big(\mathrm{NN}_j(X; \Theta_j)\big), \qquad j \in \{p, h, \ell\}, \tag{G.1}$$

where c_j^0 is a constant used to center the network at initialization. Goods-market clearing then determines the aggregate price-dividend ratio directly:

$$p_Y(X) = \left[\sum_{j \in \mathcal{J}} x_j c_j(X)\right]^{-1}. \tag{G.2}$$

Thus, p_Y is not an additional function to be approximated.

Given the networks and their first and second derivatives, we recover the remaining equilibrium objects state by state. The investors' portfolio first-order conditions, risk-market clearing, and Itô consistency of Equation (G.2) form a linear system for the active investors' risk exposures, the return volatility of the aggregate claim, and the market price of risk. When the low-risk-aversion investor's constraint binds, we solve simultaneously for its multiplier by imposing complementary

slackness. The aggregate pricing equation then determines the risk-free rate, and the portfolio and consumption choices imply the drift and diffusion of the wealth distribution. Because these equilibrium conditions are solved algebraically at every state, the only residual conditions used for training are the three consumption HJB equations. We minimize their mean squared normalized residuals over states drawn from the interior of the wealth-share simplex and from the calibrated support of passive demand.

Each consumption–wealth network is a fully connected feed-forward network with three hidden layers of 32 neurons and SiLU activation functions. The exponential transformation in Equation (G.1) guarantees positivity, while the constants c_j^0 center the networks near the corresponding state-independent solution at initialization. We draw each training batch using Latin-hypercube sampling over the wealth-share simplex, imposing a small interior floor on each wealth share, and sample passive demand over its calibrated support. The behavioral model additionally samples the belief state over its training support. First and second derivatives of the networks with respect to the state variables are computed by automatic differentiation.

Let $\mathcal{R}_j^{\mathrm{HJB}}(S;\Theta)$ denote investor j's HJB residual at state S, after the portfolio first-order conditions, market-clearing conditions, and Itô consistency conditions have been solved at that state. The first-stage objective is

$$\widehat{\mathcal{L}}_1(\Theta) = \frac{1}{N}\sum_{n=1}^{N}\sum_{j\in\{p,h,\ell\}}\left(\frac{\mathcal{R}_j^{\mathrm{HJB}}(S^n;\Theta)}{\rho}\right)^2. \tag{G.3}$$

Normalizing by the rate of time preference puts the three residuals on a common relative scale. We minimize Equation (G.3) with ADAM and clipped gradients, using fresh state draws at each update and a smoothly declining learning rate. The portfolio and market-clearing equations do not enter the loss because they are imposed directly, up to numerical precision, at every training state.

The first-stage calculation can be summarized as follows:

Algorithm 1 Aggregate-equilibrium training.

Initialize the three consumption–wealth networks with parameters Θ.
for $k = 1, \ldots, K_1$ **do**
  Draw a batch of states $\{S^n\}_{n=1}^N$.
  Evaluate $c_j(S^n)$ and their first and second derivatives.
  Use goods-market clearing to recover $p_Y(S^n)$.
  Solve the portfolio first-order conditions, risk-market clearing, and Itô consistency for portfolios, return volatility, and the market price of risk.
  Recover the interest rate and the drift and diffusion of the state vector.
  Evaluate the three HJB residuals and the loss in Equation (G.3).
  Update Θ using one ADAM step.
end for

The behavioral economy uses the expanded state vector

$$\widetilde{X} = (x_\ell, x_h, \overline{\alpha}_p, m).$$

The approximated objects remain the three consumption–wealth ratios in Equation (G.1), now evaluated at $\widetilde{X}$. Portfolio and consumption optimality use the subjective generator and perceived market price of risk derived in Appendix D.2.1; market clearing, realized wealth dynamics, and valuation use the objective measure. Goods-market clearing continues to determine p_Y directly. Hence, the belief distortion increases the dimension of the state space from three to four but does not increase the number of unknown equilibrium functions. Setting $m = 0$ and $\theta = 0$ reduces every equation to its rational counterpart, providing a direct implementation check.

