Dissecting the Aggregate Market Elasticity*

Victor Duarte† Goutham Gopalakrishna‡ Mahyar Kargar§

Jiacui Li¶ Dejanir Silva‖

July 10, 2026

Abstract

We study aggregate stock market elasticity in a general equilibrium model with heterogeneous investors, passive demand, and financial constraints. In a frictionless economy, interest rate movements fully offset risk premium responses and the macro elasticity of the endowment claim is infinite. Frictions break this offset, generating finite price impact. The equity claim has larger price impact since dividends are a levered share of output. A state global perturbation method characterizes the drivers of aggregate elasticity in closed form. We solve the full dynamic model with deep learning. The calibrated economy matches the equity premium, return volatility, and countercyclical price impact.

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

JEL Classification: G12, G11, G21.

*We thank Andy Atkeson, Tim Johnson, Leonid Kogan, Ralph Koijen, Tyler Muir, Stavros Panageas, Yinan Su, Fabrice Tourre, and Pierre-Olivier Weill; and conference and seminar participants at AFA 2026 Annual Meeting, Liquidity in Macroeconomics Workshop at St. Louis Fed, SED 2026 Annual Meeting, 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). Moreover, the ratio of return volatility to flows, which we refer to as the "volatility multiplier," is itself time varying and countercyclical, rising sharply during periods of financial stress. This pattern points to a state-dependent relationship between prices and portfolio flows.

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

To address these questions, we develop a general equilibrium model with heterogeneous investors, passive demand, and financial constraints. We study the *aggregate (macro) elasticity* of the stock market, together with its inverse, the *price impact*, defined as the change in total stock market value in response to a $1 flow between (short-term) bonds and stocks.[2] Unlike micro elasticities, which capture substitution across individual stocks, aggregate elasticity reflects market wide reallocations between risky and risk-free assets. As a result, understanding aggregate elasticity requires tracking how the risk-free rate and the risk premium respond jointly to portfolio flows. General equilibrium is therefore essential to the analysis, not merely a refinement. A flow between bonds and stocks is by construction a shift in the relative demand for the two assets, so the price of the risk-free asset, the interest rate, cannot be held fixed while the stock price adjusts; it is itself a market clearing price. Capturing this joint adjustment of the interest rate and the risk premium is what makes the aggregate elasticity problem both important and difficult, since it requires solving for stock prices and the interest rate simultaneously as portfolio flows move through the market.

Our first main contribution is to characterize the aggregate elasticity of the endowment claim, the claim on aggregate output. This joint response of the interest rate and the risk premium is central to it. We show that, in the absence of frictions, the macro elasticity of the endowment

[1] See LeRoy and Porter (1981), Shiller (1981, 1992), and Cochrane (1992); Chen, Dou, and Kogan (2024) for a recent discussion; and Atkeson, Heathcote, and Perri (2026) for a recent macroeconomic accounting perspective on the persistent rise in U.S. corporate valuation ratios.

[2] We use "aggregate elasticity" and "macro elasticity" interchangeably. The price impact is the inverse of the elasticity: a more inelastic market has a larger price impact, so that a given flow moves prices by more. Gabaix and Koijen (2023) refer to the price impact as the "multiplier"; a price impact of 5, for example, means that a $1 flow into stocks raises their aggregate value by about $5.

claim is infinite, regardless of how price inelastic individual investors are.[3] Consider an outflow of funds from stocks to bonds. The resulting increase in the risk premium tends to lower stock prices. At the same time, the associated decline in wealth would reduce consumption. With a fixed endowment, this adjustment cannot occur in equilibrium. The interest rate therefore falls to clear the goods market. In frictionless economies, it adjusts until it exactly offsets the increase in the risk premium, leaving the price of the risky asset unchanged. This result holds regardless of how inelastic individual demand is. Aggregate elasticity is therefore infinite even when individual demands are highly inelastic.

This result highlights a central distinction between individual and aggregate elasticities. In the inelastic markets framework of Gabaix and Koijen (2023), the risk-free rate is constant and aggregate elasticity is the equity-weighted average of individual elasticities, so inelastic individual demand directly produces an inelastic aggregate market.[4] Building on that insight, we allow the interest rate to adjust endogenously and characterize a complementary general equilibrium channel through which aggregate elasticity also depends on how portfolio flows affect saving incentives. When flows do not alter these incentives, movements in the risk-free rate fully offset movements in the risk premium, so asset prices remain unchanged.

Aggregate elasticity becomes finite when risk is misallocated, that is, when investors hold too much or too little risk relative to the efficient allocation. Investors who are already overexposed to risk are less willing to save in assets that further increase their exposure, so portfolio reallocation alters saving incentives and interest rates can no longer fully offset changes in risk premia. This generates a nonzero price impact, even in general equilibrium. Frictions thus reduce aggregate elasticity. Even with frictions, such as passive demand and leverage constraints, the near cancellation between interest rates and risk premia remains quantitatively important, so the endowment claim exhibits a price impact of only 0.65 in our calibration. We establish these results first in a transparent two period model in Section 2 and then extend them to a fully dynamic economy in Section 3.

Because the interest rate plays a central role, a natural question is whether these conclusions survive once we recognize that the short-term rate is set by the central bank rather than determined in a frictionless bond market. We address this concern directly in Appendix D, where we embed

[3] A portfolio flow shock is an asymmetric shock to risky asset demand, affecting only a subset of investors and triggering portfolio reallocation in equilibrium. Asymmetric demand shocks of this kind also drive portfolio rebalancing in Alvarez and Atkeson (2018), who study a general equilibrium asset pricing model with idiosyncratic shocks to investor risk tolerance.

[4] In their basic model, the risk-free rate is taken as exogenous. In their general equilibrium extension, the rate is endogenously pinned down by the consumption Euler equation, but the assumption of i.i.d. growth implies that it is constant in equilibrium and therefore does not respond to flows. This is a natural starting point given the well-known difficulty of jointly determining asset prices and interest rates in heterogeneous agent general equilibrium with portfolio flows.

the mechanism in a production economy with sticky prices and an inflation-targeting central bank that sets the nominal rate through a policy rule. The central bank optimally implements exactly the real interest rate that prevails in our baseline endowment economy, so the rate movements we study describe the behavior of the natural rate of interest in response to portfolio flows. Aggregate elasticity is therefore unaffected by the introduction of an interest-rate-setting central bank, and the demand-based mechanism we emphasize operates in the same way once monetary policy is taken into account.

Studying aggregate elasticity in dynamic models with frictions is challenging because closed-form solutions are typically not available. In Section 4, we develop a tractable state-global perturbation method that yields closed-form, state-dependent characterizations of aggregate elasticity, building on Kogan and Uppal (2001) and Silva (2020). The approximation is local in the exogenous state variables but global in the endogenous wealth distribution, so it preserves the state dependence that drives why price impact rises in bad times. The analysis shows that elasticity varies with the wealth distribution and the allocation of risk across investors, with several reinforcing amplification mechanisms. Preference heterogeneity amplifies price impact by making risk misallocation state dependent: the efficient passive share rises when active wealth is low, so the gap between the passive sector's actual and efficient holdings widens in bad times. Binding leverage constraints concentrate aggregate risk on a smaller pool of unconstrained investors, further amplifying these effects. Excess volatility, in turn, results from the interaction of portfolio flow shocks with finite, state-dependent elasticity: neither inelasticity nor flows alone generate excess volatility, both are required.

Our second main contribution is to characterize the aggregate elasticity of the equity claim. Dividends are modeled as a time varying share of aggregate output (e.g., Abel, 1999; Menzly, Santos, and Veronesi, 2004), so equity is a levered claim on aggregate fundamentals. The rest of aggregate output is paid out to corporate debt holders, so the equity claim and the bond claim together add up to the endowment claim. In contrast to the endowment claim, the equity claim exhibits a substantially larger price impact of 5.26 in our calibration, close to the empirical benchmark of about 5 in Gabaix and Koijen (2023). This difference arises because the general equilibrium offset between the risk-free rate and the risk premium that renders the endowment claim highly elastic does not extend to the dividend claim. Since dividends are levered, the dividend risk premium responds more strongly to portfolio flows than the endowment risk premium. Consequently, changes in risk premia are not fully offset, and equity prices respond strongly to flows. This result highlights that modeling the dividend claim, rather than the claim on aggregate consumption as is standard in macro-finance, is essential for generating realistic price impact.

To incorporate the dividend claim in a tractable way, we establish conditions under which a version of the mutual fund theorem holds in our economy with frictions, so investors optimally hold

the dividend and bond claims at market capitalization weights. The equilibrium then admits a block recursive structure: the stochastic discount factor from the aggregate endowment claim economy can be used to price the dividend and bond claims in a second step. The problem therefore separates into an aggregate equilibrium and a simpler pricing problem, which keeps the dimensionality of the equilibrium manageable even with multiple state variables.

For the quantitative analysis in Section 5, we solve the full dynamic model using the deep neural network method of Duarte, Duarte, and Silva (2024) together with the vectorized automatic differentiation approach of Gopalakrishna and Wu (2026).[5] Unlike the perturbation method, which holds the wealth distribution fixed, the full solution accounts for the endogenous response of the wealth distribution to portfolio shocks when computing the aggregate price impact. We calibrate the economy to U.S. Flow of Funds sectoral holdings, following the aggregation in Koijen, Richmond, and Yogo (2024), and to the magnitude and persistence of quarterly equity flows as in Gabaix and Koijen (2023).

The calibrated model is quantitatively successful: it jointly matches the magnitude of aggregate price impact, standard asset pricing moments, and sectoral balance sheets, all from a single calibration. The dividend claim price impact is 5.26, close to the empirical benchmark of about 5 in Gabaix and Koijen (2023), and the model simultaneously delivers an equity premium of 5.14%, return volatility of 17.11%, and a risk-free rate of 2.01%, all close to their empirical counterparts. By contrast, the endowment claim price impact is only 0.65, an order of magnitude smaller, in line with the analytical distinction between aggregate and equity claims. Even so, the aggregate claim is far less elastic than classical frictionless models imply: a price impact of 0.65 corresponds to a macro elasticity of roughly 1.5 (the inverse of the price impact), well below the effectively infinite elasticity of the frictionless benchmark. Financial frictions therefore leave a quantitatively meaningful imprint even on the claim on aggregate output, not only on the levered equity claim. The model also matches sectoral asset to equity ratios in the Flow of Funds: the constrained active sector is levered, the unconstrained active sector is conservative, and the passive sector holds about 45 percent of risky assets. These successes are tightly connected. The same frictions that make the market sufficiently inelastic to match the empirical price impact also generate the equity premium and the level of return volatility, so the inelastic markets hypothesis and the standard asset pricing puzzles are addressed by a single mechanism.

The calibration also reveals substantial state dependence in price impact. Conditional on bad states, defined as those in which the price-dividend ratio falls below its 25th percentile, the model implied price impact rises to 5.41, larger than the unconditional value. A predictive regression of cumulative excess returns on the price-dividend ratio gives a slope that is roughly 2.3 times larger

[5] Fernández-Villaverde (2026) provides a recent overview of deep learning methods for solving dynamic economic models.

in bad states, and a standard log-linear return decomposition delivers a long-horizon discount rate share of 0.90, consistent with the evidence in Cochrane (2011). Counterfactuals trace these results to the joint role of preference heterogeneity, an inefficient average passive share, and flow shocks. Removing preference heterogeneity reduces the dividend claim price impact to 0.51. Setting the average passive risky share equal to unity reduces the dividend claim price impact to 2.55. Leverage constraints are important for the conditional dynamics, since they make price impact rise sharply when risk-bearing capacity is low. None of these moments can be matched in isolation: it is the interaction of frictions with portfolio flows that delivers the quantitative success.

These results connect to a broader empirical literature on demand-based asset pricing. Haddad, Huebner, and Loualiche (2025b) show that the rise of passive investing has made demand for individual stocks more inelastic, a finding that we extend to the aggregate level. Lochstoer and Muir (2026) provide complementary causal evidence that the secular rise of indexing has roughly doubled U.S. equity market volatility since the 1950s, and our counterfactuals offer a structural mechanism for that pattern. Parker, Schoar, and Sun (2023) document that the spread of target-date funds has dampened return volatility through more contrarian flows; in our framework, this corresponds to weakening the correlation between passive flows and fundamentals, which reduces both price impact and volatility. Taken together, the quantitative analysis supports the inelastic markets hypothesis of Gabaix and Koijen (2023) as a unified account of price impact, return volatility, and return predictability, in a general equilibrium economy with endogenous interest rates.

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 (GIV) identification of the aggregate price impact, which they estimate at about five, 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 analysis takes the risk-free rate as constant, a natural starting point given the well-known difficulty of jointly determining asset prices and interest rates in heterogeneous agent general equilibrium. Building on their findings, we tackle this joint determination directly: we develop a structural general equilibrium model in which heterogeneous investors, financial frictions, and an endogenous interest rate together generate inelasticity of the magnitudes they document, and we use deep learning methods to solve the resulting dynamic equilibrium. Johnson (2006) also studies aggregate price impact in a representative agent Lucas economy by perturbing risky asset supply, and finds finite elasticity

even in the frictionless benchmark.[6] 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 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 shows that even small cross-elasticities between equities and bonds generate large effects on macro elasticities, and it extends the analysis to dynamic settings where the persistence of shocks and the evolution of risk premia shape both short-run and long-run adjustments (Binsbergen, David, and Opp, 2025; He, Kondor, and Li, 2025; Davis, Kargar, Li, and Silva, 2026). In related work, Chien, Cole, and Lustig (2012) study whether intermittent portfolio rebalancing can account for the countercyclical volatility of the market price of risk, a channel similar in spirit to the role our passive sector plays in bad times.[7] Finally, our paper relates to preferred-habitat models (Vayanos and Vila, 2021; Kekre, Lenel, and Mainardi, 2025; Ray, Droste, and Gorodnichenko, 2024), which emphasize intermediary frictions and market segmentation. Our framework highlights wealth effects as a distinct channel through which shifts in aggregate wealth alter portfolio demand and amplify price movements.

The rest of the paper is organized as follows. Section 2 develops a transparent 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 reports the quantitative analysis, the counterfactuals, and the predictability and conditional price impact results. 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

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

[7] In their model, passive investors who do not rebalance after aggregate shocks push more risk onto a smaller pool of active investors in downturns, amplifying movements in risk premia. We share this risk-concentration channel but study its consequences for aggregate elasticity and price impact rather than for the volatility of risk compensation.

specify investor demands directly, rather than deriving them from utility functions. A micro-founded version of these demands, embedded in a dynamic heterogeneous-agents model, will be presented in Section 3.

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

The budget constraints of investor j are

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

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

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

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

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

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

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

Consumption is specified as

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

where $r \equiv \log R_f$ is the log risk-free rate. The consumption-wealth ratio $\overline{c}_j(\cdot)$ depends on both r and 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. We assume that $\overline{c}_j(\cdot)$ does not depend on the passive portfolio share $\overline{\alpha}_p$.

The market clearing conditions are

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

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

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

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

Investor j's demand for the risky asset is

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

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

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 only on the sum $(r + \pi)$, so its sensitivity to interest rates equals its sensitivity to risk premia. Assumption 1 imposes the same structure in this two-period setting, ensuring consistency with standard micro-founded models.

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

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

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

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

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

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

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

2.1 General equilibrium implications of portfolio flows

We now examine how the price-dividend ratio p and the riskless return r respond to a portfolio flow from the riskless bond into the risky asset. Let $\overline{\alpha}_p^*$ 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* is given by

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

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

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

Equilibrium. Let $s_j \equiv \omega_j Q_j^*$ 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 $\frac{\partial c_p}{\partial p}$ is not too large.

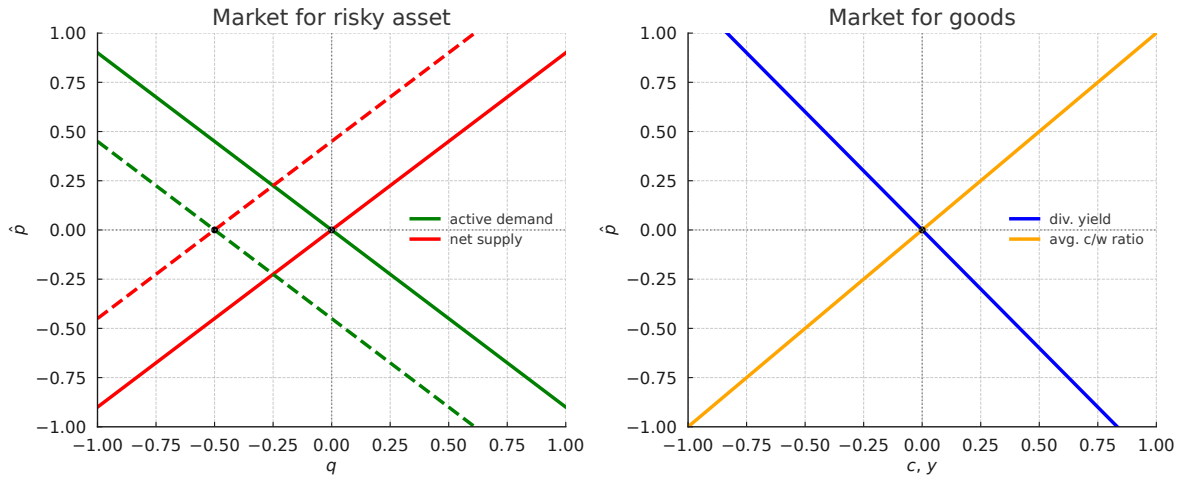


Figure 1. Equilibrium effects of a portfolio-flow shock in the risky-asset market (left) and goods market (right).

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

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

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

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

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

where

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

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

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

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

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

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

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

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

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

The disconnect between individual and aggregate elasticities arises from a general equilibrium effect. The market clearing condition for the risky asset market determines the risk premium, that is, the difference between the price of the risky asset and the risk-free rate. It does not pin down the individual responses of the price of the risky and riskless assets.

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

Consider a positive flow shock to the risky asset (e.g., funds shift toward stocks, so $f > 0$). Lower bond demand increases r, while increased demand for the risky asset reduces the risk premium. In equilibrium these effects offset exactly, leaving the price-dividend ratio unchanged. If the rise in r were smaller than the fall in the risk premium, there would be excess demand for goods

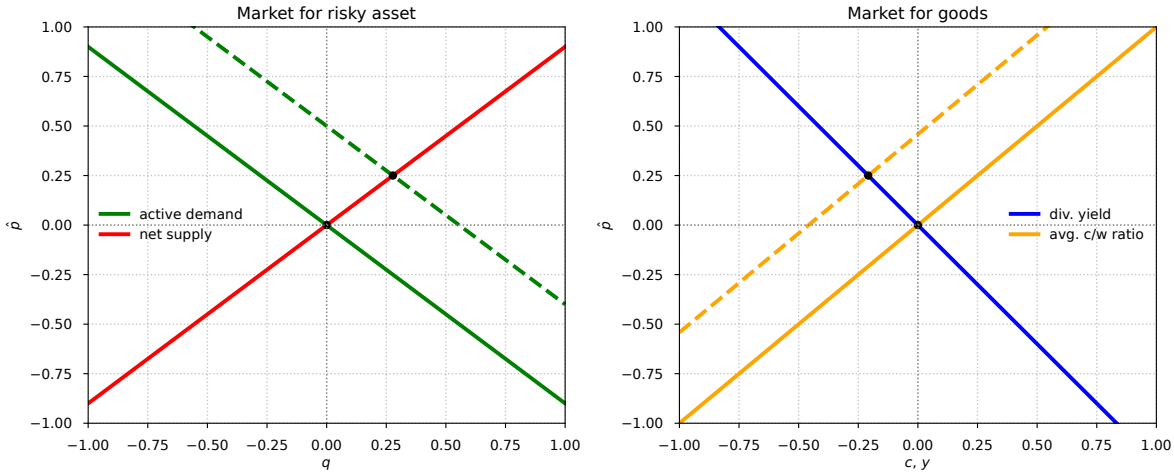


Figure 2. Equilibrium effects of a fundamental shock in the risky-asset market (left) and goods market (right).

(equivalently, excess bond supply).[9] Simultaneous clearing of risky and risk-free asset markets therefore requires infinite aggregate elasticity in this model.

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 dividend growth: $\hat{\mu} \equiv \mu - \mu^*$. In this case, the demand for the risky asset is still given by (4), but the demand shifter for active investors becomes $f_a = \frac{\partial \log \overline{Q}_a}{\partial \log \overline{\mu}} \times \hat{\mu}$. Similarly, the consumption-wealth ratio is given by:

$$\hat{c} = \chi_0 + \chi_p \hat{p}, \qquad \text{where} \qquad \chi_0 = -\chi_p \hat{\mu}.$$

Crucially, a fundamental shock simultaneously shifts the demand for the risky asset and the consumption-wealth ratio schedule. Figure 2 shows the equilibrium response to this shock. In contrast to a portfolio-flow shock, a fundamental shock generates a change in the price-dividend ratio whenever it affects the consumption-wealth ratio. For instance, in models with unit EIS, we have $\chi_p = \chi_0 = 0$, so the price-dividend ratio remains unchanged.[10] This result sheds light on why the market was infinitely elastic in response to a portfolio-flow shock, as we assumed those shocks do not affect the consumption-wealth ratio.

The need for a micro-founded demand system. The analysis underscores the central role of cross-price elasticity. When allowed to adjust, the interest rate has the potential to fully offset the impact of portfolio flows on the price-dividend ratio. The aggregate market can therefore be

[9] The equilibrium can be equivalently described in terms of the market for risky assets and the market for bonds, instead of the goods market. Appendix A.1 presents the bond-market view in detail.

[10] An analogous argument applies to uncertainty shocks. With unit EIS, higher uncertainty leaves prices unchanged. Higher uncertainty lowers the price-dividend ratio only when EIS exceeds one (see Bansal and Yaron, 2004).

infinitely elastic, even if individual investors are inelastic, so long as the average consumption-wealth ratio is unaffected by the flow shock f.

As with the fundamental shock, aggregate elasticity becomes finite when portfolio-flow shocks shift both the net-supply curve and the average consumption-wealth ratio. Understanding aggregate elasticity therefore requires a framework that links portfolio-reallocation shocks to both risky-asset demand and savings behavior. To this end, we develop a dynamic heterogeneous-agent asset pricing model and introduce a new methodology for deriving investor demand within this setting.

3 A Dynamic Asset Pricing Model with Passive Investors

We now develop a dynamic heterogeneous-agent asset pricing model that provides a micro-founded demand system. The economy extends the framework of Section 2 by embedding passive and active investors in a full general equilibrium setting.

Passive investors hold exogenously specified, stochastic portfolio shares in risky assets. Shocks to these portfolio shares generate portfolio flow (i.e., noise trading) shocks that are imperfectly correlated with cash flow shocks, so the economy is driven by both fundamental and non-fundamental demand shocks. Active investors, in contrast, optimally rebalance their portfolios subject to financial frictions. We model two types of active investors that differ in risk aversion: a high risk aversion (unconstrained) investor and a low risk aversion investor subject to a leverage constraint.

Unlike the model in Section 2, investors trade two distinct claims, a *dividend claim* (equity) and a *risky corporate bond*, whose payoffs sum to the aggregate endowment. Modeling the dividend process lets us study the response of equity prices, rather than of the aggregate claim, to portfolio flows. This distinction plays an important role quantitatively: the price impact of the dividend claim differs substantially from that of the aggregate claim and is key to matching the data.

3.1 Environment

Financial markets. We begin by specifying the aggregate endowment process. The aggregate endowment Y_t follows a geometric Brownian motion:

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

where $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}^{1\times d}$ are constants.[11] 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.

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

In the baseline model, $d = 3$, capturing three sources of risk: (i) aggregate endowment shocks, (ii) portfolio-flow shocks to passive investors, and (iii) shocks to the dividend share.

We consider a representative firm that issues two securities whose payoffs sum to the aggregate endowment. The first is a *dividend claim* (equity), with payoff flow $D_t = s_t Y_t$. The dividend share s_t follows the diffusion process:

$$ds_t = \mu_{s,t}\, dt + \sigma_{s,t}\, dZ_t, \tag{9}$$

given $s_0 > 0$, where the drift $\mu_{s,t} = \mu_s(s_t) \in \mathbb{R}$ and the diffusion $\sigma_{s,t} = \sigma_s(s_t) \in \mathbb{R}^{1\times d}$ are functions of the dividend share.[12] Dividends inherit both aggregate risk and additional variation from movements in the dividend share. By Itô's lemma, dividends follow:

$$\frac{dD_t}{D_t} = \mu_{D,t}\, dt + \sigma_{D,t}\, dZ_t, \tag{10}$$

where

$$\sigma_{D,t} = \sigma_D(s_t) = \sigma + \frac{\sigma_s(s_t)}{s_t}, \qquad \mu_{D,t} = \mu_D(s_t) = \mu + \frac{\mu_s(s_t)}{s_t} + \frac{\sigma\, \sigma_s(s_t)^\top}{s_t}.$$

Both the drift and the volatility of dividend growth are therefore functions of the dividend share s_t.

This specification allows dividends to be more volatile than consumption, a well-documented feature of the data (see, e.g., Abel 1999; Campbell 1999), and it generates predictable dividend growth and state-dependent dividend volatility through the dependence of $\mu_{D,t}$ and $\sigma_{D,t}$ on s_t.

The second claim is a *corporate bond* (risky debt), with payoff flow $B_t = (1 - s_t)Y_t$. Since $D_t + B_t = Y_t$, the bond payoff is smoother than aggregate consumption, reflecting the safer nature of debt and its lower exposure to aggregate risk. This decomposition allows us to distinguish between claims with different risk exposures, with important implications for price impact.

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

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

Since $D_t + B_t = Y_t$, the combined claim replicates the aggregate endowment, with price $P_{Y,t} =$

[12] One example is the levered-equity specification of Abel (1999), with $D_t = Y_t^\phi$ for a leverage parameter $\phi \geq 1$, so that $s_t = Y_t^{\phi-1}$, $\mu_s(s_t) = (\phi-1)\left[\mu + \frac{1}{2}(\phi-2)\|\sigma\|^2\right] s_t$, and $\sigma_s(s_t) = (\phi-1)\, s_t\, \sigma$, implying constant dividend growth volatility $\sigma_{D,t} = \phi\, \sigma$. Another is the mean-reverting specification of Menzly et al. (2004), with $\mu_s(s_t) = \kappa_s(\overline{s} - s_t)$ and $\sigma_s(s_t) = s_t(1 - s_t)\, \sigma_s$, where $\overline{s} \in (0,1)$, $\kappa_s > 0$, and $\sigma_s \in \mathbb{R}^{1\times d}$ are constants; this specification ensures $s_t \in (0,1)$ at all times and yields a stationary share process, so consumption and dividends are cointegrated.

$P_{D,t} + P_{B,t}$. The return on the aggregate claim follows:

$$dR_{Y,t} = \frac{dP_{Y,t} + Y_t\, dt}{P_{Y,t}} = \mu_{R_Y,t}\, dt + \sigma_{R_Y,t}\, dZ_t.$$

Investors also have access to a risk-free asset paying rate r_t and, to ensure markets are dynamically complete, a derivative in zero net supply that spans risks not directly traded through the dividend and bond claims. In particular, we assume the derivative does not load on dividend-share shocks. As in a futures contract, there is no initial cash exchange to enter into the derivative contract. The derivative provides risk exposure $\sigma_{R_F,t} \in \mathbb{R}^{1\times d}$, orthogonal to the aggregate claim, $\sigma_{R_F,t}\, \sigma_{R_Y,t}^\top = 0$, so that it provides independent risk exposure, with normalization $\|\sigma_{R_F,t}\| = \|\sigma_{R_Y,t}\|$.

Stochastic discount factor. Absence of arbitrage implies the existence of a stochastic discount factor (SDF), Λ_t, which prices all assets in the economy:

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

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

Absence of arbitrage implies that risk premia satisfy $\pi_{k,t} = \sigma_{R_k,t}\, \eta_t^\top$ for each claim $k \in \{D, B, F\}$. Let Σ_t denote the 3×3 matrix collecting the risk exposures of all traded assets:

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

and $\pi_t \equiv [\pi_{D,t}, \pi_{B,t}, \pi_{F,t}]^\top$, so that $\pi_t = \Sigma_t\, \eta_t^\top$.

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

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

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

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

redistributed across cohorts without affecting aggregate resources. This mortality structure ensures a nondegenerate stationary wealth distribution.

Preferences are recursive as in Duffie and Epstein (1992), with type-j risk aversion γ_j and a common elasticity of intertemporal substitution (EIS) ψ. The aggregator is given by:

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

where ρ denotes the discount factor.

Each investor chooses consumption $C_{j,t}$ and portfolio allocations across the dividend claim, the bond claim, and the derivative, subject to the constraints specified below.

Passive investors. Passive investors allocate a fraction $1-\overline{\alpha}_{p,t}$ of their wealth to the risk-free asset and a fraction $\overline{\alpha}_{p,t}$ to a *market index* fund, which holds the dividend and bond claims in proportion to their market-capitalization weights. Passive investors do not trade the derivative.

The share of wealth invested in the market index fund, $\overline{\alpha}_{p,t}$, evolves according to an Ornstein-Uhlenbeck process:

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

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

Portfolio flows may comove with cash-flow shocks or arise independently. We focus on the case where $\sigma = (\sigma_1, 0, 0)$ for aggregate endowment risk, $\sigma_{\alpha_p} = (\sigma_{\alpha_p,1}, \sigma_{\alpha_p,2}, 0)$ for the passive-share process, and $\sigma_s = (\sigma_{s,1}, 0, \sigma_{s,3})$ for the dividend share process. The first shock represents aggregate endowment risk, the second represents non-fundamental portfolio-flow shocks, and the third represents dividend-share shocks. We allow portfolio-flow shocks and dividend-share shocks to load on the aggregate endowment risk factor. This implies that aggregate endowment risk can be correlated with both portfolio flows and dividend-share shocks.

The exposure of the passive share to aggregate shocks can reflect *imperfect rebalancing* or *performance-induced flows*. Without rebalancing, a positive return on the risky asset mechanically increases the passive share when $\overline{\alpha}_{p,t} < 1$. Similarly, investors may increase risky exposure after good performance. This specification aligns with empirical evidence on passive behavior: households exhibit substantial portfolio inertia (Ameriks and Zeldes, 2004; Brunnermeier and Nagel, 2008); flows chase performance (Chevalier and Ellison, 1997; Dannhauser and Pontiff, 2024);

and regulatory changes can sharply increase equity allocations via TDF adoption (Parker, Schoar, Cole, and Simester, 2025), illustrating portfolio-flow shocks not tied to cash flow fundamentals.

Active investors. Investors of type $j = 1, 2$ are *active investors*. Active investors optimally rebalance their portfolios, choosing allocations across the dividend claim, the bond claim, and the derivative, $\alpha_{j,t} \equiv [\alpha_{j,D,t}, \alpha_{j,B,t}, \alpha_{j,F,t}]^\top$. Type-2 investors have low risk aversion, $\gamma_2 < \gamma_1$, and therefore have an incentive to lever up their exposure to risky assets. They borrow short-term to increase their risk exposure, subject to a leverage constraint that limits total portfolio risk:

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

where $\sigma_{j,t} \equiv \alpha_{j,t}^\top \Sigma_t = \alpha_{j,D,t}\,\sigma_{R_D,t} + \alpha_{j,B,t}\,\sigma_{R_B,t} + \alpha_{j,F,t}\,\sigma_{R_F,t}$ is the investor's total risk exposure and $\|\cdot\|$ denotes the Euclidean norm. This constraint resembles a Value-at-Risk (VaR) limit commonly applied to banks and leveraged institutions (see, e.g., Adrian and Shin 2014 and Brunnermeier and Pedersen 2009). Type-1 investors are more risk averse and do not face a leverage constraint.

Investors' problem. An investor of type $j = 0, 1, 2$ chooses consumption $C_{j,t}$ and portfolio allocations $\alpha_{j,t}$ to solve the problem:

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

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

with $W_{j,t} \geq 0$ and $\alpha_{j,t} \in \Omega_{j,t}$, where $\sigma_{j,t} = \alpha_{j,t}^\top \Sigma_t$ is the investor's risk exposure, $W_{j,t}$ denotes the investor's wealth, and $\Omega_{j,t}$ is the investment set for type-j investor.

Portfolio allocations are restricted to lie in an *investment set* $\Omega_{j,t}$, which depends on investor type and captures differences in financial constraints across investors.[13] For passive investors, $j = 0$, holdings of the market index fund induce indirect holdings of the dividend and bond claims, so their investment set is given by:

$$\Omega_{0,t} = \left\{ \alpha_{0,t} : \ \alpha_{0,D,t} = \overline{\alpha}_{p,t} \frac{P_{D,t}}{P_{Y,t}},\ \alpha_{0,B,t} = \overline{\alpha}_{p,t} \frac{P_{B,t}}{P_{Y,t}},\ \alpha_{0,F,t} = 0 \right\},$$

Type-1 investors are unconstrained, so their investment set is given by $\Omega_{1,t} = \mathbb{R}^3$. For type-2

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

investors, the investment set reflects their leverage constraint:

$$\Omega_{2,t} = \left\{\alpha_{2,t} : \|\Sigma_t^\top \alpha_{2,t}\| \leq \overline{\sigma}\right\}.$$

Market clearing and equilibrium. We define equilibrium as follows.

Definition 1. *A competitive equilibrium consists of stochastic processes, all adapted to the filtration generated by Z_t. These include the aggregate endowment Y, the dividend share s, the prices P_D and P_B, the risk-free rate r, the market price of risk η_t, the weight $\overline{\alpha}_p$, and, for each investor $j \in \{0, 1, 2\}$, wealth W_j, consumption C_j, and portfolio allocations α_j. These processes satisfy the following conditions:*

(i) The aggregate endowment Y_t evolves according to (8), with $Y_0 > 0$. The dividend share s_t evolves according to (9).

(ii) Given $\{P_{D,t}, P_{B,t}, r_t, \eta_t\}$ and the passive share $\overline{\alpha}_{p,t}$, the choices $(C_{j,t}, \alpha_{j,t})$ solve investor j's problem in (12). The weight on the market index fund, $\overline{\alpha}_{p,t}$, evolves according to (11).

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

$$\sum_{j=0}^{2} \omega_j C_{j,t} = Y_t, \qquad \sum_{j=0}^{2} \omega_j Q_{j,k,t} = 1, \qquad \sum_{j=0}^{2} \omega_j B_{j,t} = 0, \qquad \sum_{j=0}^{2} \omega_j F_{j,t} = 0,$$

for $k \in \{D, B\}$. Holdings are defined as follows. The number of units of the dividend and bond claims held by investor j are $Q_{j,D,t} = \alpha_{j,D,t} W_{j,t}/P_{D,t}$ and $Q_{j,B,t} = \alpha_{j,B,t} W_{j,t}/P_{B,t}$. Riskless bond holdings are given by $B_{j,t} = (1 - \alpha_{j,D,t} - \alpha_{j,B,t}) W_{j,t}$. The notional position in the derivative is $F_{j,t} = \alpha_{j,F,t} W_{j,t}$.[14]

3.2 Equilibrium characterization

We now characterize equilibrium in the dynamic model. We derive the optimality conditions that govern portfolio choice and consumption decisions, and show how they jointly determine prices and risk premia. We then establish a version of the *mutual fund theorem* in our economy with frictions. This result implies that investors hold the dividend and bond claims in proportion to their market-capitalization weights, so that aggregate equilibrium objects are independent of the dividend-share shock. As a result, the equilibrium admits a *block-recursive* structure, in which we first solve for the aggregate equilibrium and then recover the prices of individual claims. This representation will be central for our analytical characterization of aggregate market elasticity.

[14] As there is no initial cash exchange for the derivative, wealth invested in the risk-free asset equals total wealth minus the amount invested in the dividend and bond claims.

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 share. The optimal risk exposure for active investors $j = 1, 2$ 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 $\nu_{1,t} = 0$. 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 $\nu_{2,t}$ satisfies the complementary slackness condition:

$$\nu_{2,t}\left(\|\sigma_{2,t}\| - \overline{\sigma}\right) = 0.$$

For passive investors, risk exposure is given by

$$\sigma_{0,t} = \overline{\alpha}_{p,t}\sigma_{R_Y,t},$$

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:

$$\sum_{j=0}^{2} \omega_j W_{j,t} = P_{Y,t}.$$

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:

$$\sum_{j=0}^{2} x_{j,t} c_{j,t} = \frac{1}{p_{Y,t}}, \qquad \sum_{j=0}^{2} x_{j,t}\sigma_{j,t} = \sigma_{R_Y,t},$$

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:

$$\underbrace{\sum_{j=1}^{2} \frac{x_{j,t}}{x_{a,t}} \sigma_{j,t}}_{\substack{\text{active demand} \\ \text{for risk}}} = \underbrace{\chi_{a,t} \sigma_{R_Y,t}}_{\substack{\text{net supply} \\ \text{of risk}}},$$

where $\chi_{a,t} \equiv \frac{1-x_{0,t}\overline{\alpha}_{p,t}}{1-x_{0,t}}$ captures the share of aggregate risk that must be absorbed by active investors, and $x_{a,t} \equiv 1 - x_{0,t}$ 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 $h_{a,t} = \sum_{j=1}^{2} \frac{x_{j,t}}{x_{a,t}} \frac{1-\gamma_j^{-1}}{\psi-1} \frac{\sigma_{c_j,t}}{1+\nu_{j,t}}$ is the average hedging demand of active investors, adjusted for leverage constraints, and $\gamma_{a,t}$ is the average effective risk aversion:

$$\gamma_{a,t} \equiv \left[\sum_{j=1}^{2} \frac{x_{j,t}}{x_{a,t}} \frac{1}{\gamma_j (1+\nu_{j,t})} \right]^{-1}.$$

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:

$$r_t = \rho + \psi^{-1} \left(\mu_{P_Y,t} + \xi_t \right) + (1 - \psi^{-1}) \sum_{j=0}^{2} x_{j,t} \frac{\gamma_j}{2} \|\sigma_{j,t}\|^2 - \pi_{Y,t}, \tag{16}$$

where $\xi_t \equiv \sum_{j=0}^{2} x_{j,t}\,\xi_{j,t}$.

The first term reflects time preference. The second term captures intertemporal substitution. The third term and fourth terms capture the effect of uncertainty 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 $p_{D,t} \equiv P_{D,t}/Y_t$ and $p_{B,t} \equiv P_{B,t}/Y_t$ for the dividend and corporate bond claims, respectively. The valuation ratio of the aggregate claim is $p_{Y,t} \equiv P_{Y,t}/Y_t$, which satisfies $p_{Y,t} = p_{D,t} + p_{B,t}$.

The pricing condition for each claim $k \in \{D, B\}$ is given by

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

where $\pi_{k,t} = \sigma_{R_k,t}\,\eta_t^{\top}$ is the risk premium on claim k, $s_{D,t} = s_t$ and $s_{B,t} = 1 - s_t$. The terms $\mu_{p_k,t}$ and $\sigma_{p_k,t}$ denote the drift and diffusion of the valuation ratios, respectively.

State variables. The aggregate state vector consists of both endogenous and exogenous variables. The endogenous state variables are the wealth shares of active investors, $x_{1,t}$ and $x_{2,t}$, which summarize the distribution of wealth. The exogenous state variables are the passive portfolio share $\overline{\alpha}_{p,t}$ and the dividend share s_t.

Hence, the aggregate state vector is $X_t = (x_t, \overline{\alpha}_{p,t}, s_t)$, where $x_t = (x_{1,t}, x_{2,t})$. The processes $\overline{\alpha}_{p,t}$ and s_t evolve according to (11) and (9), respectively. To obtain $\mu_{X,t}$ and $\sigma_{X,t}$, we need to solve for the drift and diffusion of the endogenous state variables, $x_{1,t}$ and $x_{2,t}$.

By Itô's lemma, the law of motion for $x_{j,t}$, $j = 1, 2$, 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, $\sigma_{R_k,t} = \sigma + \sigma_{p_k,t}$, where $\sigma_{p_k,t}$ arises from movements in the valuation ratio $p_{k,t}$ for asset $k \in \{D, B\}$. By Itô's lemma,

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

Equation (18) shows that endogenous volatility depends on the sensitivity of the valuation ratio to three groups of variables: (i) the wealth distribution, (ii) the passive portfolio share, and (iii) the dividend share. The first term is standard in heterogeneous-agent models (see Panageas, 2020). The

third term is typical of models with multiple assets. 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 amplifies the sensitivity of prices to portfolio flows and therefore increases endogenous volatility.

Given the volatility of individual claims, the volatility of the aggregate claim follows from aggregation:

$$\sigma_{R_Y,t} = \sum_{k=D,B} \frac{p_{k,t}}{p_{Y,t}} \sigma_{R_k,t}.$$

It remains to determine $\sigma_{R_F,t}$. We make three assumptions about $\sigma_{R_F,t}$: (i) it does not load on the dividend-share shock, so $\sigma_{R_F,t,3} = 0$; (ii) it is orthogonal to the aggregate claim, so $\sigma_{R_F,t}\, \sigma_{R_Y,t}^{\top} = 0$; (iii) it has the same norm as the aggregate claim, so $\|\sigma_{R_F,t}\| = \|\sigma_{R_Y,t}\|$. These conditions uniquely determine $\sigma_{R_F,t}$ given $\sigma_{R_Y,t}$.

Markov equilibrium. 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 ratios $p_k(X_t)$. 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 five partial differential equations.

These equations jointly determine the consumption-wealth ratios $\{c_j(X_t)\}_{j=0}^{2}$ and the valuation ratios $\{p_k(X_t)\}_{k\in\{D,B\}}$ as functions of four state variables: the two active wealth shares $x_{1,t}$ and $x_{2,t}$, the passive portfolio share $\overline{\alpha}_{p,t}$, and the dividend share s_t.

3.2.2 Block recursivity and the mutual fund theorem

We show that the system of PDEs admits a block-recursive structure. First, we solve for the consumption-wealth ratios and the price-dividend ratio of the aggregate claim in a reduced set of states $\hat{X}_t = (x_t, \overline{\alpha}_{p,t})$. Second, we solve for the valuation ratios of the dividend and corporate bond claims in the full set of states $X_t = (x_t, \overline{\alpha}_{p,t}, s_t)$.