G.3.2 Valuation of the dividend claim

Conditional on the first-stage solution, we approximate the dividend claim's price-dividend ratio $p_D = P_D/D$ with a separate positive neural network. For either state vector $S \in \{X, \widetilde{X}\}$, write

$$\frac{dp_D(S_t)}{p_D(S_t)} = \mu_{p_D,t}\, dt + \sigma_{p_D,t}\, dZ_t.$$

The objective pricing condition implies

$$0 = 1 + p_D \left[\mu + \mu_{p_D} + \phi\sigma\sigma_{p_D}^\top - r - \eta\left(\phi\sigma + \sigma_{p_D}\right)^\top\right]. \tag{G.4}$$

All coefficients in this linear valuation equation are supplied by the first-stage equilibrium. The corporate-bond claim is then the residual claim on aggregate output. In the behavioral economy, Equation (G.4) is evaluated under the objective measure; the subjective cash-flow and valuation drifts generate the same pricing equation because the change-of-measure terms cancel, as shown in Appendix D.2.1.

The valuation network has the same three-layer architecture as the consumption–wealth networks and uses an exponential output transformation to impose $p_D > 0$. Holding the first-stage networks fixed, we evaluate the coefficients in Equation (G.4) on a global set of collocation states. Automatic differentiation supplies the first and second derivatives needed to compute μ_{p_D} and σ_{p_D} through Itô's lemma. If $\mathcal{R}_D^{\text{val}}(S;\Theta_D)$ denotes the resulting pricing residual, we train the valuation network by minimizing the mean squared relative residual

$$\widehat{\mathcal{L}}_2(\Theta_D) = \frac{1}{N}\sum_{n=1}^{N}\left(\frac{\mathcal{R}_D^{\text{val}}(S^n;\Theta_D)}{p_D(S^n;\Theta_D)}\right)^2. \tag{G.5}$$

Thus, the valuation stage updates only Θ_D; none of the aggregate-equilibrium objects are retrained. As a solution diagnostic, we evaluate both stages on states not used in the corresponding parameter update. We also compare the dividend claim's network value with a forward Monte Carlo present-value calculation using the model's stochastic discount factor and verify that the discounted terminal-value contribution becomes small at long horizons.

The valuation stage is:

Algorithm 2 Dividend-claim valuation.

Load and freeze the trained aggregate-equilibrium networks.
Initialize the dividend price–dividend-ratio network with parameters Θ_D.
for $k = 1, \ldots, K_2$ **do**
    Draw a batch from the valuation collocation states.
    Evaluate p_D and its first and second state derivatives.
    Recover μ_{p_D} and σ_{p_D} using Itô's lemma.
    Evaluate Equation (G.4) and the loss in Equation (G.5).
    Update Θ_D using one ADAM step.
end for

G.3.3 Simulation and reported statistics

After solving each economy, we simulate the state variables and recover prices, returns, portfolios, and consumption at each simulated state. The unconditional moments in Tables 2–4 are computed from these simulated panels after discarding the burn-in observations. Price impact is estimated at quarterly frequency using Equation (26), instrumenting aggregate flows with the component of the passive-demand innovation orthogonal to the fundamental shock. The impulse responses, Campbell–Shiller decomposition, state-dependent price-impact estimates, and return-predictability regressions use the same saved solutions and simulation panels.

The replication package separates the computationally intensive solution of the four economies from the simulation and post-processing steps. Conditional on the saved solutions, it reproduces the tables, impulse responses, Campbell–Shiller decomposition, state-dependent price-impact estimates, and return-predictability results reported in Section 6.

G.3.4 Solution accuracy

Table A.1 reports the distribution of the relative equation errors of the trained solutions of the two full-frictions economies, evaluated on 20,000 states drawn in two ways: from the training distribution, which is the same law that generates the training batches, and from each model's ergodic distribution, using a post-burn-in subsample of the simulated panel of Section 6. The ergodic weighting is the economically relevant one because it evaluates the solution where the economy actually lives, while the training weighting also probes rarely visited corners of the state space.

Table A.1. Solution accuracy.
Relative equation errors on 20,000 evaluation states: HJB residuals (normalized by ρ, pooled across investor types) and valuation-PDE residuals (normalized by p_D). "Training" draws states from the training-batch distribution; "ergodic" draws them from the simulated stationary distribution. Portfolio first-order and market-clearing conditions hold by construction; their residual RMS is approximately 1×10^{-7}. Diagnostics are computed on the default-architecture (32×3), seed-0 solution of each economy.