This provides a substantial simplification, as the equilibrium allocation and pricing kernel are identical to those of an economy with a single claim on the aggregate endowment and a derivative. At the same time, the block-recursive structure allows us to recover the pricing implications for individual claims. In particular, we can characterize how portfolio flows affect the market elasticity of the dividend claim.

An important implication of the block-recursive structure is that a version of the *mutual fund theorem* holds in our economy. Despite financial frictions, investors optimally hold risky claims in proportion to their market-capitalization weights, as in the classic mutual fund theorem.

Proposition 2 (Block recursivity and the mutual fund theorem)**.** *Define the reduced set of states as* $\hat{X}_t = (x_t, \overline{\alpha}_{p,t})$*, where* $x_t = (x_{1,t}, x_{2,t})$ *is the vector of wealth shares of the active investors. Then,*

(i) ***Block recursivity:*** *The consumption-wealth ratios,* $\{c_{j,t}(\hat{X}_t)\}_{j=0}^{2}$*, and the price-dividend ratio for the aggregate claim,* $p_Y(\hat{X}_t)$*, are a function of the reduced set of states* $\hat{X}_t$*, and they satisfy the following system of PDEs:*

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

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

(ii) ***Mutual fund theorem:*** *The portfolio shares of active investors are given by:*

$$
\alpha_{j,k} = \alpha_{j,Y}\frac{p_k}{p_Y}, \quad \forall k \in \{D, B\},
$$

so $Q_{j,D} = Q_{j,B}$*, where* $\alpha_{j,Y}$ *denotes the holdings of the market index fund for investor* j*, which satisfies the condition*

$$
\alpha_{j,Y} = \frac{\sigma_j \sigma_{R_Y}^\top}{\sigma_{R_Y}\sigma_{R_Y}^\top} \tag{20}
$$

Proposition 2 shows that the equilibrium with separate dividend and bond claims is equivalent to that of an economy with a single claim on the aggregate endowment and a derivative.

Two assumptions are key to this result. First, the aggregate endowment process is independent of the dividend share s_t. This assumption is consistent with, e.g., Menzly et al. (2004), but differs from the two-trees model of Cochrane, Longstaff, and Santa-Clara (2008), where the dividend share directly affects consumption volatility.

Second, markets are dynamically complete. In equilibrium, investors are not exposed to the dividend-share shock, which can be perfectly diversified by holding the dividend and bond claims in proportion to their market-capitalization weights. Investors can therefore obtain any desired exposure to the fundamental and portfolio-flow shocks by trading the market index fund and the derivative. Without the derivative, investors would need to deviate from market-capitalization weights to adjust their exposure to these risks, at the cost of bearing dividend-share risk.

Pricing the individual claims. Having solved for the aggregate objects $\{c_{j,t}(\hat{X}_t)\}_{j=0}^{2}$ and $p_Y(\hat{X}_t)$, we next recover the valuation of individual claims. The individual valuation ratios are functions of the full state vector $X_t = (x_t, \overline{\alpha}_{p,t}, s_t)$, as they load on the dividend-share shock.

Given the aggregate equilibrium objects $r(\hat{X}_t)$ and $\eta(\hat{X}_t)$, and using Itô's lemma to express the drift and diffusion of valuation ratios as functions of the state dynamics and their derivatives, the pricing condition in Equation (17) yields a linear PDE for $p_{D,t}(X_t)$ and $p_{B,t}(X_t)$.

Thus, individual claim prices are obtained in a second step. This separation highlights that portfolio flows affect aggregate quantities through the reduced state vector, while their impact on individual asset prices operates through valuation in the full state space.

This completes the characterization of the equilibrium with separate dividend and bond claims. 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 aggregate market elasticity is infinite around efficient allocations, even when individual investors have finite demand elasticities. Thus, portfolio frictions play a crucial role in delivering a 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 risky asset 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,*

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

$c_j^* = (p_Y^*)^{-1}$, *and* $\alpha_{j,Y}^* = 1$ *for* $j = 0, 1, 2$.

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. When $\epsilon = 1$, we recover the full heterogeneous-agent economy with frictions.

First, investors have heterogeneous risk aversion:

$$\gamma_j = \gamma\big(1 + \hat{\gamma}_j \epsilon\big), \qquad \sum_{j=0}^{2} \omega_j \hat{\gamma}_j = 0,$$

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 risky asset. 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 $\sigma_p = \hat{\sigma}_p \epsilon$ 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 dividend share is given by $s = \overline{s} + \hat{s}\epsilon$ and the diffusion coefficient satisfies $\sigma_s = \hat{\sigma}_s \epsilon$, where $\overline{s}$ is the reference point for the dividend share. Hence, the perturbation of the dividend share is analogous to the one for the passive portfolio share.

Fourth, the leverage constraint coefficient is perturbed according to

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

so that the tightness of the constraint is of order $O(\epsilon)$, 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. From the block recursivity property in Proposition 2, the consumption-wealth ratios and the price-dividend ratio of the aggregate claim depend on the reduced set of states $\hat{X}_t = (x_t, \overline{\alpha}_{p,t})$. We expand them to second order in ϵ as follows:

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

for $j = 0, 1, 2$.

Unlike standard DSGE perturbations, we do not assume that endowment volatility σ is small. Moreover, the approximation is local in the exogenous state variables $\overline{\alpha}_p$ and s, 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 (2020).[15]

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.[16] 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: $\mu_{\hat{X}} = O(\epsilon)$ and $\sigma_{\hat{X}} = O(\epsilon)$. 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: $p_{Y,\hat{X}}(\hat{X}) = O(\epsilon)$ and $c_{j,\hat{X}} = O(\epsilon)$.

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

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

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 risky assets. By Proposition 2, investors hold the dividend and bond claims in market-capitalization weights. Hence, risky asset holdings coincide across claims: $Q_{j,D,t} = Q_{j,B,t} \equiv Q_{j,t}$, where[17]

$$Q_j(x, \overline{\alpha}_p; \epsilon) = \alpha_{j,Y}(x, \overline{\alpha}_p; \epsilon)\,\frac{x_j}{\omega_j}.$$

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

[16] 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).

[17] For a discussion of the implications of the mutual fund theorem for trading volume, see Lo and Wang (2000).

Define the proportional deviation of risky asset demand from the benchmark allocation as

$$q_{j,t} \equiv \frac{Q_j(x_t, \overline{\alpha}_{p,t}; \epsilon) - Q_j(x_t, \overline{\alpha}_{p,t}; 0)}{Q_j(x_t, \overline{\alpha}_{p,t}; 0)} = \alpha_{j,Y,t} - 1.$$

Expanding the risk exposure in Equation (14), and combining it with the expression for $\alpha_{j,Y}$ in Equation (20), yields the first-order correction to the portfolio share of active investors:

$$\alpha_{j,Y,t} - 1 = \frac{\hat{\pi}_{Y,t}}{\gamma \|\sigma\|^2} - (\hat{\gamma}_j + \hat{\nu}_{j,t})\,\epsilon + O(\epsilon^2),$$

where $\hat{\pi}_{Y,t} \equiv \pi_{Y,t} - \pi^*_Y$ denotes the deviation of the risk premium from the benchmark, and $\hat{\nu}_{j,t}$ denotes the Lagrange multiplier on the leverage constraint. Hence, since hedging demand is second order in ϵ, the optimal portfolio share depends only on the 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:

$$\hat{\pi}_{Y,t} = -\hat{r}_t - \frac{1}{p^*_Y}\hat{p}_{Y,t} + O(\epsilon^2),$$

where $\hat{r}_t \equiv r_t - r^*$ 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:

$$q_{j,t} = -\zeta_{j,p}\,\hat{p}_{Y,t} - \zeta_{j,r}\,\hat{r}_t + f_{j,t} + O(\epsilon^2), \tag{21}$$

where, for unconstrained investors,

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

For passive and constrained investors, own- and cross-price elasticities vanish, $\zeta_{j,p} = \zeta_{j,r} = 0$. 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 (21) 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 captures the difference between the price of risky and risk-free assets.

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 $f_{0,t}$, 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 $\hat{\pi}_{Y,t} = -\hat{r}_t - \frac{1}{p_Y^*}\hat{p}_{Y,t}$ 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

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

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 asset and goods markets:

$$\sum_{j=0}^{2} x_{j,t}\,q_{j,t} = 0, \qquad \sum_{j=0}^{2} x_{j,t}\,\hat{c}_{j,t} = -y^*\,\hat{p}_{Y,t},$$

where $y^* = 1/p_Y^*$.

Aggregating the demand system across investors yields

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

where $\zeta_{k,t} \equiv \sum_{j=0}^{2} x_{j,t}\zeta_{j,k}$, for $k \in \{p, r\}$, $f_t \equiv \sum_{j=0}^{2} x_{j,t}f_{j,t}$ and $\chi_{0,t} \equiv \sum_{j=0}^{2} x_{j,t}\chi_{j,0}$.

Aggregate market elasticity. We define the *inverse aggregate market elasticity* as the proportional price response of the aggregate claim to a flow shock f_t:

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

The market is elastic when it can absorb large flows with only modest price movements. This notion differs from the standard micro elasticity, which measures the response of an individual investor's demand holding everything else fixed (see, e.g., Koijen and Yogo, 2019). Instead, aggregate market elasticity is an equilibrium object: a flow shock affects multiple asset prices simultaneously and induces endogenous adjustments in risk premia, interest rates, and savings behavior. Our focus is therefore on the aggregate equilibrium response to portfolio-flow shocks rather than on individual demand elasticities.[18]

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:

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

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 asset 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

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

[18] For a discussion of why the price impact of portfolio flows is, in general, not the same as the inverse of the micro elasticity of individual demand, see, e.g., He et al. (2025).

Proposition 3 shows that the main lesson from the simple model in Section 2 extends to the dynamic economy. Portfolio inflows into the risky asset reduce the risk premium, while the shift out of bonds raises the risk-free 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 of this result is that the aggregate market elasticity is independent of individual investors' own-price elasticities $\zeta_{j,p}$. By increasing average risk aversion γ, individual investors can be made arbitrarily inelastic, 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 the demand for the risky asset, the risk premium must adjust in equilibrium to clear the risky asset market. But this only pins down the *difference* in the response of the price of risky and risk-free assets. To determine the absolute level of the price-dividend ratio and the interest rate, 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 $\sigma_{j,t}$:

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

The term in brackets is equal to zero at the privately optimal portfolio, as implied by the investor's optimality condition in Equation (14). By the envelope theorem, small changes in $\alpha_{j,t}$ around the optimum have negligible effects on welfare and hence do not affect $c_{j,t}$.

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

This reasoning also 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.

Definition of aggregate market elasticity. Up to second order, the price-dividend ratio for the aggregate claim takes the form:

$$\hat{p}_{Y,t} = a_Y(x_t)\frac{(\overline{\alpha}_{p,t}-1)^2}{2} + b_Y(x_t)(\overline{\alpha}_{p,t}-1) + c_Y(x_t) + O(\epsilon^3), \tag{22}$$

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

The aggregate market elasticity corresponds to the inverse of the price impact of a passive flow, i.e., a change in $x_{0,t}(\overline{\alpha}_{p,t}-1)$. In this case, the (inverse) aggregate market elasticity is given by:[19]

$$\varepsilon_{Y,t}^{-1} \equiv \frac{1}{x_{0,t}}\frac{\partial \hat{p}_{Y,t}}{\partial \overline{\alpha}_{p,t}} = \frac{a_Y(x_t)}{x_{0,t}}(\overline{\alpha}_{p,t}-1) + \frac{b_Y(x_t)}{x_{0,t}} + O(\epsilon^2).$$

Hence, to characterize the market elasticity, we need to determine the coefficients $a_Y(x_t)$ and $b_Y(x_t)$.

Diffusion and drift of the price-dividend ratio. The diffusion of the price-dividend ratio is

$$\sigma_{p_Y} = \frac{p_{Y,\overline{\alpha}_p}}{p_Y}\sigma_p + O(\epsilon^3),$$

using the fact that $p_{Y,x} = O(\epsilon^2)$ and $\sigma_x = O(\epsilon)$ in the case without preference heterogeneity.

We can then write return volatility as follows:

$$\sigma_{R_Y,t} = \sigma + x_{0,t}\varepsilon_{Y,t}^{-1}\sigma_p + O(\epsilon^3). \tag{23}$$

The expression above shows that there is a tight connection between aggregate market elasticity, portfolio-flow shocks, and excess return volatility in our economy. To generate excess volatility, the aggregate market must be sufficiently inelastic and the economy must be exposed to portfolio-flow shocks. If the market is either perfectly elastic, $\varepsilon_{Y,t}^{-1} = 0$, or not exposed to flow shocks, $\sigma_p = 0$, there is no excess volatility.

[19] In this definition, we consider the effect of a change in $\overline{\alpha}_{p,t}$ holding $x_{0,t}$ fixed. Appendix B.3 shows that the endogenous response of $x_{0,t}$ to a portfolio-flow shock affects prices only at higher order. Hence, it does not contribute to aggregate market elasticity up to second order.

The drift of the price-dividend ratio is given by

$$\mu_{p_Y} = \frac{p_{Y,\overline{\alpha}_p}}{p_Y}\mu_{\overline{\alpha}_p} + \frac{1}{2}\frac{p_{Y,\overline{\alpha}_p\overline{\alpha}_p}}{p_Y}\sigma_p\sigma_p^\top + O(\epsilon^3).$$

Given the process for portfolio share $\overline{\alpha}_{p,t}$ and the coefficients in Equation (22), we obtain:

$$\mu_{p_Y,t} = \left[a_Y(x_t)(\overline{\alpha}_{p,t} - 1) + b_Y(x_t)\right]\kappa_p(1 - \overline{\alpha}_{p,t}) + \frac{1}{2}a_Y(x_t)\sigma_p\sigma_p^\top + O(\epsilon^3).$$

Demand for the risky asset. The second-order demand for the risky asset still takes the form in Equation (21), with the same definition of own and cross-price elasticities. The demand shifter for passive investors is simply $f_{0,t} = \overline{\alpha}_{p,t} - 1$, while the demand shifter for active investors now captures the effect of price dynamics:

$$f_{j,t} = \underbrace{\zeta_{j,r}(\mu_{p_Y,t} + \sigma_{p_Y,t}\sigma^\top)}_{\text{expected price appreciation}} + \underbrace{\frac{1-\gamma^{-1}}{\psi-1}\frac{\sigma_{c_j,t}\sigma^\top}{\sigma\sigma^\top}}_{\text{hedging demand}} - \underbrace{2\frac{\sigma_{p_Y,t}\sigma^\top}{\sigma\sigma^\top}}_{\text{endogenous volatility}}.$$

Given the current price-dividend ratio $\hat{p}_{Y,t}$ and interest rate $\hat{r}_t$, active investors have a higher demand for risky assets when there is a positive expected price appreciation or hedging demand. Holding expected returns fixed, greater endogenous volatility reduces demand for risky assets. The fact that portfolio-flow shocks generate endogenous volatility and hedging demand in a dynamic setting, thereby affecting risky-asset demand, is closely related to the results in He et al. (2025).

Consumption-wealth ratio. Consider next the consumption-wealth ratio. Aggregating (13) across investors, letting $c_t \equiv \sum_{j=0}^{2} x_{j,t}c_{j,t}$, yields

$$c_t = \psi\rho + (1-\psi)\underbrace{\left[r_t + \pi_{Y,t} - \frac{\gamma\|\sigma_{R_Y,t}\|^2}{2}\sum_{j=0}^{2}x_{j,t}\alpha_{j,Y,t}^2\right]}_{\text{risk-adjusted expected return}} + \xi_t.$$

The expression above illustrates how the *misallocation* of the risky asset affects the consumption-wealth ratio. Under the efficient allocation $\alpha_{j,Y,t} = 1$, so $\sum_j x_{j,t}\alpha_{j,Y,t}^2 = 1$. Any deviation from this efficient allocation creates dispersion in portfolios, leading to $\sum_j x_{j,t}\alpha_{j,Y,t}^2 > 1$ by Jensen's inequality, which lowers the risk-adjusted expected return. When $\psi > 1$, this reduction weakens incentives to save and raises the consumption-wealth ratio for a given $r_t + \pi_{Y,t}$.

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 $\chi_p = (\psi - 1)y^*$ and $\chi_{0,t}$ is given by

$$\chi_{0,t} = (\psi - 1)\frac{\gamma\|\sigma\|^2}{2}\frac{x_0}{x_a}(\overline{\alpha}_{p,t} - 1)^2 - \psi\left[\mu_{p_Y,2} + (1-\gamma)\sigma_{p_Y,2}\sigma^\top\right].$$

The first term captures the effect of the misallocation of the risky asset. It is equal to zero when $\overline{\alpha}_{p,t} = 1$, that is, when passive investors initially hold the efficient portfolio. The second term captures the effect of price dynamics. Importantly, both terms respond to portfolio-flow shocks.

The aggregate market elasticity. As in the first-order analysis, while the market clearing condition for the risky asset is informative about the risk premium, the market clearing condition for goods pins down the price-dividend ratio. In particular, the price-dividend ratio is proportional to the consumption-wealth ratio shifter $\chi_{0,t}$:

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

Notice that the risky asset demand elasticities $\zeta_{j,p}$ and $\zeta_{j,r}$ do not enter into the expression above, so they do not directly drive the price-dividend ratio or the aggregate market elasticity. Instead, the price impact is governed by how portfolio flows impact the consumption-wealth ratio.

Using the expression for $\mu_{p_Y,t}$ and $\sigma_{p_Y,t}$, we can write the shifter $\chi_{0,t}$ as a state-dependent quadratic function of the passive portfolio share $\overline{\alpha}_{p,t}$. Matching coefficients, we can solve for $a_Y(x_t)$, $b_Y(x_t)$, and $c_Y(x_t)$. Given these coefficients, we can then determine the aggregate market elasticity. The next proposition provides the main result.

Proposition 4 (Aggregate elasticity: inefficient passive demand)**.** *Let $x_{a,t} = 1 - x_{0,t}$ denote the wealth share of active investors. Without preference heterogeneity or leverage constraints, the inverse aggregate market elasticity is given by:*

$$\varepsilon_{Y,t}^{-1} = \frac{1-\psi^{-1}}{y^* + 2\kappa_p}\frac{\gamma\|\sigma\|^2}{x_{a,t}}\left[1 - \overline{\alpha}_{p,t} + \frac{\gamma - 1}{y^* + \kappa_p}\sigma_p\sigma^\top\right] + O\left(\epsilon^2\right). \tag{24}$$

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 (24).

First, the market is infinitely elastic in the unit EIS case, $\psi = 1$. As it is well known, the

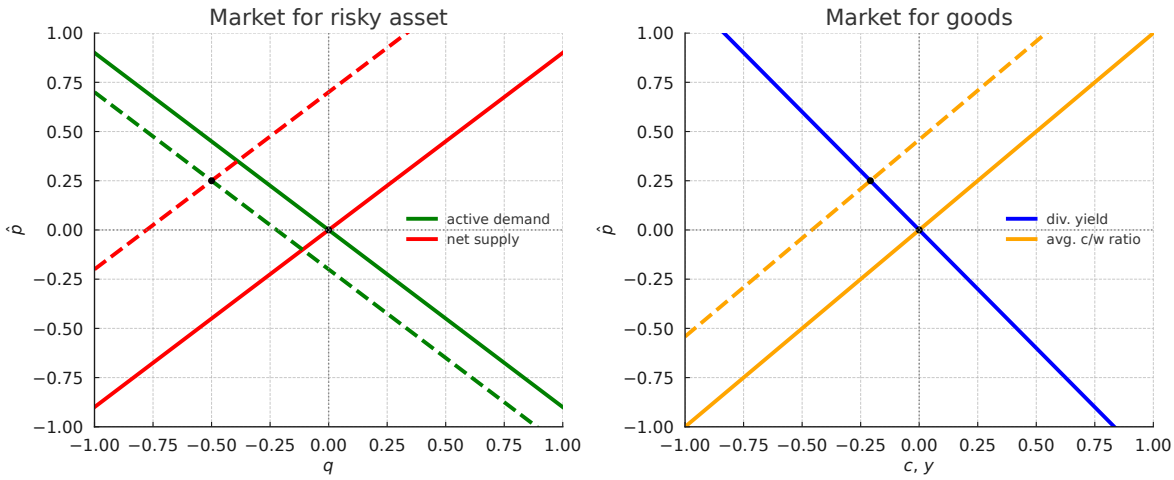


Figure 3. Equilibrium with inefficient risk allocation: risky asset (left) and goods (right) markets.

price-dividend ratio is pinned down by the subjective discount rate in this case, so all 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 $\sigma_p = 0$, aggregate market elasticity is finite as long as the initial allocation of the risky asset 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 the demand for the risky asset 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 asset reduces risk misallocation. When $\psi > 1$, this increases the incentive to save, reducing the consumption–wealth ratio.[20] 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, price impact is positive even if the initial allocation of the risky asset is efficient, $\overline{\alpha}_{p,t} = 1$, when portfolio-flow shocks are positively correlated with fundamental shocks, $\sigma_p \sigma^\top > 0$. The mechanism again operates through the consumption-wealth ratio. As shown in Equation (23), portfolio-flow shocks create excess volatility in the risky asset return, which through a precautionary effect reduces the consumption-wealth ratio when $\psi > 1$, raising the price-dividend ratio.[21]

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 under capitalized. Moreover, a reduction in active wealth share raises the price impact by more

[20] 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).

[21] In the case $\sigma_p \sigma^\top = 0$, the impact of σ_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.

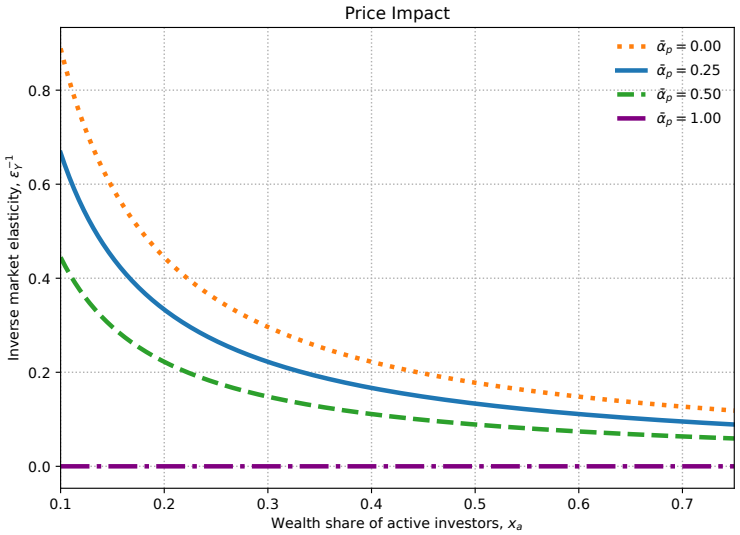


Figure 4. Price impact: inefficient passive demand.

when passive investors are further from the efficient allocation. For simplicity, we focus on the case $\sigma_p = 0$, so there is no price impact when the passive portfolio is initially efficient.

Fifth, the market elasticity depends on the persistence of portfolio-flow shocks, κ_p. In particular, the market becomes infinitely elastic when portfolio-flow shocks are transitory, that is, when $\kappa_p \to \infty$. This highlights that the mechanism in the model is fundamentally distinct from the market microstructure literature, which studies short-horizon trading frictions rather than equilibrium reallocations of aggregate risk over longer horizons. In contrast, our mechanism is closer to a recent literature emphasizing how the *market macrostructure* affects asset prices (see, e.g., Haddad and Muir, 2025).

Risk premium and interest rate. The next proposition characterizes the asset pricing implications in more detail.

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*

$$\eta_t = \gamma \chi_{a,t} \sigma + \left[1 - H_{a,t}\right] \gamma x_{0,t} \varepsilon_{Y,t}^{-1} \sigma_p + O(\epsilon^3), \tag{25}$$

where

$$\chi_{a,t} \equiv \frac{1 - x_{0,t}\overline{\alpha}_{p,t}}{1 - x_{0,t}}, \qquad H_{a,t} \equiv \frac{1 - \gamma^{-1}}{\psi - 1}\left[\frac{\psi}{x_{a,t}} - 1\right].$$

2. *The risk premium is given by*

$$\pi_{Y,t} = \eta_t \sigma^\top + \gamma x_{0,t} \varepsilon_{Y,t}^{-1} \sigma_p \sigma^\top + O(\epsilon^3). \tag{26}$$

3. *The interest rate is given by*

$$r_t = r^* + (1 - \psi^{-1}) \frac{\gamma \|\sigma\|^2}{2} \frac{x_0}{x_a} (\overline{\alpha}_{p,t} - 1)^2 - (\eta_t - \eta^*) \sigma^\top + O(\epsilon^3).$$

Equation (25) 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 asset, less aggregate risk must be absorbed by active investors, reducing the market price of risk.

The second term captures the effect of endogenous volatility, $x_{0,t}\varepsilon_{Y,t}^{-1}\sigma_p$. This implies that portfolio-flow shocks command a non-zero risk premium in equilibrium. The magnitude of this effect is attenuated by the hedging demand of active investors, summarized by the coefficient $H_{a,t}$.

The risk premium reflects movements in both the price and the quantity of risk, as shown in Equation (26). These two terms depend on the endogenous volatility generated by the interaction of portfolio-flow shocks and inelastic aggregate demand. Thus, the same mechanism that generates excess volatility also generates a risk premium for portfolio-flow shocks.

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.

Most importantly, the interest rate does not perfectly offset movements in the market price of risk, because it is affected by the same risk-misallocation term that drives aggregate market elasticity. This is a manifestation of the general equilibrium mechanism characterized in Proposition 4.

4.4 The role of monetary policy

So far, we have characterized aggregate market elasticity in an endowment economy in which the interest rate adjusts endogenously to clear the goods market. In practice, however, short-term interest rates are controlled by the central bank. A natural question is therefore whether introducing monetary policy changes the determination of aggregate market elasticity. In particular, if the central bank sets the interest rate, does the endogenous interest-rate adjustment mechanism survive?

In Appendix D, we consider an extension of the baseline model with production, sticky prices,

and monetary policy. We derive two main results. First, we show that the flexible-price equilibrium of the production economy is identical to the equilibrium of the endowment economy in Section 3. Hence, all our results carry over directly to the production setting.

Second, we introduce sticky prices and a central bank that follows a Taylor rule targeting zero inflation. We show that an inflation-targeting central bank implements the same interest rate as in the baseline model. Therefore, aggregate market elasticity in the economy with monetary policy is identical to the one characterized in the baseline model.

To understand this result, consider a portfolio-flow shock from stocks to bonds. Suppose the central bank were to keep the real interest rate fixed. If the portfolio-flow shock does not directly affect the consumption-wealth ratio, the decline in asset prices would reduce consumption, leading to a fall in aggregate demand and output.[22] The resulting recession would create deflationary pressures. To stabilize inflation, the central bank would have to reduce the real interest rate by exactly the same amount implied by the baseline model.

Hence, our results characterize the behavior of the *natural* interest rate, defined as the real interest rate in the flexible-price economy, in response to portfolio-flow shocks. An inflation-targeting central bank therefore tracks the same natural-rate adjustment mechanism characterized in the baseline model.

4.5 Amplification mechanisms

We now characterize aggregate market elasticity in richer environments, capturing the role of preference heterogeneity, leverage constraints, and multiple claims. These ingredients amplify the sensitivity of prices to portfolio flows by weakening the economy's effective risk-bearing capacity and generating additional state dependence in the allocation of risk.

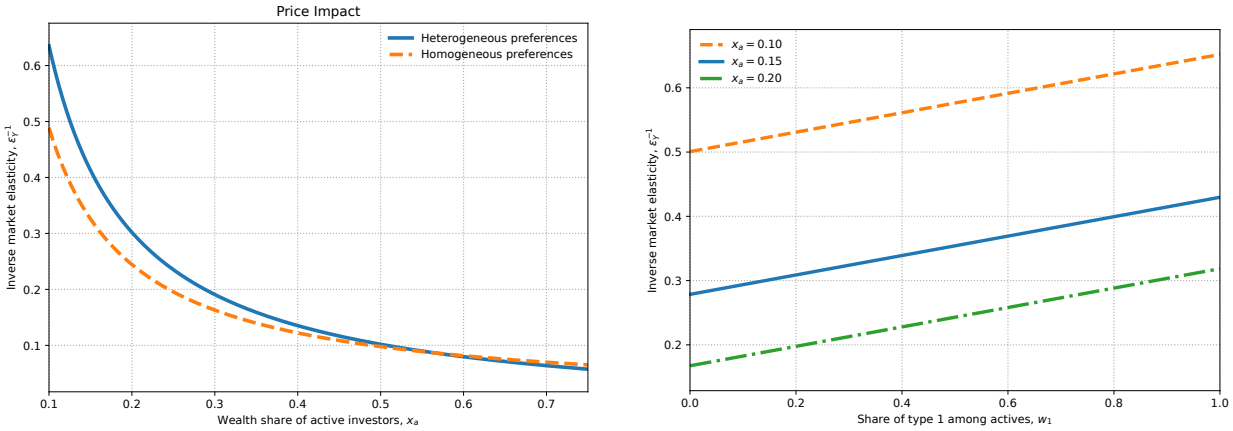
Preference heterogeneity and leverage constraints. The next proposition characterizes the aggregate market elasticity in the case with preference heterogeneity and leverage constraints.

Proposition 6 (Aggregate elasticity: heterogeneity and leverage constraints)**.** *Suppose* $\overline{\alpha} = 1 + \hat{\alpha}\epsilon$. *Then, the inverse aggregate market elasticity is given by*

$$\varepsilon_{Y,t}^{-1} \;=\; \left(1-\psi^{-1}\right)\frac{\gamma\,\|\sigma\|^2}{y^*+2\kappa_p}\,\frac{1-x_{c,t}}{x_{u,t}}\left[\overline{\alpha}_{p,t}^{opt}-\overline{\alpha}_{p,t}+\frac{\kappa_p(\overline{\alpha}_{p,t}^{opt}-\overline{\alpha})}{y_{Y,0}+\kappa_p}+\frac{(\gamma-1)\sigma_p\sigma^\top}{y_{Y,0}+\kappa_p}\right]\;+\;O\Big(\epsilon^2\Big),$$

where $x_{c,t}$ *is the wealth share of constrained active investors,* $x_{u,t} \equiv x_{a,t} - x_{c,t}$ *is the wealth share of*

[22] The mechanism is related to models with wealth effects from asset price revaluations, as in Caballero and Simsek (2020) and Caramp and Silva (2026). For evidence on this mechanism, see, e.g., Cieslak and Vissing-Jorgensen (2021) and Chodorow-Reich, Nenov, and Simsek (2021).



(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.

unconstrained active investors, and $\overline{\alpha}_{p,t}^{\text{opt}}$ *is the passive investors' optimal portfolio share, given by*

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

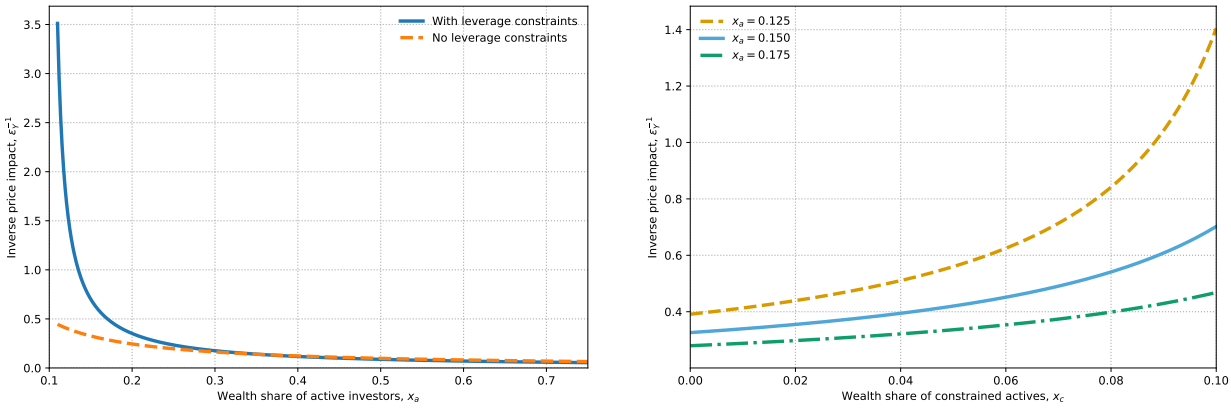
with $\mathbb{E}_t^u[\gamma_j]$ *denoting the wealth-weighted average risk aversion of unconstrained active investors.*

Proposition 6 characterizes aggregate market elasticity in the full model. 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 corresponds to the portfolio that satisfies the first-order condition for the passive investor's portfolio choice problem. In the absence of preference heterogeneity, the optimal passive share is simply $\overline{\alpha}_{p,t}^{\text{opt}} = 1$, so we recover the result from Proposition 4 after imposing $\overline{\alpha} = 1$.

Preference heterogeneity and leverage constraints imply that the optimal passive share is time-varying. In particular, the optimal passive share tends to rise in bad times. When passive investors are more risk averse than active investors ($\gamma_0 > \mathbb{E}_t^u[\gamma_j]$) and leverage constraints are slack, passive investors should hold more of the risky asset when active investors are less capitalized:

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

Hence, the market becomes endogenously more inelastic after shocks that reduce active wealth. First, active investors' risk-bearing capacity declines. Second, the gap between the optimal and actual passive portfolio shares widens, pushing the economy farther from the efficient allocation. Together, these forces generate a sharp increase in price impact as $x_{a,t}$ declines, as shown in the left panel of Figure 5.



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

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

Figure 6. Price impact under binding leverage constraints. Panel (a) compares the cases with and without leverage constraints (holding $x_{c,t} = 0.10$). Panel (b) plots the price impact as a function of $x_{c,t}$ for different values of $x_{a,t}$.

The distribution of wealth among active investors also matters. An increase in the share of high risk-aversion active investors, $w_{1,t} \equiv x_{1,t}/x_{a,t}$, 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 $x_{a,t} = 1 - x_{0,t}$. To isolate the role of constraints, we set $\gamma_0 = \mathbb{E}_t^u[\gamma_j]$, removing pure preference heterogeneity effects. As $x_{a,t}$ declines, price impact rises more sharply under binding constraints. Panel (b) holds $x_{a,t}$ fixed and varies the constrained share $x_{c,t}$. Markets therefore become more inelastic as a larger share of active investors becomes constrained.

Price impact of the dividend 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. The next proposition characterizes the elasticity of the dividend claim, for the case in which the dividend share is mean reverting, as in Menzly et al. (2004).

Proposition 7 (Aggregate elasticity: dividend-claim)**.** *Suppose the dividend share follows the mean-reverting process of Menzly et al. (2004),* $ds_t = \kappa_s(\bar{s} - s_t)\,dt + s_t(1 - s_t)\,\sigma_s\,dZ_t$*, with constants* $\bar{s} \in (0,1)$*,* $\kappa_s > 0$*, and* $\sigma_s \in \mathbb{R}^{1\times d}$*. Let* $\varepsilon_{D,t}^{-1} \equiv \frac{1}{x_{0,t}} \frac{\partial \hat{p}_{D,t}}{\partial \bar{\alpha}_{p,t}}$ *denote the inverse elasticity of the dividend claim and define the dividend-share beta:* $\beta_s \equiv \bar{s}(1-\bar{s})\frac{\sigma_s \sigma^\top}{\sigma\sigma^\top}$*. Then,*

$$\varepsilon_{D,t}^{-1} = \varepsilon_{Y,t}^{-1} + \frac{\beta_s}{\bar{s}} \frac{y^*}{y^* + \kappa_p} \frac{\zeta_{r,t}^{-1}}{y^* + \kappa_s} + O(\epsilon^2),$$

where $\zeta_{r,t}^{-1} \equiv \frac{\gamma\|\sigma\|^2}{x_{u,t}}$ *is the inverse of the cross-elasticity of the demand for risky assets.*

Proposition 7 shows that the inverse elasticity of the dividend claim is the sum of the inverse elasticity of the aggregate claim and an additional term capturing the effect of dividend-share shocks. Hence, the equity market can exhibit finite elasticity even when the aggregate market is infinitely elastic.

Consider, for example, the case where $\overline{\alpha}_{p,t} = 1$ and $\sigma_p = 0$, so that $\varepsilon_{Y,t}^{-1} = 0$. In this case, the interest rate exactly offsets the change in the risk premium of the aggregate claim. However, the interest rate does not offset the change in the risk premium of the dividend claim.

When the dividend-share beta is positive, the equity premium responds more strongly to portfolio-flow shocks than the risk premium on the aggregate claim. As shown in Proposition 3, the response of $\pi_{Y,t}$ to a portfolio-flow shock is determined by the inverse of the cross-elasticity of risky asset demand, $\zeta_{r,t}^{-1}$. The magnitude of the amplification depends on the levered beta of the dividend claim, summarized by the ratio $\beta_s/\overline{s}$.

Finally, the price impact depends both on the persistence of portfolio-flow shocks, governed by κ_p, and on the persistence of dividend-share shocks, governed by κ_s. This result helps reconcile the relatively small elasticity of the aggregate claim with the large empirical price impact observed in equity markets.

Taking stock. This section provides a detailed characterization of the determinants of aggregate market elasticity in a general equilibrium setting with frictions. Aggregate market elasticity is finite only when portfolio flows directly affect 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 price impact in bad times by reducing effective risk-bearing capacity. Equity markets exhibit larger price impact because dividends are a levered claim on aggregate output, so equity prices respond more strongly to portfolio flows than the aggregate claim. We next provide a quantitative evaluation of how these different mechanisms interact to generate inelastic markets.

5 Quantitative Implications

We now evaluate the quantitative implications of the dynamic model of Section 3. The exercise has two goals. The first is to assess whether a calibration disciplined by U.S. data on sectoral portfolio holdings and passive equity flows can simultaneously deliver an empirically plausible price impact and the standard asset pricing moments, which are difficult to match jointly. The second is to quantify how much of return predictability and the variance of the price-dividend ratio

is attributable to how portfolio flows interact with preference heterogeneity, leverage constraints, and passive-sector risk misallocation. We do the latter by comparing the baseline economy to a set of counterfactuals that turn off these frictions one at a time.

For the quantitative analysis, we calibrate a version of the dividend process in which the drift and volatility of dividend growth in Equation (10) are constant, $\mu_{D,t} = \mu_D$ and $\sigma_{D,t} = \sigma_D$. Dividend growth is then i.i.d., more volatile than consumption growth, and imperfectly correlated with it. This specification is consistent with the dividend dynamics in Bansal and Yaron (2004), once their long-run risk and stochastic volatility components are shut down, and we discipline the dividend moments using the targets reported there.

We solve the full dynamic system numerically using the neural-network method of Duarte et al. (2024), which accommodates the three state variables: passive demand and the wealth shares of the constrained and unconstrained active sectors. We follow Gopalakrishna and Wu (2026) and use vectorized automatic differentiation to compute derivatives and approximate the set of equilibrium conditions. Appendix F.3 describes the numerical procedure in detail.

5.1 Calibration

We aggregate Flow of Funds (FoF) sectors into three ultimate sectors: a passive sector, a constrained active sector, and an unconstrained active sector.[23] The passive sector is modeled as an exogenous Ornstein-Uhlenbeck process for the portfolio share $\overline{\alpha}_{p,t}$ in Equation (11). The two active sectors differ both in risk aversion and in whether they are subject to a leverage constraint. Table 1 reports the calibrated parameter values, and the next subsection describes the moments used to discipline them.

Figure 7 motivates two of the calibration targets. The left panel plots quarterly aggregate equity flows, scaled by lagged market capitalization, together with realized return volatility. Appendix F.2 describes the construction of the flow series in detail, following Appendix D of Gabaix and Koijen (2023). The right panel reports the "volatility multiplier," defined as the ratio of realized return volatility to quarterly flows. The multiplier is a reduced-form analog of the price impact that our calibration aims to reproduce.

5.2 Moments

We discipline the calibration using three blocks of moments: technology parameters, preference and cross-sectional parameters, and the passive-demand process.

[23] We follow the aggregation in Koijen et al. (2024). Appendix F.1 describes the sector mapping and the treatment of pass-through vehicles in detail.

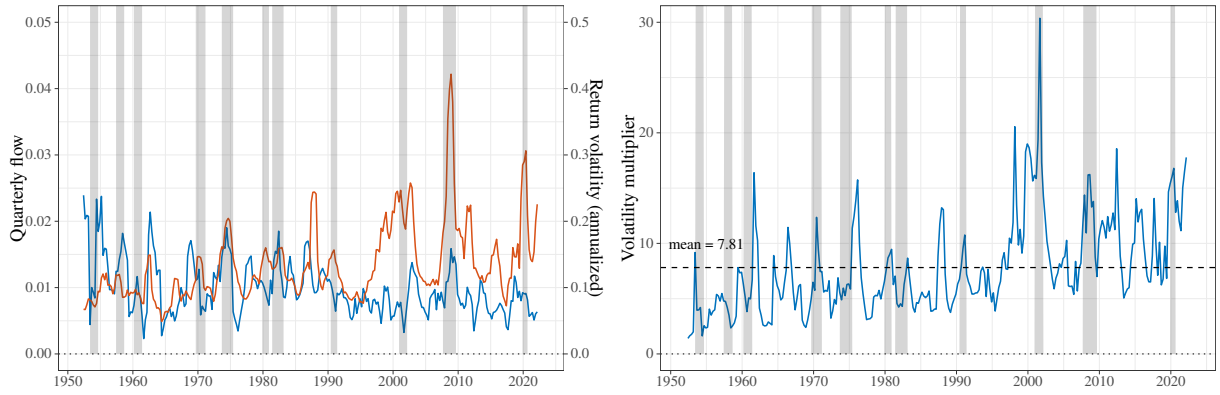


Figure 7. Quarterly flows and return volatility.
The left panel plots quarterly aggregate equity flows in blue and realized return volatility in red. Quarterly flows are scaled by lagged aggregate market capitalization, Appendix F.2 describes the construction of the series in detail. The right panel plots the "volatility multiplier," defined as the ratio of return volatility to flow; this multiplier is a reduced-form analog of the price impact targeted in the calibration. Source: Flow of Funds and CRSP.