Equation	Weighting	RMS	Median	p95	p99	Max
Rational, full frictions						
HJB residual	training	2.8×10^{-3}	8.3×10^{-4}	5.0×10^{-3}	1.1×10^{-2}	1.2×10^{-1}
HJB residual	ergodic	3.8×10^{-3}	1.8×10^{-3}	8.5×10^{-3}	1.2×10^{-2}	1.8×10^{-2}
Valuation residual	training	2.2×10^{-4}	8.0×10^{-5}	4.1×10^{-4}	8.2×10^{-4}	6.0×10^{-3}
Valuation residual	ergodic	2.5×10^{-4}	1.2×10^{-4}	5.2×10^{-4}	7.9×10^{-4}	1.2×10^{-3}
Behavioral, full frictions						
HJB residual	training	5.1×10^{-3}	1.7×10^{-3}	8.9×10^{-3}	2.0×10^{-2}	1.0×10^{-1}
HJB residual	ergodic	6.2×10^{-3}	3.3×10^{-3}	1.3×10^{-2}	2.3×10^{-2}	3.9×10^{-2}
Valuation residual	training	1.2×10^{-3}	7.4×10^{-4}	2.1×10^{-3}	3.8×10^{-3}	1.7×10^{-2}
Valuation residual	ergodic	1.1×10^{-3}	7.8×10^{-4}	1.9×10^{-3}	2.8×10^{-3}	5.1×10^{-3}

Three independent checks complement the residuals. First, as a closed-form cross-check, we apply the valuation stage of Algorithm 2 to the aggregate endowment claim instead of the dividend claim. It must reproduce the first-stage valuation ratio p_Y, which is known in closed form given the consumption–wealth networks. Re-solving the corresponding PDE with a fresh network gives a mean absolute error of 0.13% in the rational economy (95th percentile 0.35%, maximum 0.69%) and 0.15% in the behavioral economy (0.33%, 0.99%) on an independent sample from the training law. This provides an end-to-end validation of the valuation stage against an exact benchmark. Second, for Monte Carlo pricing and transversality, we start from states drawn from the ergodic distribution and simulate the stochastic discount factor and dividend stream forward for 300 years, using 4,000 antithetic paths per state. The network valuation deviates from the Monte Carlo present value by 1.5% in the rational economy and 3.5% in the behavioral economy on average, against Monte Carlo standard errors of 1.8% and 2.7%. The discounted terminal-value contribution is on the order of 10^{-8} of the price, ruling out the bubble family of solutions to the valuation equation. Third, for ergodic coverage, the simulated panels give $\alpha_p \in [0.24, 0.77]$, strictly inside the training interval $[0.2, 0.8]$; $(x_\ell, x_h) \in [0.24, 0.57] \times [0.11, 0.24]$, well inside the sampled simplex; and a belief state within $[-0.017, 0.016]$, deep inside its training domain $[-0.10, 0.10]$. The reported moments therefore never rely on network extrapolation.

G.3.5 Robustness to seeds and architectures

To separate approximation error from sampling noise, we re-solve both full-frictions economies end to end, including both training stages and the full simulation protocol of Section 6, across three training seeds and five network architectures. These are the default 32×3 and four perturbations of it: narrower 16×3, wider 64×3, shallower 32×2, and deeper 32×4, applied to the consumption–wealth and valuation networks alike. Across the fifteen seed–architecture combinations, the dividend-claim price impact is 2.18 ± 0.04 (mean ± standard deviation) in the rational economy and 6.19 ± 0.06 in the behavioral economy; the endowment-claim price impact is 0.099 ± 0.001 and 0.535 ± 0.003. The equity premium varies by less than 0.19 percentage points, return volatility by less than 0.35 percentage points, and the risk-free rate by about 0.02 percentage points across all thirty solutions.

The only discernible outlier is the shallow two-layer variant, whose ergodic HJB residuals are two to three times larger than the default's; excluding it, the rational price impact is 2.16 ± 0.02. The residual diagnostics of Table A.1 flag precisely the cells whose moments move. For comparison, holding a solution fixed and redrawing only the simulation shocks five times moves the measured price impact by ± 0.05 in the rational economy and ± 0.06 in the behavioral economy. The dispersion across independent numerical solutions is indistinguishable from pure simulation noise, indicating that approximation error does not materially affect the reported moments. Each end-to-end solution, including both training stages, the simulated panel, and the accuracy diagnostics above, runs in about twelve minutes for the rational economy and fourteen minutes for the behavioral economy on a single consumer GPU (NVIDIA RTX 5090).