Technology. We set the endowment growth rate to $\mu = 1.8\%$ and endowment volatility to $\|\sigma\| = 3.0\%$, consistent with U.S. aggregate consumption data (Campbell and Cochrane, 1999). We set the expected dividend growth μ_D to 1%, close to the endowment growth rate. We set the volatility of dividend growth $\|\sigma_D\|$ to 14.3%, and the correlation of dividend and consumption growth ν_D to 59%, both close to the corresponding moments reported in Bansal and Yaron (2004).

Preferences and cross section of holdings. We set a common elasticity of intertemporal substitution $\psi = 1.5$ for all agent types, in line with the value used in Bansal and Yaron (2004). The effective discount rate ρ is chosen to match the average real risk-free rate, yielding $\rho = 0.021$. The sector-specific risk aversion coefficients γ_j are chosen to match sectoral *asset shares*, defined as the value of risky holdings of sector j divided by the aggregate value of risky holdings. This leads to an average risk aversion of 10.52, within the range of values used in the macro-finance literature. The sector masses ω_j are chosen to match the asset-to-equity ratios of the active sectors. The agent turnover rate $\kappa = 0.02$ is calibrated so that the wealth share of the passive sector in the simulated ergodic distribution matches its empirical counterpart of about 0.39. The leverage constraint tightness $\overline{\sigma}$ is set so that the asset-to-equity ratio of the constrained sector matches its Flow of Funds counterpart.

Passive demand. The long-run mean $\overline{\alpha}$ is set to the household-weighted aggregate implied by our FoF aggregation, which yields $\overline{\alpha} = 0.52$. This value is intermediate between the one-third estimate in Chinco and Sammon (2024), inferred from index-reconstitution events, and the roughly two-thirds estimate in Koijen et al. (2024), which uses active-share measures in the spirit of Cremers and Petajisto (2009). The volatility $\|\sigma_p\|$ is disciplined by the observed dispersion of quarterly

	Parameter	Choice
	Preferences & distribution	
ψ	EIS	1.50
γ_0	Risk aversion of passive investors	15.86
γ_1	Risk aversion of unconstrained investors	35.52
γ_2	Risk aversion of constrained investors	6.15
ρ	Effective discount rate	0.021
κ	Agent turnover rate	0.02
ω_0	Share of passive agents	0.65
ω_1	Share of unconstrained agents	0.30
ω_2	Share of constrained agents	0.05
	Technology	
μ	Endowment growth rate	0.018
$\|\sigma\|$	Endowment volatility	0.030
$\|\sigma_D\|$	Dividend volatility	0.143
μ_D	Dividend growth rate	0.01
ν_D	Corr. of dividend and consumption shock	0.59
	Passive demand	
$\overline{\alpha}$	Mean	0.52
κ_p	Mean reversion parameter	0.14
σ_p	Passive demand shock exposures	0.015
ν_p	Corr. of passive demand and fundamental shock	0.76
	Leverage constraints	
$\overline{\sigma}$	Tightness of the leverage constraint	0.07

Table 1. Parameter values.
This table reports the parameter values used to calibrate the baseline economy. Risk aversion parameters and sectoral masses (γ_j, ω_j) are chosen to match average sectoral wealth shares and sectoral "betas" from time-series regressions of risky asset holdings on aggregate risky supply. Technology parameters $(\mu, \sigma, \mu_D, \sigma_D, \nu_D)$, where $\nu_D = \frac{\sigma\sigma_D^\top}{\|\sigma\|\,\|\sigma_D\|}$, are standard macro–finance targets. Passive-demand parameters $(\overline{\alpha}, \kappa_p, \sigma_p, \nu_p)$, where $\nu_p = \frac{\sigma\sigma_p^\top}{\|\sigma\|\,\|\sigma_p\|}$, are disciplined by the average passive portfolio share, household portfolio inertia (Brunnermeier and Nagel, 2008), and the volatility and correlation of quarterly aggregate equity flows with market returns (Gabaix and Koijen, 2023). The leverage-constraint tightness $\overline{\sigma}$ targets the empirical asset-to-equity ratio of the constrained sector.

aggregate equity flows shown in the left panel of Figure 7. We construct this series from Flow of Funds sectoral holdings as detailed in Appendix F.2: we 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. The resulting series replicates the quarterly flows in their Figure D.7. The persistence κ_p is informed by evidence on household portfolio inertia. Using PSID data, Brunnermeier and Nagel (2008) regress the actual change in the household risky asset share on a counterfactual change that assumes no rebalancing, and find an inertia coefficient near 0.75, implying that portfolio shares adjust slowly to capital gains and to

Moment	Model	Data
Panel A. Dividend-claim moments		
Risk premium (%)	5.1	6.3
Risk-free rate (%)	2.0	0.9
Return volatility (%)	17.1	19.4
Quarterly flows (%)	0.6	0.7
Price impact	5.26	5.0
Panel B. Endowment-claim moments		
Risk premium (%)	1.4	—
Return volatility (%)	3.5	—
Price impact	0.65	—
Panel C. Sector balance-sheet moments		
Assets/Equity (unconstrained)	0.2	0.5
Assets/Equity (constrained)	1.9	1.6
Asset share (passive, %)	43.8	38.8
Asset share (unconstrained, %)	17.6	20.1
Asset share (constrained, %)	38.5	29.2

Table 2. Unconditional moments.
This table reports model-implied and empirical unconditional moments for the dividend claim and the endowment claim in the baseline economy. The empirical moments for the risk premium, risk-free rate, and return volatility are from Bansal and Yaron (2004). The price impact estimate is from Gabaix and Koijen (2023). Model moments are computed from a 1,200-year monthly simulation of the calibrated economy. The price impact is estimated as the slope coefficient $\hat{b}_1$ of the regression $\Delta \log(P/D)_t = a_0 + b_1 \, \overline{\alpha}^s_{p,t} \, x_{p,t-1} + e_t$, where $\overline{\alpha}^s_{p,t}$ is the innovation to passive demand. Asset/Equity denotes the ratio of risky asset holdings to equity for the constrained and unconstrained sectors; asset shares report the fraction of the aggregate risky market held by each sector. All values are annualized.

flow-induced changes in wealth.[24] We choose κ_p so that the half-life of the passive-demand process in the simulated model matches the half-life implied by this inertia coefficient. Finally, we estimate the loading $\nu_p = \frac{\sigma \sigma_p^\top}{\|\sigma\| \|\sigma_p\|}$ of passive-flow shocks on the aggregate shock indirectly by regressing the quarterly passive-flow series on contemporaneous market returns. This yields a model-implied correlation between flows and returns of 0.006, close to the empirical value of 0.011.

5.3 Unconditional moments

Price impact regression. We simulate the calibrated economy at monthly frequency for 1,200 years and report unconditional moments in Table 2. 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

[24] Parker et al. (2023) show that the adoption of target-date funds has further stabilized household portfolio shares.

passive demand, scaled by the passive sector's lagged wealth share:

$$\Delta \log(P/D)_t = a_0 + b_1 \, \overline{\alpha}^s_{p,t} \, x_{p,t-1} + e_t, \tag{27}$$

where $\overline{\alpha}^s_{p,t}$ denotes the innovation to passive demand.

Equation (27) 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 and mirrors the size-weighting in their granular instrumental variable (GIV) estimator.[25]

In our model, passive-demand innovations are observed and exogenous by construction, so OLS directly identifies the model-implied counterpart of the GIV estimate without requiring an instrument.

Price impact and asset pricing moments. The model simultaneously matches a broad set of empirical facts, including asset pricing moments, the magnitude of portfolio flows, price impact, and sectoral balance sheets.

As shown in Panel A of Table 2, the model generates a large price impact, with an estimated slope $\hat{b}_1 = 5.26$, close to the empirical benchmark of about 5 (Gabaix and Koijen, 2023). For the dividend claim, the model also matches standard asset pricing moments, delivering a risk premium of 5.14%, a risk-free rate of 2.01%, and a return volatility of 17.11%. As we show below, these two successes are closely related: generating realistic price impact for the dividend claim is crucial for matching asset pricing moments such as the risk premium and return volatility.

A key driver of these results is the explicit modeling of the dividend process. Panel B of Table 2 shows that the price impact for the aggregate endowment claim is only 0.65, an order of magnitude smaller than for the dividend claim. This implies an aggregate market elasticity of roughly 1.5, substantially smaller than the infinite elasticity implied by frictionless models, but still much larger than the empirical elasticity of about 0.2 estimated for the equity market. While financial frictions generate finite aggregate market elasticity, the explicit modeling of the dividend process is essential for matching the larger price impact observed in equity markets.

The model also matches sectoral balance sheets, as shown in Panel C of Table 2. The constrained sector is levered (asset-to-equity ratio of 1.9 in the model versus 1.6 in the data), the unconstrained sector is conservative (0.2 versus 0.5), and the passive sector holds about 43.8% of risky assets.

[25] Gabaix and Koijen (2023), following Gabaix and Koijen (2024), decompose sector-level flows into a common component and an idiosyncratic shock, and construct the granular instrumental variable (GIV) as the difference between size-weighted and equal-weighted averages of these idiosyncratic shocks. This instrument isolates exogenous variation in aggregate flows and addresses the simultaneity bias in OLS, since prices and flows are jointly determined.

5.4 Counterfactuals: the role of frictions

To quantify the role of each friction, we compare the baseline economy to a sequence of counterfactuals that progressively shut down individual frictions. We vary one parameter at a time, holding all others fixed at their baseline values in Table 1, so each column isolates the marginal contribution of a given friction.

The baseline economy (column 1) simultaneously delivers a large price impact, a sizable risk premium, and high return volatility, consistent with the data. Removing the leverage constraint (column 2) reduces the price impact to 1.21 and attenuates both the risk premium and return volatility, but these remain substantial. Eliminating preference heterogeneity (column 3) reduces the price impact sharply to 0.51, implying a very large market elasticity.

The next counterfactual vary passive demand. Setting $\overline{\alpha} = 1$ (column 4) raises the passive sector's exposure to risk and reduces the degree of risk misallocation, lowering the price impact to 2.55.

Finally, when passive flows are uncorrelated with fundamentals (column 5), so that $\sigma\sigma_p^\top = 0$, the model generates a small price impact for the dividend claim. This points to the importance of imperfect rebalancing and performance-induced flows, which induce correlation between passive flows and fundamentals, in generating price impact.

These results point to a tight connection between price impact, risk premium, and return volatility. The model generates asset pricing moments consistent with the data only when the market is sufficiently inelastic. In versions of the model with modest or no price impact, the risk premium and return volatility are also smaller. These results highlight that matching both price impact and asset pricing moments requires the interaction of frictions with shocks to passive demand.

Related evidence on passive behavior. These results connect closely to recent empirical evidence on the effects of passive investing. Haddad et al. (2025b) show that the rise of passive investing has made the demand for individual stocks more inelastic. Our results extend this insight to the aggregate level, showing that the presence of passive investors plays a central role in determining aggregate market elasticity.

Lochstoer and Muir (2026) provide causal evidence that the secular rise of market-index trading has roughly doubled U.S. stock market volatility since the 1950s, even as macroeconomic fundamentals have become calmer. Our counterfactuals offer a complementary structural mechanism, highlighting how an increased importance of portfolio flows amplifies both price impact and return volatility.

Parker et al. (2023) show that the introduction of target-date funds (TDFs) led to more contrarian flows, ultimately dampening volatility. In our framework, the expansion of TDFs can be interpreted

	Baseline	No leverage constraints	No preference heterogeneity	High passive portfolio share	Flows uncorrelated with fundamentals
r (%)	2.01	2.31	2.14	2.43	2.07
π_D (%)	5.14	3.32	3.90	3.82	4.92
$\sigma_{R,D}$ (%)	17.11	14.98	14.55	16.06	16.82
π_Y (%)	1.43	1.15	1.34	1.08	1.41
$\sigma_{R,Y}$ (%)	3.55	3.33	3.07	3.21	3.52
PI (Y)	0.65	0.35	0.08	0.18	0.07
PI (D)	5.26	1.21	0.51	2.55	1.24
Cond. PI (D)	5.41	1.56	0.57	3.48	1.48
Γ	10.52	11.37	10.52	13.31	10.58

Table 3. Counterfactual experiments.
This table reports annualized model-implied moments for the baseline economy and four counterfactuals. *No leverage constraints* sets $\overline{\sigma} = 10{,}000$, so leverage constraints are slack. *No preference heterogeneity* sets $\gamma_c = \gamma_p = \gamma_u = 10.52$, eliminating heterogeneity in risk aversion, combined with the slack leverage constraint. *High passive portfolio share* sets $\overline{\alpha} = 1$. *Flows uncorrelated with fundamentals* sets $\nu_p = 0$, so passive-demand shocks are orthogonal to fundamentals. All other parameters are as in Table 1. Price impacts PI (Y) and PI (D) denote the slope of the regression in equation (27) for the endowment and dividend claim, respectively. Conditional price impact restricts the regression to "bad states" in which $(P/D)_t$ is below its 25th percentile. The variable Γ reports the unconditional mean of the wealth-weighted harmonic mean of the risk aversion parameters.

as reducing the correlation between passive portfolio shares and fundamentals. Imperfect rebalancing induces a positive correlation between passive demand and return shocks, which is weakened by the rebalancing behavior of TDFs. Consistent with this evidence, our counterfactuals show that reducing this correlation lowers both price impact and return volatility.

5.5 Return predictability and the variance of the price-dividend ratio

The tight link between price impact and return volatility documented above has a natural counterpart in the literature. Since Cochrane (1992), excess volatility has been understood as reflecting variation in discount rates rather than in expected cash flows. In an inelastic market, portfolio flows move the price-dividend ratio without changing expected dividend growth, so the resulting variation must be absorbed by expected returns. The P/D ratio should therefore forecast future excess returns, with a slope that is itself informative about market inelasticity. We examine this prediction in the simulated baseline economy.

Figure 8 plots cumulative future excess returns against the current P/D ratio. Table 4 reports the corresponding predictive regression

$$\sum_{k=1}^{T} r_{t+k} - r_t = a + b\,(P/D)_t + \varepsilon_{t+T}, \qquad T = 20 \text{ quarters},$$

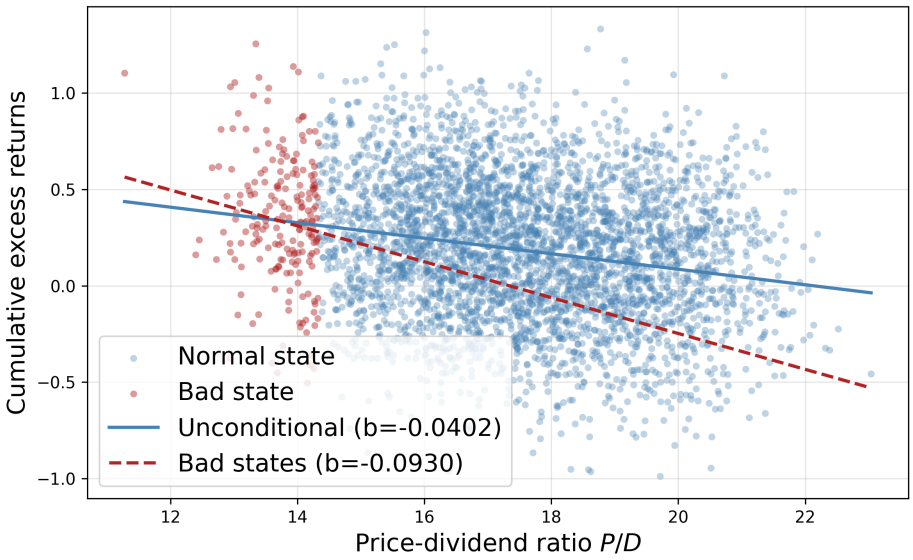


Figure 8. Return predictability.
Scatter of cumulative future excess returns ($T = 20$ quarters) against the current price-dividend ratio in the simulated baseline economy. Blue dots denote normal states; red dots denote bad states with $(P/D)_t$ below the 5th percentile of its long-run distribution. Lines show the corresponding OLS fits.

estimated on the full sample and on "bad states" in which $(P/D)_t$ falls below its 5th percentile. The slope is negative in both samples and about 2.3 times larger in magnitude in bad states. Risk premia are therefore countercyclical: when the P/D ratio is low, expected returns are high, and the P/D ratio is especially informative about them.

Following Campbell and Shiller (1988) and Cochrane (2011), we decompose the unconditional variance of $pd_t \equiv \log(P/D)_t$ as

$$\mathrm{Var}(pd_t) = \mathrm{Cov}\left(pd_t, \sum_{k=1}^{\infty} \rho^{k-1} \Delta d_{t+k}\right) - \mathrm{Cov}\left(pd_t, \sum_{k=1}^{\infty} \rho^{k-1} r_{t+k}\right).$$

Scaling by $\mathrm{Var}(pd_t)$ gives the discount-rate share

$$\beta_r = -\frac{\mathrm{Cov}\left(pd_t, \sum_{k=1}^{\infty} \rho^{k-1} r_{t+k}\right)}{\mathrm{Var}(pd_t)},$$

which we compute at a 60-quarter horizon. In the baseline economy, $\beta_r = 0.90$: nearly all variation in pd_t is driven by expected returns.

Taken together, these results provide a demand-based explanation of the excess volatility puzzle of Shiller (1981) and LeRoy and Porter (1981). Consistent with the inelastic market hypothesis of Gabaix and Koijen (2023), modest portfolio flows generate pronounced price fluctuations and predictable returns, even in an economy where the interest rate responds endogenously to those

	Dependent variable: $\sum_{k=1}^{20}(r_{t+k} - r_t)$	
	(1) Full sample	(2) Bad states
price-dividend ratio, $(P/D)_t$	−0.0402 (0.0015)	−0.0930 (0.0268)
R^2	0.045	0.016
Observations	14,400	720
Bad-state threshold	$(P/D)_t <$ 5th percentile	

Table 4. Return predictability regression.
OLS estimates of the predictive regression $\sum_{k=1}^{T} r_{t+k} - r_t = a + b\,(P/D)_t + \varepsilon_{t+T}$ ($T = 20$ quarters), based on a 1,200-year monthly simulation of the calibrated baseline economy. Column (1): full sample. Column (2): "bad states" with $(P/D)_t$ below the 5th percentile. Standard errors in parentheses.

flows. Only economies with sufficiently inelastic markets jointly match return volatility, price impact, and return predictability, so market inelasticity is a central ingredient for matching standard macro-finance moments.

5.6 Conditional price impact and dynamics

Averaging across states masks the state dependence of aggregate elasticity emphasized in Section 4. To trace this dependence, we compute the model-implied *conditional* price impact, plotted in Figure 9. We fix passive demand $\overline{\alpha}_{p,t}$ and the dividend share s_t at their unconditional means, vary one wealth-share coordinate at a time, and apply a fixed one-quarter passive-flow shock. The shock is propagated through the dynamic transition of the state variables. We then recompute $\log(P/D)$ at the shocked and unshocked states, and measure the conditional price impact as the resulting change in $\log(P/D)$ per unit of passive-flow shock.

The left panel of Figure 9 shows that the price impact rises sharply as the wealth share of the unconstrained active sector $x_{u,t}$ falls. When the sector most capable of absorbing risk is small, the market is less elastic and flow shocks translate into larger price movements. The right panel shows that the price impact also rises monotonically in the wealth share of the constrained active sector $x_{c,t}$. As more of the active sector hits its leverage limit, the residual risk-bearing capacity in $x_{u,t}$ shrinks and price impact grows. These patterns mirror the analytical characterization in Proposition 6 and explain why the conditional price impact reported in Table 3 (the row labeled "Cond. price impact (D)" with a value of 5.41) is larger than its unconditional counterpart.

Combined with the counterfactual results, Figure 9 makes clear that the very states in which risk-bearing capacity is depleted, with low x_u or high x_c, are the states in which discount rates are most volatile and the P/D ratio is most informative about future returns.

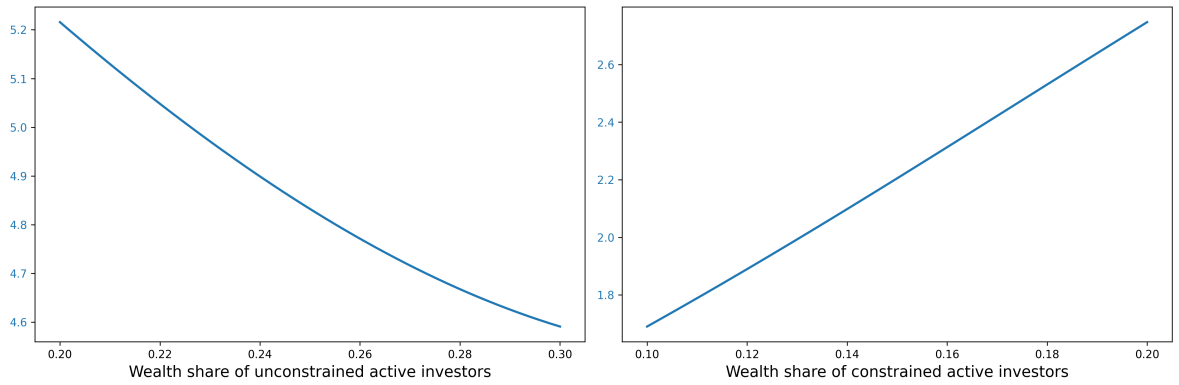


Figure 9. Conditional price impact.
Model-implied price impact of the dividend claim evaluated off the ergodic mean. In each panel, $\overline{\alpha}_{p,t}$ and s_t are fixed at their unconditional means and one wealth-share coordinate is varied. Conditional price impact is computed as the change in $\log(P/D)$ per unit of passive-flow shock, where the shock is propagated one quarter forward through the state-variable dynamics.

5.7 Takeaways

Three central results emerge from the quantitative analysis. First, when calibrated to sectoral FoF holdings and passive equity flows, the model jointly matches aggregate price impact, the volatility of dividend-claim returns, and the risk premium. Second, heterogeneity and risk misallocation are essential for matching the asset pricing moments. Removing either preference heterogeneity or risk misallocation reduces price impact and return volatility. Third, leverage constraints further amplify these effects by tightening risk-sharing when risk-bearing capacity is low. Finally, the model delivers a demand-based explanation of the excess volatility puzzle and return predictability, based on how portfolio flows interact with inelastic markets.

6 Conclusion

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

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

Macro elasticity is both state-dependent and time-varying. Passive investors raise price impact, leverage constraints further tighten risk-bearing capacity, and preference heterogeneity amplifies

price impact in bad times by making risk misallocation state dependent. Inelasticity alone does not generate excess volatility, nor do flows in infinitely elastic markets; both elements are needed to explain observed fluctuations and the countercyclical volatility multiplier.

Our analytical results rely on a state-global perturbation method that provides closed-form characterizations of elasticity, while the quantitative analysis solves the full dynamic model with deep neural networks. Calibrated to U.S. Flow of Funds sectoral holdings and to the volatility and persistence of quarterly equity flows, the model jointly matches the empirical price impact of equity flows, standard asset pricing moments, and sectoral balance sheets, while the endowment claim remains an order of magnitude more elastic than the equity claim, confirming the analytical distinction between the two. The calibration also reproduces the countercyclical behavior of price impact and the dominant role of discount rate variation in return predictability. Counterfactuals trace these results to the interaction of frictions with portfolio flows: the same frictions that make markets inelastic also generate the equity premium and the level of return volatility, so the inelastic markets hypothesis and standard asset pricing puzzles are addressed by a single mechanism.

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

References

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

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

Alvarez, Fernando, and Andy Atkeson, 2018, The risk of becoming risk averse: A model of asset pricing and trade volumes, Working Paper.

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.

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

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

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

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

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

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.

Campbell, John Y., and Robert J. Shiller, 1988, The dividend–price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 1, 195–228.

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.

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

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

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

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.

Cochrane, John H., 2011, Presidential address: Discount rates, *Journal of Finance* 66, 1047–1108.

Cochrane, John H., Francis A. Longstaff, and Pedro Santa-Clara, 2008, Two trees, *Review of Financial Studies* 21, 347–385.

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.

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.

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.

Fernández-Villaverde, Jesús, 2026, Deep learning for solving economic models, *Journal of Economic Literature* Forthcoming.

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.

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

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

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

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

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

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

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

Lo, Andrew W., and Jiang Wang, 2000, Trading volume: Definitions, data analysis, and implications of portfolio theory, *Review of Financial Studies* 13, 257–300.

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

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

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

Parker, Jonathan A, Antoinette Schoar, Allison 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, 2020, The risk channel of unconventional monetary policy, Working Paper.

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.

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 partial equilibrium 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

$$0 = \max_{C_j, \alpha_j} f_j(C_j, V_j) + V_{j,W}\left[rW_j + \pi^\top \alpha_j W_j - C_j\right] + V_{j,X}\mu_X + \frac{1}{2}V_{j,WW}W_j^2\|\sigma_j\|^2 + V_{j,WX}W_j\sigma_X\sigma_j^\top + \frac{1}{2}\sum_{k=1}^{3}\sigma_{X,k}^\top V_{j,XX}\sigma_{X,k},$$

subject to

$$\alpha_j = \overline{\alpha}_p[\alpha_{D,t}^m, \alpha_{B,t}^m, 0]^\top \qquad \text{for } j = 0, \qquad\qquad \sigma_j\sigma_j^\top \leq \overline{\sigma}^2 \qquad \text{for } j = 2.$$

given $\sigma_j = \alpha_j^\top \Sigma$, 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 a $N \times 1$ vector, the diffusion σ_X is a $N \times 3$ matrix, while σ_R is a 1×3 vector. The notation $\sigma_{X,k}$ denotes the $k-$th column of σ_X, that is, the exposure to the $k-$th Brownian motion.

The optimality condition for consumption and portfolio share are given by

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

where the second optimality condition holds for $j = 1, 2$, 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 share for an active investor is given by

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

using 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

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

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

Let $p_{D,t} = \frac{P_{D,t}}{Y_t}$ and $p_{B,t} = \frac{P_{B,t}}{Y_t}$ denote the normalized prices of the dividend and bond claims, respectively. The normalized price of the aggregate claim is $p_{Y,t} = \frac{P_{Y,t}}{Y_t}$. The pricing condition for asset $k \in \{D, B\}$ is given by

$$r_t + \sigma_{R_k,t}\eta_t^\top = \frac{1}{p_{k,t}} + \mu + \mu_{p_k,t} + \sigma\sigma_{p_k,t}^\top,$$

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

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

The expected return and risk exposure on the aggregate claim are given by $\mu_{R_Y,t} = \alpha_{D,t}^m \mu_{R_D,t} + \alpha_{B,t}^m \mu_{R_B,t}$ and $\sigma_{R_Y,t} = \alpha_{D,t}^m \sigma_{R_D,t} + \alpha_{B,t}^m \sigma_{R_B,t}$, respectively.

A.2.3 Aggregate state variable

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

$$x_{j,t} \equiv \frac{\omega_j W_{j,t}}{\sum_{i=0}^{2}\omega_i W_{i,t}}.$$

We define the aggregate state variable as $X_t = (x_{1,t}, x_{2,t}, \overline{\alpha}_{p,t}, s_t)$.

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

$$\sum_{j=0}^{2}\omega_j W_{j,t}\alpha_{j,k,t} = P_{k,t}, \qquad \sum_{j=0}^{2}\omega_j W_{j,t}(1 - \alpha_{j,D,t} - \alpha_{j,B,t}) = 0,$$

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

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

$$\sum_{j=0}^{2}\omega_j W_{j,t} = P_{Y,t}.$$

The market clearing condition for goods is given by

$$\sum_{j=0}^{2} x_{j,t} c_{j,t} = \frac{1}{p_{Y,t}}.$$

We can then write the market clearing condition for the equity claim, corporate bond claim, and the derivative as follows:

$$\sum_{j=0}^{2} x_{j,t} \alpha_{j,D,t} = \alpha_{D,t}^{m}, \qquad \sum_{j=0}^{2} x_{j,t} \alpha_{j,B,t} = \alpha_{B,t}^{m}, \qquad \sum_{j=0}^{2} x_{j,t} \alpha_{j,F,t} = 0.$$

Let $\alpha_t^m = [\alpha_{D,t}^m, \alpha_{B,t}^m, 0]^\top$ denote the market weights of the dividend, bond, and derivative claims. We can then write the market clearing conditions above in a compact way:

$$\sum_{j=0}^{2} x_{j,t} \alpha_{j,t} = \alpha_t^m.$$

Multiplying both sides by Σ and using the fact that $\sigma_{R_Y,t} = \omega_{D,t}\sigma_{R_D,t} + \omega_{B,t}\sigma_{R_B,t}$, we obtain

$$\sum_{j=0}^{2} x_{j,t} \sigma_{j,t} = \sigma_{R_Y,t}.$$

Law of motion of the aggregate state variable. The law of motion of $\overline{\alpha}_{p,t}$ is given by (11). The law of motion of s_t is given by (9). 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

$$\frac{dx_{j,t}}{x_{j,t}} = \left[r_t + \pi_t^\top \alpha_{j,t} - c_{j,t} - \mu_{P_Y,t} - (\sigma_{j,t} - \sigma_{R_Y,t})\, \sigma_{R_Y,t}^\top + \kappa \frac{\omega_j - x_{j,t}}{x_{j,t}} \right] dt + (\sigma_{j,t} - \sigma_{R_Y,t})\, dZ_t.$$

using $\mu_{P_Y,t} = \mu_{R_Y,t} - \mu_{R_D,t}\alpha_{D,t}^m - \mu_{R_B,t}\alpha_{B,t}^m$.

A.2.4 Block recursivity and the mutual fund theorem

To compute the equilibrium, one needs to solve a system of 5 partial differential equations (PDEs), involving the three consumption-wealth ratios $\{c_{j,t}(X_t)\}_{j=0}^2$ and the two price-dividend

ratios $\{p_{k,t}(X_t)\}_{k=D,B}$. These functions depend on 4 state variables, the 2−dimensional vector x_t, the portfolio share of passive investors $\overline{\alpha}_{p,t}$, and the dividend share s_t.

The system of PDEs satisfies an important *block-recursive* structure. We can first solve for a reduced system of PDEs involving (c_j, p_Y) in the set of states $\hat{X}_t = (x_{1,t}, x_{2,t}, \overline{\alpha}_{p,t})$. Given the solution of the reduced system of PDEs, we can then solve for $p_{D,t}$ and $p_{B,t}$ in the full set of states $X_t = (x_{1,t}, x_{2,t}, \overline{\alpha}_{p,t}, s_t)$. Moreover, the solution satisfies the mutual fund theorem, that is, investors hold risky claims in proportion to their market-capitalization weights. This result is summarized in Proposition 2.

We provide below the proof of Proposition 2 in two steps. First, we derive the system of PDEs and show the block-recursive structure in Lemma A.1. Second, we show the mutual fund theorem holds in Lemma A.2.

Block recursivity. We next derive the system of PDEs. We will show that one can solve for a reduced system of PDEs involving (c_j, p_Y). This is the crucial in the obtaining the mutual fund theorem in our setting.

Lemma A.1. *The consumption-wealth ratios, $\{c_{j,t}(X_t)\}_{j=0}^2$, 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, $\sigma_{x_1,t}$ and $\sigma_{x_2,t}$, 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, $\mu_{x_1,t}$ and $\mu_{x_2,t}$, 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:

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

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

$$\sigma_{R_Y} = \chi_{\sigma_{R_Y},0} + \chi_{\sigma_{R_Y},1}\sigma_{x_1} + \chi_{\sigma_{R_Y},2}\sigma_{x_2},$$

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

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

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

$$\sigma_{c_j} = \chi_{\sigma_{c_j},0} + \chi_{\sigma_{c_j},1}\sigma_{x_1} + \chi_{\sigma_{c_j},2}\sigma_{x_2},$$

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

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

The market price of risk

$$\eta = \gamma_a \left[\chi_a \sigma_{R_Y} - \sum_{j=0}^{2} x_j \frac{1 - \gamma_j^{-1}}{\psi - 1} \frac{\sigma_{c_j}}{1 + \nu_j} \right],$$

where $\chi_a = \frac{1-(1-x_1-x_2)\overline{\alpha}_p}{x_1+x_2}$ is a known function of the aggregate state variables X.

We can then write η as follows:

$$\eta = \chi_{\eta,0} + \chi_{\eta,1}\sigma_{x_1} + \chi_{\eta,2}\sigma_{x_2},$$

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

$$\chi_{\eta,k} = \gamma_a \left[\chi_a \chi_{\sigma_{R_Y},k} - \sum_{j=0}^{2} x_j \frac{1 - \gamma_j^{-1}}{\psi - 1} \frac{\chi_{\sigma_{c_j},k}}{1 + \nu_j} \right].$$

The risk exposure of investor $j = 0, 1, 2$ is given by

$$\begin{aligned} \sigma_0 &= \overline{\alpha}_p \sigma_{R_Y} \\ \sigma_j &= \frac{1}{1 + \nu_j} \left[\frac{\eta}{\gamma_j} + \frac{1 - \gamma_j^{-1}}{\psi - 1} \sigma_{c_j} \right], \qquad j = 1, 2. \end{aligned}$$

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

$$\sigma_j = \chi_{\sigma_j,0} + \chi_{\sigma_j,1} \sigma_{x_1} + \chi_{\sigma_j,2} \sigma_{x_2},$$

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

$$\begin{aligned} \chi_{\sigma_0,k} &= \overline{\alpha}_p \chi_{\sigma_{R_Y},k} \\ \chi_{\sigma_j,k} &= \frac{1}{1 + \nu_j} \left[\frac{\chi_{\eta,k}}{\gamma_j} + \frac{1 - \gamma_j^{-1}}{\psi - 1} \chi_{\sigma_{c_j},k} \right], \qquad j = 1, 2. \end{aligned}$$

The risk exposure of the endogenous state variables is given by

$$\sigma_{x_j} = x_j \left[\sigma_j - \sum_{k=0}^{2} x_k \sigma_k \right],$$

using the fact that $\sigma_{R_Y} = \sum_{k=0}^{2} x_k \sigma_k$.

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

$$\begin{bmatrix} 1 - x_1 \sum_{k=0}^{2} \left(\chi_{\sigma_1,1} - \chi_{\sigma_k,1} \right) x_k & -x_1 \sum_{k=0}^{2} \left(\chi_{\sigma_1,2} - \chi_{\sigma_k,2} \right) x_k \\ -x_2 \sum_{k=0}^{2} \left(\chi_{\sigma_2,1} - \chi_{\sigma_k,1} \right) x_k & 1 - x_2 \sum_{k=0}^{2} \left(\chi_{\sigma_2,2} - \chi_{\sigma_k,2} \right) x_k \end{bmatrix} \begin{bmatrix} \sigma_{x_1} \\ \sigma_{x_2} \end{bmatrix} = \begin{bmatrix} x_1 \sum_{k=0}^{2} \left(\chi_{\sigma_1,0} - \chi_{\sigma_k,0} \right) x_k \\ x_2 \sum_{k=0}^{2} \left(\chi_{\sigma_2,0} - \chi_{\sigma_k,0} \right) x_k \end{bmatrix}.$$

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:

$$\mu_{c_j} = \frac{c_{j,X}}{c_j}\mu_X + \frac{1}{2}\sum_{k=1}^{3} \sigma_{X,k}^\top \frac{c_{j,XX}}{c_j}\sigma_{X,k},$$
$$\mu_{p_Y} = \frac{p_{Y,X}}{p_Y}\mu_X + \frac{1}{2}\sum_{k=1}^{3} \sigma_{X,k}^\top \frac{p_{Y,XX}}{p_Y}\sigma_{X,k}.$$

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 $\xi = \sum_{j=0}^{2} x_j \xi_j$ and the interest rate:

$$r = \rho + \psi^{-1}\left(\mu_{p_Y} + \mu + \sigma\sigma_{p_Y}^\top + \xi\right) + (1-\psi^{-1})\sum_{j=0}^{2} x_j \frac{\gamma_j}{2}\|\sigma_j\|^2 - \pi_Y.$$

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). □

A corollary of the above proposition is that the consumption-wealth ratio and price-dividend ratio are independent of the dividend share s.

Corollary 1. *The consumption-wealth ratio c_j and price-dividend ratio p_Y are a function of the reduced state vector $\hat{X} = (x_1, x_2, \overline{\alpha}_p)$.*

Proof. The dividend share s does not appear in the expressions for PDE system for (c_j, p_Y). They only appear indirectly through μ_s and σ_s. If (c_j, p_Y) are a function of $\hat{X}$, then their derivatives with respect to s are zero, and s completely drops out of the PDE system for (c_j, p_Y). □

The mutual fund theorem We next show that a version of the mutual fund theorem holds in our economy.

Lemma A.2 (Mutual fund theorem)**.** *The portfolio shares of active investors are given by:*

$$\alpha_{j,k} = \alpha_{j,Y} \frac{p_k}{p_Y}, \quad \forall k \in \{D, B\},$$

where $\alpha_{j,Y}$ denotes the holdings of the market index fund for investor j, which satisfies the condition

$$\alpha_{j,Y} = \frac{\sigma_j \sigma_{R_Y}^\top}{\sigma_{R_Y} \sigma_{R_Y}^\top}.$$

Moreover, the holdings of the different risky claims are equalized for each investor: $Q_{j,D,t} = Q_{j,B,t}$.

Proof. We can implement the risk exposure $\sigma_{j,t}$ by holding the market index fund and a derivative. Let $\hat{\sigma}_{j,t} = [\sigma_{j,t,1}, \sigma_{j,t,2}]$, $\hat{\Sigma}_t = \begin{bmatrix} \hat{\sigma}_{R_Y,t} \\ \hat{\sigma}_{R_F,t} \end{bmatrix}$, and $\hat{\alpha}_{j,t} = [\alpha_{j,Y,t}, \alpha_{j,F,t}]^\top$, then we can define the portfolio shares of the market index fund and the derivative as:

$$\hat{\sigma}_{j,t} = \hat{\alpha}_{j,t}^\top \hat{\Sigma}_t \Rightarrow \hat{\alpha}_{j,t}^\top = \hat{\sigma}_{j,t} \hat{\Sigma}_t^{-1}.$$

We can then write the risk exposure $\sigma_{j,t}$ as:

$$\sigma_{j,t} = \alpha_{j,Y,t} \sigma_{R_Y,t} + \alpha_{j,F,t} \sigma_{R_F,t}.$$

Using the fact that $\sigma_{R_Y,t} = \frac{P_{D,t}}{P_{Y,t}} \sigma_{R_D,t} + \frac{P_{B,t}}{P_{Y,t}} \sigma_{R_B,t}$, we obtain:

$$\sigma_{j,t} = \alpha_{j,Y,t} \frac{P_{D,t}}{P_{Y,t}} \sigma_{R_D,t} + \alpha_{j,Y,t} \frac{P_{B,t}}{P_{Y,t}} \sigma_{R_B,t} + \alpha_{j,F,t} \left(\frac{P_{B,t}}{P_{Y,t}} \sigma_{R_Y,t} + \frac{P_{D,t}}{P_{Y,t}} \sigma_{R_D,t} \right).$$

Hence, this implies that the portfolio shares of dividend and bond claims are given by:

$$\alpha_{j,D,t} = \alpha_{j,Y,t} \frac{P_{D,t}}{P_{Y,t}}, \quad \alpha_{j,B,t} = \alpha_{j,Y,t} \frac{P_{B,t}}{P_{Y,t}},$$

where $\alpha_{j,Y,t}$ is the portfolio share of the market index fund defined.

Multiplying the equation for $\sigma_{j,t}$ by $\sigma_{R_Y,t}^\top$, we obtain:

$$\sigma_{j,t}\sigma_{R_Y,t}^\top = \alpha_{j,Y,t}\sigma_{R_Y,t}\sigma_{R_Y,t}^\top \Rightarrow \alpha_{j,Y,t} = \frac{\sigma_{j,t}\sigma_{R_Y,t}^\top}{\sigma_{R_Y,t}\sigma_{R_Y,t}^\top},$$

using the orthogonality condition $\sigma_{R_F,t}\sigma_{R_Y,t}^\top = 0$.

Similarly, multiplying the equation for $\sigma_{j,t}$ by $\sigma_{R_F,t}^\top$, we obtain:

$$\alpha_{j,F,t} = \frac{\sigma_{j,t}\sigma_{R_F,t}^\top}{\sigma_{R_F,t}\sigma_{R_F,t}^\top} = \frac{\sigma_{j,t}\sigma_{R_F,t}^\top}{\|\sigma_{R_F,t}\|^2}.$$

The holdings of the different risky claims are given by:

$$Q_{j,D,t} = \alpha_{j,D,t}\frac{W_{j,t}}{P_{D,t}} = \alpha_{j,Y,t}\frac{W_{j,t}}{P_{Y,t}}, \qquad Q_{j,B,t} = \alpha_{j,B,t}\frac{W_{j,t}}{P_{B,t}} = \alpha_{j,Y,t}\frac{W_{j,t}}{P_{Y,t}}.$$

Hence, the holdings of the dividend claim and the bond claim are equalized for each investor: $Q_{j,D,t} = Q_{j,B,t}$.

Pricing the individual claims. The normalized price of the dividend claim satisfies the pricing condition:

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

where

$$\mu_{p_D,t} = \frac{p_{D,X,t}}{p_{D,t}}\mu_{X,t} + \frac{1}{2}\sum_{k=1}^{3}\sigma_{X,k,t}^\top\frac{p_{D,XX,t}}{p_{D,t}}\sigma_{X,k,t}, \qquad \sigma_{p_D,t} = \frac{p_{D,X,t}}{p_{D,t}}\sigma_{X,t}.$$

□

A.2.5 Closed-form solution to individual claims

I consider next the characterization of the price of the individual claims when the interest rate and the price of risk are constant. This corresponds to the case in the economy without preference heterogeneity and passive investors are fully invested in the risky asset, $\overline{\alpha}_{p,t} = 1$. To obtain closed-form solutions, analogous to the ones in Menzly et al. (2004), we focus in the special case $\gamma = 1$. The closed-form solutions provide a useful reference to compare with the perturbation solution derived in the next section.

Pricing the individual claims. The price of the dividend claim satisfies the pricing condition:

$$P_{D,t} = \mathbb{E}_t \left[\int_t^\infty \frac{\Lambda_\tau}{\Lambda_t} s_\tau Y_\tau d\tau \right],$$

The SDF and the aggregate endowment follow the processes:

$$\frac{d\Lambda_t}{\Lambda_t} = -r^* dt - \eta^* dZ_t, \qquad \frac{dY_t}{Y_t} = \mu dt + \sigma dZ_t.$$

We can solve for the price of the dividend claim in closed form in the case $\gamma = 1$, so $\eta^* = \sigma$. In this case, we can then write the SDF and the aggregate endowment as:

$$\Lambda_t = \exp\left(-\int_0^t \left(r^* + \frac{\sigma\sigma^\top}{2} \right) ds - \sigma Z_t \right) \Lambda_0, \qquad Y_t = \exp\left(\int_0^t \left(\mu - \frac{\sigma\sigma^\top}{2} \right) ds + \sigma Z_t \right) Y_0.$$

Combining the expressions for Λ_t and Y_t, we obtain:

$$P_{D,t} = \int_t^\infty e^{-\delta(\tau - t)} \mathbb{E}_t \left[s_\tau \right] d\tau Y_t,$$

where $\delta = r^* + \sigma\sigma^\top - \mu$, using the fact that Z_t cancels out in the product $\Lambda_t Y_t$.

The conditional expectation of the dividend share is given by

$$\mathbb{E}_t \left[s_\tau \right] = \overline{s} + (s_t - \overline{s}) e^{-\kappa_s (\tau - t)}.$$

The price of the dividend claim is given by

$$P_{D,t} = \left[\frac{\overline{s}}{\delta} + \frac{s_t - \overline{s}}{\delta + \kappa_s} \right] Y_t$$

where $Y_t = \exp\left(\int_0^t \left(\mu - \frac{\sigma\sigma^\top}{2} \right) ds + \sigma Z_t \right) Y_0$.

Hence, the normalized price of the dividend and corporate bond claims are given by

$$p_{D,0}(X) = \frac{1}{\delta} \left[\frac{\delta}{\delta + \kappa_s} s_t + \frac{\kappa_s}{\delta + \kappa_s} \overline{s} \right], \qquad p_{B,0}(X) = \frac{1}{\delta} \left[\frac{\delta}{\delta + \kappa_s} (1 - s_t) + \frac{\kappa_s}{\delta + \kappa_s} (1 - \overline{s}) \right].$$

Portfolio share of individual claims. The portfolio share of investor j on the dividend and bond claims are given by

$$\alpha_{j,D,t} = \frac{p_{D,t}}{p_{Y,t}}, \qquad \alpha_{j,B,t} = \frac{p_{B,t}}{p_{Y,t}}.$$

In the benchmark economy with $\gamma = 1$, we have that

$$\alpha_{j,D,0}(X) = \frac{\delta}{\delta + \kappa_s} s + \frac{\kappa_s}{\delta + \kappa_s} \overline{s}, \qquad \alpha_{j,B,0}(X) = \frac{\delta}{\delta + \kappa_s} (1 - s) + \frac{\kappa_s}{\delta + \kappa_s} (1 - \overline{s}),$$

using the fact that $\delta = 1/p_{Y,0}(X)$ when $\gamma = 1$.

A.3 Derivations for Section 4: Perturbation Solution

In this section, we discuss the derivations of the perturbation solution in detail. Consistent with the block-recursivity property in Proposition 2, the consumption-wealth ratios and the price-dividend ratio of the aggregate claim are independent of the dividend share s.

For these variables, we can focus on the reduced set of states $\hat{X} = (x, \overline{\alpha}_p)$, where $x = (x_1, x_2)$ is the vector of wealth shares of the active investors. To price the individual claims, we need to consider the full set of states $X = (x, \overline{\alpha}_p, s)$.

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 $\sum_{j=0}^{2} \omega_j \hat{\gamma}_j = 0$.

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

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

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

Second, the dividend share is written as $s = \overline{s} + \hat{s}\epsilon$, where $\hat{s}$ follows the process induced by

$$ds = \kappa_s(\overline{s} - s)dt + s(1 - s)\sigma_s \epsilon dZ_t.$$

The case $\epsilon = 0$ corresponds to the benchmark economy, in which there is no preference heterogeneity and the exogenous state variables are constant. In particular, passive investors are fully invested in the risky asset and the dividend share is fixed at $s = \overline{s}$.

We expand the consumption-wealth ratio of investor $j \in \{0, 1, 2\}$ and the price-dividend ratio of the aggregate claim as follows:

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

For each individual claim $k \in \{D, B\}$, we use the expansion

$$p_k(x, \overline{\alpha}_p, s; \epsilon) = p_{k,0}(x, \hat{\alpha}_p, \hat{s}) + p_{k,1}(x, \hat{\alpha}_p, \hat{s})\epsilon + p_{k,2}(x, \hat{\alpha}_p, \hat{s})\epsilon^2 + O(\epsilon^3).$$

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

Lemma A.3 (Basic properties of the perturbation solution)**.** *The perturbation solution satisfies the following properties:*

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

$$c_{j,0}(x, \hat{\alpha}_p) = c_j^*, \qquad p_{Y,0}(x, \hat{\alpha}_p) = p_Y^*, \qquad p_{k,0}(x, \hat{\alpha}_p, \hat{s}) = p_k^*, \qquad k \in \{D, B\},$$

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

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:*

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

   where $\psi_{j,l}(x)$ and $\Psi_{j,l}(x)$, for $j \in \{0, 1, 2, Y\}$ and $l \in \{0, \hat{\alpha}_p, \hat{\alpha}_p^2\}$, are functions of x.

4. *The first-order and second-order corrections to the price-dividend ratio for the individual claim can be written as:*

$$\begin{aligned} p_{k,1}(x, \hat{\alpha}_p, \hat{s}) &= \psi_{k,0}(x) + \psi_{k,\hat{\alpha}_p}(x)\hat{\alpha}_p + \psi_{k,\hat{s}}(x)\hat{s}, \\ p_{k,2}(x, \hat{\alpha}_p, \hat{s}) &= \Psi_{k,0}(x) + \Psi_{k,\hat{\alpha}_p}(x)\hat{\alpha}_p + \Psi_{k,\hat{s}}(x)\hat{s} + \Psi_{k,\hat{\alpha}_p^2}(x)\hat{\alpha}_p^2 + \Psi_{k,\hat{\alpha}_p\hat{s}}(x)\hat{\alpha}_p\hat{s} + \Psi_{k,\hat{s}^2}(x)\hat{s}^2, \end{aligned}$$

   where $\psi_{k,l}(x)$ and $\Psi_{k,l}(x)$, for $k \in \{D, B\}$ and $l \in \{0, \hat{\alpha}_p, \hat{\alpha}_p^2, \hat{s}, \hat{\alpha}_p\hat{s}, \hat{s}^2\}$, 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 $(x, \hat{\alpha}_p, \hat{s})$. 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. Properties 3 and 4 describe a solution that is global in the endogenous wealth shares and local in the exogenous state variables.

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 price-dividend ratio of the aggregate claim. Finally, we derive the corresponding corrections to the price-dividend ratios of the individual claims.

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 is slack. 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 $\sigma_{c_j,0}(X) = \sigma_{p_Y,0}(X) = 0_{1\times 3}$, where $0_{1\times 3}$ is a 1×3 vector of zeros.

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 $\sigma_{p_Y,t} = 0_{1\times 3}$ 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

$$\mu_{x_j,0}(X) = 0, \qquad \sigma_{x_j,0}(X) = 0_{1\times 3}, \qquad \forall j \in \{1,2\}.$$

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

Derivative claim. The risk exposure of the derivative claim is given by

$$\sigma_{R_F,0} = [\sigma^1_{R_F,0}, \sigma^2_{R_F,0}, \sigma^3_{R_F,0}].$$

These three quantities are defined by the normalization conditions

$$\sigma_{R_Y,0}\sigma^\top_{R_F,0} = 0, \qquad \sigma_{R_F,0}\sigma^\top_{R_F,0} = \sigma_{R_Y,0}\sigma^\top_{R_Y,0}, \qquad \sigma^3_{R_F,0} = 0.$$

Using $\sigma_{R_Y,0} = [\sigma^1, 0, 0]$, we obtain

$$\sigma^1_{R_F,0} = 0, \qquad \sigma^2_{R_F,0} = \sigma^1, \qquad \sigma^3_{R_F,0} = 0.$$

Portfolio shares. The portfolio shares on the aggregate claim and the derivative satisfy

$$\sigma_{j,0} = \alpha_{j,Y}\sigma_{R_Y,0} + \alpha_{j,F}\sigma_{R_F,0}.$$

Using $\sigma_{j,0} = \sigma_{R_Y,0}$, we obtain $\alpha_{j,Y} = 1$ and $\alpha_{j,F} = 0$ for $j = 0, 1, 2$.

□

A.3.2 Proof of Proposition 3

Proof. Consistent with Proposition 2, the consumption-wealth ratios and the price-dividend ratio of the aggregate claim are independent of the dividend-share shock s. We can then focus on the reduced set of states $\hat{X} = (x, \overline{\alpha}_p)$, where $x = (x_1, x_2)$ is the vector of wealth shares of the 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:

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

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

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 $j = 1, 2$. Hence, the drift and diffusion of the state variables are first-order in ϵ, that is, $\mu_{\hat{X}} = O(\epsilon)$ and $\sigma_{\hat{X}} = O(\epsilon)$.

We characterize next the derivatives of the consumption-wealth ratios and the price-dividend ratio with respect to $\hat{X}$. 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:

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

This implies that $\sigma_{R_Y,1} = 0_{1\times 3}$.

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

$$\sigma_{0,1} = \hat{\alpha}_p \sigma.$$

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

$$\sum_{j=0}^{2} x_j \sigma_{j,1} = 0$$

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

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

$$\eta_1 = \gamma\sigma \left[\sum_{j=1}^{2} \frac{x_j}{x_a}(\hat{\gamma}_j + \hat{\nu}_j) - \frac{x_0}{x_a}\hat{\alpha}_p \right],$$

where $x_a = 1 - x_0$.

The risk exposure for the first active investor is

$$\sigma_{1,1} = -\left[\frac{x_2}{x_a}\left(\hat{\gamma}_1 - \hat{\gamma}_2 - \hat{\nu}_2\right) + \frac{x_0}{x_a}\hat{\alpha}_p\right]\sigma.$$

and the risk exposure for the second active investor is

$$\sigma_{2,1} = \left[\frac{x_1}{x_a}\left(\hat{\gamma}_1 - \hat{\gamma}_2 - \hat{\nu}_2\right) - \frac{x_0}{x_a}\hat{\alpha}_p\right]\sigma.$$

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 1 is unconstrained and investor 2 may be constrained, so $\hat{\nu}_1 = 0$ and $\hat{\nu}_2 \geq 0$. For investor 2, using the expression for $\sigma_{2,1}$ above, the first-order leverage constraint is

$$\frac{x_1}{x_a}\left(\hat{\gamma}_1 - \hat{\gamma}_2 - \hat{\nu}_2\right) - \frac{x_0}{x_a}\hat{\alpha}_p \leq \hat{\sigma}.$$

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

$$\hat{\nu}_2 = \hat{\gamma}_1 - \hat{\gamma}_2 - \frac{x_0}{x_1}\hat{\alpha}_p - \frac{x_a}{x_1}\hat{\sigma}.$$

If the constraint is slack, then $\hat{\nu}_2 = 0$ and

$$\frac{x_1}{x_a}\left(\hat{\gamma}_1 - \hat{\gamma}_2\right) - \frac{x_0}{x_a}\hat{\alpha}_p < \hat{\sigma},$$

or, equivalently,

$$\hat{\gamma}_1 - \hat{\gamma}_2 - \frac{x_0}{x_1}\hat{\alpha}_p - \frac{x_a}{x_1}\hat{\sigma} < 0.$$

Hence, the first-order Lagrange multiplier is

$$\hat{\nu}_2 = \max\left\{\hat{\gamma}_1 - \hat{\gamma}_2 - \frac{x_0}{x_1}\hat{\alpha}_p - \frac{x_a}{x_1}\hat{\sigma}, 0\right\}.$$

When the constraint binds, substituting the expression for $\hat{\nu}_2$ into the market price of risk yields

$$\eta_1 = \gamma\sigma\left[\hat{\gamma}_1 - \frac{x_2}{x_1}\hat{\sigma} - \frac{x_0}{x_1}\hat{\alpha}_p\right].$$

When the constraint is slack, $\hat{\nu}_2 = 0$, and the market price of risk is

$$\eta_1 = \gamma\sigma\left[\sum_{j=1}^{2}\frac{x_j}{x_a}\hat{\gamma}_j - \frac{x_0}{x_a}\hat{\alpha}_p\right].$$

Equivalently, both cases can be summarized as

$$\eta_1 = \gamma\sigma\left[\sum_{j=1}^{2}\frac{x_j}{x_a}(\hat{\gamma}_j + \hat{\nu}_j) - \frac{x_0}{x_a}\hat{\alpha}_p\right],$$

with $\hat{\nu}_1 = 0$ and $\hat{\nu}_2$ given above.

Risk premium. The first-order risk premium is

$$\pi_{Y,1} = \eta_1\sigma^\top = \gamma\|\sigma\|^2\left[\sum_{j=1}^{2}\frac{x_j}{x_a}(\hat{\gamma}_j + \hat{\nu}_j) - \frac{x_0}{x_a}\hat{\alpha}_{p,1}\right].$$

It increases with the average risk aversion of active investors, with tighter (binding) leverage constraints, and when active investors bear a larger share of aggregate risk.

The derivative of the risk premium with respect to the portfolio flow is given by

$$\frac{\partial\hat{\pi}_{Y,t}}{\partial f_t} = \frac{1}{x_{0,t}}\frac{\partial\pi_{Y,1}}{\partial\hat{\alpha}_p} = -\frac{\gamma\|\sigma\|^2}{x_{u,t}} = -\frac{1}{\zeta_{r,t}},$$

using the fact that $\zeta_{r,t} = \frac{x_{u,t}}{\gamma\|\sigma\|^2}$, and $x_{u,t} = x_{a,t}$ when the leverage constraint is slack and $x_{u,t} = x_{1,t}$ when the constraint is binding.

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

$$\mu_{c_j} = \frac{c_{j,\hat{X}}}{c_j}\mu_{\hat{X}} + \frac{1}{2}\sum_{k=1}^{3}\sigma_{\hat{X},k}^{\top}\frac{c_{j,\hat{X}\hat{X}}}{c_j}\sigma_{\hat{X},k},$$
$$\mu_{p_Y} = \frac{p_{Y,\hat{X}}}{p_Y}\mu_{\hat{X}} + \frac{1}{2}\sum_{k=1}^{3}\sigma_{\hat{X},k}^{\top}\frac{p_{Y,\hat{X}\hat{X}}}{p_Y}\sigma_{\hat{X},k},$$

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

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

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

$$r_1 = (1-\psi^{-1})\frac{\gamma\|\sigma\|^2}{2}\sum_{j=0}^{2}x_j\hat{\gamma}_j + (1-\psi^{-1})\gamma\|\sigma\|^2\sum_{j=0}^{2}x_j\frac{\sigma_{j,1}\sigma^{\top}}{\|\sigma\|^2} - \pi_{Y,1},$$

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 $\sum_{j=0}^{2}x_j\sigma_{j,1} = 0_{1\times 3}$, we obtain

$$r_1 = -\gamma\|\sigma\|^2\left[\sum_{j=1}^{2}\frac{x_j}{x_a}(\hat{\gamma}_j + \hat{\nu}_j) - \sum_{j=0}^{2}x_j\hat{\gamma}_j - \frac{x_0}{x_a}\hat{\alpha}_{p,1} + \frac{1+\psi^{-1}}{2}\sum_{j=0}^{2}x_j\hat{\gamma}_j\right].$$

Notice that we can write the expression above as

$$r_1 = -\gamma\|\sigma\|^2\frac{1+\psi^{-1}}{2}\sum_{j=0}^{2}x_j\hat{\gamma}_j - \left[\pi_{Y,1} - \pi_{Y,1}^{fb}\right],$$

where

$$\pi_{Y,1}^{fb} = \gamma\|\sigma\|^2\sum_{j=0}^{2}x_j\hat{\gamma}_j.$$

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

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

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

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

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

$$c_{j,1} = \frac{\gamma\|\sigma\|^2}{2}(1-\psi)\left[\left(1-\psi^{-1}\right)\sum_{k=0}^{2}x_k\hat{\gamma}_k - \hat{\gamma}_j\right],$$

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

Price-dividend ratio of 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

$$\frac{p_{Y,1}}{p_Y^*} = -p_Y^*(1-\psi^{-1})\frac{\gamma\|\sigma\|^2}{2}\sum_{j=0}^{2}x_j\hat{\gamma}_j,$$

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:

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

State dynamics. The diffusion of the wealth share of investor j, $j = 1, 2$, is given by

$$\begin{aligned}\sigma_{x_j,1} &= x_j(\sigma_{j,1} - \sigma_{R,1})\\ &= x_j\left[\sum_{k=1}^{2}\frac{x_k}{x_a}(\hat{\gamma}_k + \hat{\nu}_k) - \frac{x_0}{x_a}\hat{\alpha}_{p,1} - (\hat{\gamma}_j + \hat{\nu}_j)\right]\sigma.\end{aligned}$$

The drift of the wealth share of investor j, $j = 1, 2$, is given by

$$\mu_{x_j,1} = x_j\left[(\gamma-1)\sigma\sigma_{j,1}^\top + \frac{\gamma\|\sigma\|^2}{2}(\psi-1)\left(\sum_{k=0}^{2}x_k\hat{\gamma}_k - \hat{\gamma}_j\right)\right] + \hat{\kappa}(\omega_j - x_j),$$

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

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

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

Pricing individual claims. We next solve for the valuation ratios $p_{k,t} \equiv P_{k,t}/Y_t$, $k \in \{D, B\}$. We consider the expansion

$$p_k(x, \overline{\alpha}_p, s; \epsilon) = p_{k,0} + p_{k,1}(x, \hat{\alpha}_p, \hat{s})\epsilon + O(\epsilon^2).$$

The pricing condition for claim k is

$$r + \sigma_{R_k}\eta^\top = \frac{s_k}{p_k} + \mu + \mu_{p_k} + \sigma\sigma_{p_k}^\top.$$

Let $s_D = s$ and $s_B = 1 - s$, so that $s_k = \overline{s}_k + \hat{s}_k\epsilon$, with $\overline{s}_D = \overline{s}, \overline{s}_B = 1 - \overline{s}, \hat{s}_D = \hat{s}$, and $\hat{s}_B = -\hat{s}$. At $\epsilon = 0$, we have

$$p_{k,0} = \frac{\overline{s}_k}{\delta}, \qquad \delta \equiv r^* + \gamma\|\sigma\|^2 - \mu.$$

The first-order pricing equation implies

$$\delta p_{k,1} + (r_1 + \pi_{Y,1})p_{k,0} = \hat{s}_k - \kappa_s\hat{s}\frac{\partial p_{k,1}}{\partial \hat{s}} + (1-\gamma)\sigma\frac{\partial p_{k,1}}{\partial \hat{s}}\overline{s}(1-\overline{s})\sigma_s^\top.$$

We conjecture $p_{D,1} = a_D(x) + b_D(x)\hat{s}$. Matching coefficients gives

$$b_D = \frac{1}{\delta + \kappa_s}, \qquad a_D = -(r_1 + \pi_{Y,1})\frac{\overline{s}}{\delta^2} - \frac{\gamma - 1}{\delta(\delta + \kappa_s)}\overline{s}(1-\overline{s})\sigma\sigma_s^\top.$$

Hence

$$p_{D,1} = \frac{\hat{s}}{\delta + \kappa_s} - (r_1 + \pi_{Y,1})\frac{\overline{s}}{\delta^2} - \frac{\gamma - 1}{\delta(\delta + \kappa_s)}\overline{s}(1-\overline{s})\sigma\sigma_s^\top.$$

Similarly, the normalized price of the bond claim is given by

$$p_{B,1} = -\frac{\hat{s}}{\delta + \kappa_s} - (r_1 + \pi_{Y,1})\frac{1-\overline{s}}{\delta^2} + \frac{\gamma - 1}{\delta(\delta + \kappa_s)}\overline{s}(1-\overline{s})\sigma\sigma_s^\top.$$

Adding both claims we recover the first-order correction of the aggregate claim:

$$p_{D,1} + p_{B,1} = -\frac{r_1 + \pi_{Y,1}}{\delta^2},$$

□

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

$$\kappa_p = 0, \qquad \sigma_p = 0_{1\times 3}.$$

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 $x = (x_1, x_2)$ 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: Diffusion for c_j and y_Y.

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

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

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:

$$\begin{aligned}\sigma_{y_Y}(\hat{X},\epsilon) &= \frac{y_{Y,\hat{X}}(\hat{X},\epsilon)}{y_Y(\hat{X},\epsilon)}\sigma_{\hat{X}}(\hat{X},\epsilon) \\ &= \left(1-\psi^{-1}\right)\frac{\gamma\|\sigma\|^2}{2y_{Y,0}(x,\hat{\alpha}_p)}\sum_{k=1}^{2}(\hat{\gamma}_k-\hat{\gamma}_0)\,x_k\sigma_{k,1}(x,\hat{\alpha}_p)\epsilon^2+O(\epsilon^3).\end{aligned}$$

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

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

The drift of y_Y, t is given by

$$\begin{aligned}\mu_{y_Y}(\hat{X},\epsilon) &= \frac{y_{Y,\hat{X}}(\hat{X},\epsilon)}{y_Y(\hat{X},\epsilon)}\mu_{\hat{X}}(\hat{X},\epsilon)+\sum_{k=1}^{3}\frac{1}{2}\sigma_{\hat{X},k}^{\top}(\hat{X},\epsilon)\frac{y_{Y,\hat{X}\hat{X}}(\hat{X},\epsilon)}{y_Y(\hat{X},\epsilon)}\sigma_{\hat{X},k}(\hat{X},\epsilon) \\ &= \frac{y_{Y,1,x}(x,\hat{\alpha}_p)}{y_{Y,0}(x,\hat{\alpha}_p)}\mu_{x,1}(x,\hat{\alpha}_p)\epsilon^2+O(\epsilon^3) \\ &= \left(1-\psi^{-1}\right)\frac{\gamma\|\sigma\|^2}{2y_{Y,0}(x,\hat{\alpha}_p)}\sum_{k=1}^{2}(\hat{\gamma}_k-\hat{\gamma}_0)\,x_k\mu_{x_k,1}(x,\hat{\alpha}_p)\epsilon^2+O(\epsilon^3).\end{aligned}$$

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:

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

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,

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

with

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

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

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

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

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

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

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

Substituting for $q_{j,1}$ and $q_{j,2}$ 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 1 is unconstrained and investor 2 may be constrained. Thus $\hat{\nu}_{1,1} = \hat{\nu}_{1,2} = 0$. If investor 2 is constrained, then $\hat{\nu}_{2,1} = \hat{\nu}_2$ and $\hat{\nu}_{2,2}$ is given by (A.5) with $j = 2$. If investor 2 is slack, then $\hat{\nu}_{2,1} = \hat{\nu}_{2,2} = 0$ and (A.3) applies to both active investors.

Step 3: Market price of risk.

The market price of risk is given by

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

where

$$\chi_{a,t} \equiv \frac{1 - x_0\overline{\alpha}_p}{x_a}$$

is the aggregate risky exposure that must be absorbed by active investors per unit of active wealth,

$$h_{a,t} \equiv \sum_{j=1}^{2}\frac{x_j}{x_a}\frac{1-\gamma_j^{-1}}{\psi-1}\frac{\sigma_{c_j,t}}{1+\nu_{j,t}}$$

is the average hedging demand of active investors, adjusted for leverage constraints, and

$$\gamma_{a,t} \equiv \left[\sum_{j=1}^{2} \frac{x_j}{x_a} \frac{1}{\gamma_j (1 + \nu_{j,t})} \right]^{-1}$$

is the average effective risk aversion of active investors.

We now expand each term to second order. Let

$$a_j \equiv \hat{\gamma}_j + \hat{\nu}_{j,1},$$

and define active-wealth-weighted averages by

$$\mathbb{E}_a[z_j] \equiv \sum_{j=1}^{2} \frac{x_j}{x_a} z_j.$$

Since

$$\frac{1}{\gamma_j (1 + \nu_{j,t})} = \frac{1}{\gamma} \left[1 - a_j \epsilon + \left(a_j^2 - \hat{\nu}_{j,2} - \hat{\gamma}_j \hat{\nu}_{j,1} \right) \epsilon^2 \right] + O(\epsilon^3),$$

we have

$$\gamma_{a,t} = \gamma \left[1 + \Gamma_{a,1} \epsilon + \Gamma_{a,2} \epsilon^2 \right] + O(\epsilon^3),$$

where

$$\begin{aligned} \Gamma_{a,1} &= \mathbb{E}_a[a_j], \\ \Gamma_{a,2} &= \mathbb{E}_a[\hat{\nu}_{j,2} + \hat{\gamma}_j \hat{\nu}_{j,1}] - \mathrm{Var}_a(a_j). \end{aligned}$$

Here

$$\mathrm{Var}_a(a_j) \equiv \mathbb{E}_a[a_j^2] - \mathbb{E}_a[a_j]^2.$$

Because $\overline{\alpha}_p = 1 + \hat{\alpha}_p \epsilon$ and $\kappa_p = \sigma_p = 0$ in this step,

$$\chi_{a,t} = 1 - \frac{x_0}{x_a} \hat{\alpha}_p \epsilon.$$

Moreover, since $y_Y = 1/p_Y$, the risk exposure of the aggregate claim is

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

Finally, since $\sigma_{c_j,t} = \sigma_{c_j,2}\epsilon^2 + O(\epsilon^3)$, hedging demand starts at second order:

$$h_{a,t} = h_{a,2}\epsilon^2 + O(\epsilon^3),$$

with

$$h_{a,2} = \frac{1-\gamma^{-1}}{\psi - 1}\mathbb{E}_a[\sigma_{c_j,2}].$$

Therefore,

$$\chi_{a,t}\sigma_{R_Y,t} - h_{a,t} = \sigma - \frac{x_0}{x_a}\hat{\alpha}_p\sigma\epsilon - \left(\sigma_{y_Y,2} + h_{a,2}\right)\epsilon^2 + O(\epsilon^3).$$

Combining these expansions gives

$$\begin{aligned}
\eta_0 &= \gamma\sigma, \\
\eta_1 &= \gamma\sigma\left[\mathbb{E}_a[a_j] - \frac{x_0}{x_a}\hat{\alpha}_p\right], \\
\eta_2 &= \gamma\left[\Gamma_{a,2}\sigma - \frac{x_0}{x_a}\hat{\alpha}_p\Gamma_{a,1}\sigma - \sigma_{y_Y,2} - h_{a,2}\right].
\end{aligned}$$

The first-order term coincides with the expression derived above because $a_j = \hat{\gamma}_j + \hat{\nu}_{j,1}$. The second-order term depends on $\hat{\nu}_{j,2}$ through $\Gamma_{a,2}$. Thus, when a leverage constraint binds at second order, (A.5) and the expression for η_2 form a linear system for η_2 and the second-order multiplier.

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.

$$r_t = \rho + \psi^{-1}\left(\mu_{P_Y,t} + \xi_t\right) + (1-\psi^{-1})\sum_{j=0}^{2} x_{j,t}\frac{\gamma_j}{2}\|\sigma_{j,t}\|^2 - \pi_{Y,t},$$

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

$$\xi_2 = \sum_{j=0}^{2} x_j \left[\mu_{c_j,2} + (1-\gamma)\sigma_{c_j,2}\sigma^\top\right],$$

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

$$\left[\sum_{j=0}^{2} x_j \frac{\gamma_j}{2}\|\sigma_j\|^2\right]_2 = \gamma \sum_{j=0}^{2} x_j \left[\sigma_{j,2}\sigma^\top + \frac{1}{2}\sigma_{j,1}\sigma_{j,1}^\top + \hat{\gamma}_j \sigma_{j,1}\sigma^\top\right].$$

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

$$\begin{aligned} r_2 = &-\psi^{-1}\left(\mu_{y_Y,2} + \sigma\sigma_{y_Y,2}^\top\right) + \psi^{-1}\sum_{j=0}^{2} x_j \left[\mu_{c_j,2} + (1-\gamma)\sigma_{c_j,2}\sigma^\top\right] \\ &+ (1-\psi^{-1})\gamma \sum_{j=0}^{2} x_j \left[\sigma_{j,2}\sigma^\top + \frac{1}{2}\sigma_{j,1}\sigma_{j,1}^\top + \hat{\gamma}_j\sigma_{j,1}\sigma^\top\right] - \pi_{Y,2}. \end{aligned} \tag{A.6}$$

Using market clearing for the risky claim,

$$\sum_{j=0}^{2} x_j \sigma_{j,2} = \sigma_{R_Y,2} = -\sigma_{y_Y,2},$$

we can also write

$$r_2 = -\psi^{-1}\left(\mu_{y_Y,2} + \sigma\sigma_{y_Y,2}^\top\right) + \psi^{-1}\sum_{j=0}^{2} x_j\left[\mu_{c_j,2} + (1-\gamma)\sigma_{c_j,2}\sigma^\top\right]$$
$$+ (1-\psi^{-1})\gamma\left[-\sigma_{y_Y,2}\sigma^\top + \sum_{j=0}^{2} x_j\left(\frac{1}{2}\sigma_{j,1}\sigma_{j,1}^\top + \hat{\gamma}_j\sigma_{j,1}\sigma^\top\right)\right] - \pi_{Y,2}. \tag{A.7}$$

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

$$y_{Y,2} = (1-\psi^{-1})\mu_{y_Y,2} + (1-\psi^{-1})\sigma\sigma_{y_Y,2}^\top + \psi^{-1}\sum_{j=0}^{2} x_j\left[\mu_{c_j,2} + (1-\gamma)\sigma_{c_j,2}\sigma^\top\right]$$
$$+ (1-\psi^{-1})\gamma\left[-\sigma_{y_Y,2}\sigma^\top + \sum_{j=0}^{2} x_j\left(\frac{1}{2}\sigma_{j,1}\sigma_{j,1}^\top + \hat{\gamma}_j\sigma_{j,1}\sigma^\top\right)\right].$$

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 $\sum_{j=0}^{2} x_j = 1$, the terms involving $\mu_{y_Y,2}$ cancel. The terms involving $\sigma_{y_Y,2}\sigma^\top$ also cancel.

Therefore,

$$y_{Y,2} = (1-\psi^{-1})\gamma \sum_{j=0}^{2} x_j \left(\frac{1}{2}\sigma_{j,1}\sigma_{j,1}^{\top} + \hat{\gamma}_j \sigma_{j,1}\sigma^{\top} \right). \tag{A.8}$$

Equivalently,

$$y_{Y,2} = (1-\psi^{-1})\gamma \|\sigma\|^2 \sum_{j=0}^{2} x_j \left(\frac{1}{2} m_{j,1}^2 + \hat{\gamma}_j m_{j,1} \right), \tag{A.9}$$

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 $\hat{\alpha}_p$. Write

$$m_{j,1} = \bar{m}_j - \ell_j \hat{\alpha}_p.$$

For passive investors,

$$\bar{m}_0 = 0, \qquad \ell_0 = -1,$$

so that $m_{0,1} = \hat{\alpha}_p$. For active investors,

$$\bar{m}_1 = \frac{x_2}{x_a}(\hat{\gamma}_2 - \hat{\gamma}_1 + \hat{\nu}_2), \qquad \ell_1 = \frac{x_0}{x_a},$$
$$\bar{m}_2 = \frac{x_1}{x_a}(\hat{\gamma}_1 - \hat{\gamma}_2 - \hat{\nu}_2), \qquad \ell_2 = \frac{x_0}{x_a}.$$

These definitions satisfy the first-order market-clearing condition

$$\sum_{j=0}^{2} x_j m_{j,1} = 0.$$

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

$$a_Y^0(x) = (1-\psi^{-1})\gamma\|\sigma\|^2 \sum_{j=0}^{2} x_j \left(\frac{1}{2}\bar{m}_j^2 + \hat{\gamma}_j \bar{m}_j\right),$$

$$b_Y^0(x) = -(1-\psi^{-1})\gamma\|\sigma\|^2 \sum_{j=0}^{2} x_j \ell_j \left(\bar{m}_j + \hat{\gamma}_j\right),$$

$$d_Y^0(x) = \frac{1}{2}(1-\psi^{-1})\gamma\|\sigma\|^2 \sum_{j=0}^{2} x_j \ell_j^2.$$

Using the values of ℓ_j, the quadratic coefficient can be simplified to

$$d_Y^0(x) = \frac{1}{2}(1-\psi^{-1})\gamma\|\sigma\|^2 \left[x_0 + \frac{x_0^2}{x_a}\right] = \frac{1}{2}(1-\psi^{-1})\gamma\|\sigma\|^2 \frac{x_0}{x_a}.$$

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.

Step 8: The dividend claim.

The normalized price of the dividend claim, $p_D = P_D/Y$, satisfies

$$r_t + (\sigma + \sigma_{p_D,t})\eta_t^\top = \frac{s_t}{p_{D,t}} + \mu + \mu_{p_D,t} + \sigma\sigma_{p_D,t}^\top.$$

Equivalently, after multiplying by $p_{D,t}$,

$$\left[r_t + \sigma\eta_t^\top - \mu\right] p_{D,t} = s_t + p_{D,x,t}\mu_{x,t} + p_{D,s,t}\mu_{s,t} + \frac{1}{2}\sum_{k=1}^{3}\Sigma_{D,k,t}^\top p_{D,XX,t}\Sigma_{D,k,t} + (\sigma - \eta_t)\left(p_{D,x,t}\sigma_{x,t} + p_{D,s,t}\sigma_{s,t}\right)^\top,$$

where $X = (x, s)$, $p_{D,XX}$ is the Hessian with respect to (x, s), and $\Sigma_{D,k,t}$ is the k-th column of the diffusion matrix of (x, s).

We now introduce the local coordinate

$$s = \bar{s} + \epsilon\hat{s}.$$

The function p_D is a function of the physical state (x, s), but its perturbation coefficients are functions of $(x, \hat{s})$:

$$p_D(x, s; \epsilon) = p_{D,0} + p_{D,1}(x, \hat{s})\epsilon + p_{D,2}(x, \hat{s})\epsilon^2 + O(\epsilon^3).$$

Because $\hat{s} = (s - \bar{s})/\epsilon$, derivatives with respect to the physical state s satisfy

$$p_{D,s} = p_{D,1,\hat{s}} + \epsilon p_{D,2,\hat{s}} + O(\epsilon^2),$$
$$p_{D,ss} = \epsilon^{-1} p_{D,1,\hat{s}\hat{s}} + p_{D,2,\hat{s}\hat{s}} + O(\epsilon).$$

In the first-order solution below, $p_{D,1}$ is linear in $\hat{s}$, so $p_{D,1,\hat{s}\hat{s}} = 0$ and the second derivative is well behaved:

$$p_{D,ss} = p_{D,2,\hat{s}\hat{s}} + O(\epsilon).$$

Similarly,

$$p_{D,x} = \epsilon p_{D,1,x} + \epsilon^2 p_{D,2,x} + O(\epsilon^3), \qquad p_{D,xs} = p_{D,1,x\hat{s}} + \epsilon p_{D,2,x\hat{s}} + O(\epsilon^2).$$

The dividend-share process implies

$$\mu_{s,t} = -\kappa_s(s_t - \bar{s}) = -\kappa_s \hat{s}\epsilon,$$
$$\sigma_{s,t} = s_t(1 - s_t)\sigma_s\epsilon = \bar{s}(1 - \bar{s})\sigma_s\epsilon + (1 - 2\bar{s})\hat{s}\sigma_s\epsilon^2 + O(\epsilon^3).$$

The wealth-share dynamics satisfy

$$\mu_{x,t} = \mu_{x,1}\epsilon + O(\epsilon^2), \qquad \sigma_{x,t} = \sigma_{x,1}\epsilon + O(\epsilon^2).$$

Define

$$A_t \equiv r_t + \sigma\eta_t^\top - \mu = \delta + A_1\epsilon + A_2\epsilon^2 + O(\epsilon^3),$$

where

$$\delta = r_0 + \sigma\eta_0^\top - \mu, \qquad A_1 = r_1 + \sigma\eta_1^\top, \qquad A_2 = r_2 + \sigma\eta_2^\top.$$

The payoff satisfies $s_t = \bar{s} + \hat{s}\epsilon$.

The zeroth-order condition is

$$\delta p_{D,0} = \bar{s}, \qquad \text{so} \qquad p_{D,0} = \frac{\bar{s}}{\delta}.$$

At first order, the only state-derivative terms come from the drift of s and from the order-ϵ diffusion of s evaluated at $\bar{s}$. The first-order correction satisfies

$$\delta p_{D,1} + A_1 p_{D,0} = \hat{s} - \kappa_s \hat{s} p_{D,1,\hat{s}} + (\sigma - \eta_0)\left[p_{D,1,\hat{s}}\bar{s}(1 - \bar{s})\sigma_s\right]^\top.$$

This equation is the same first-order pricing equation derived above. Its solution is

$$p_{D,1} = \frac{\hat{s}}{\delta + \kappa_s} - (r_1 + \pi_{Y,1})\frac{\bar{s}}{\delta^2} - \frac{\gamma - 1}{\delta(\delta + \kappa_s)}\bar{s}(1 - \bar{s})\sigma\sigma_s^\top.$$

At second order, using $\mu_{s,t} = -\kappa_s \hat{s}\epsilon$ and $p_{D,s} = p_{D,1,\hat{s}} + \epsilon p_{D,2,\hat{s}} + O(\epsilon^2)$, the correction satisfies

$$\begin{aligned}\delta p_{D,2} + A_1 p_{D,1} + A_2 p_{D,0} = &-\kappa_s \hat{s} p_{D,2,\hat{s}} + (\sigma - \eta_0)\left[p_{D,2,\hat{s}}\bar{s}(1 - \bar{s})\sigma_s\right]^\top \\ &+ p_{D,1,x}\mu_{x,1} + (\sigma - \eta_0)\left[p_{D,1,x}\sigma_{x,1}\right]^\top - \eta_1\left[p_{D,1,\hat{s}}\bar{s}(1 - \bar{s})\sigma_s\right]^\top \\ &+ (\sigma - \eta_0)\left[p_{D,1,\hat{s}}(1 - 2\bar{s})\hat{s}\sigma_s\right]^\top + \frac{1}{2}p_{D,2,\hat{s}\hat{s}}\bar{s}^2(1 - \bar{s})^2\sigma_s\sigma_s^\top \quad \text{(A.10)}\end{aligned}$$

The potential second-order Hessian terms involving $p_{D,1,x\hat{s}}$ and $p_{D,1,\hat{s}\hat{s}}$ vanish because $p_{D,1}$ is additively separable in $(x, \hat{s})$ and linear in $\hat{s}$. Thus, after these simplifications, the right-hand side of (A.10) is at most affine in $\hat{s}$, apart from the operator applied to $p_{D,2}$. We can solve (A.10) by matching coefficients in $\hat{s}$. Let

$$B_s \equiv \bar{s}(1 - \bar{s}), \qquad \Lambda_s \equiv (\sigma - \eta_0)\sigma_s^\top = (1 - \gamma)\sigma\sigma_s^\top.$$

Write the first-order solution as

$$p_{D,1}(x, \hat{s}) = a_1(x, \hat{\alpha}_p) + b_1\hat{s},$$

where

$$b_1 = \frac{1}{\delta + \kappa_s},$$

and

$$a_1(x, \hat{\alpha}_p) = -(r_1 + \pi_{Y,1})\frac{\bar{s}}{\delta^2} - \frac{\gamma - 1}{\delta(\delta + \kappa_s)}B_s\sigma\sigma_s^\top.$$

Equivalently, using $A_1 = r_1 + \sigma\eta_1^\top = r_1 + \pi_{Y,1}$,

$$a_1(x, \hat{\alpha}_p) = -\frac{A_1 p_{D,0}}{\delta} + \frac{\Lambda_s B_s}{\delta(\delta + \kappa_s)}.$$

Conjecture that the second-order correction is quadratic in $\hat{s}$:

$$p_{D,2}(x, \hat{s}) = a_2(x, \hat{\alpha}_p) + b_2(x, \hat{\alpha}_p)\hat{s} + c_2(x, \hat{\alpha}_p)\hat{s}^2.$$

Matching the coefficient on $\hat{s}^2$ gives

$$(\delta + \kappa_s)c_2 = 0,$$

so

$$c_2 = 0.$$

Thus, in this benchmark case, the second-order correction is affine in $\hat{s}$.

Matching the coefficient on $\hat{s}$ gives

$$(\delta + \kappa_s) b_2 = -A_1 b_1 + \Lambda_s (1 - 2\bar{s}) b_1,$$

so

$$b_2(x, \hat{\alpha}_p) = \frac{\Lambda_s (1 - 2\bar{s}) - A_1}{(\delta + \kappa_s)^2}.$$

Finally, matching the constant term gives

$$\delta a_2 = -A_1 a_1 - A_2 p_{D,0} + \Lambda_s B_s b_2 + \mathcal{M}_x - \eta_1 \left(b_1 B_s \sigma_s\right)^\top,$$

where

$$\mathcal{M}_x \equiv a_{1,x} \mu_{x,1} + (\sigma - \eta_0) \left[a_{1,x} \sigma_{x,1}\right]^\top.$$

Therefore,

$$a_2(x, \hat{\alpha}_p) = \frac{1}{\delta} \left[-A_1 a_1 - A_2 p_{D,0} + \Lambda_s B_s b_2 + \mathcal{M}_x - \eta_1 \left(b_1 B_s \sigma_s\right)^\top\right].$$

Combining the coefficient restrictions, the second-order correction is

$$p_{D,2}(x, \hat{s}) = a_2(x, \hat{\alpha}_p) + b_2(x, \hat{\alpha}_p) \hat{s}.$$

The term proportional to $\eta_1 (b_1 B_s \sigma_s)^\top$ captures the interaction between first-order changes in the market price of risk and the first-order dividend-share exposure of the dividend claim.

Step 9: Pricing impact for the dividend claim.

We now compute the derivative of the second-order dividend-claim price with respect to the passive investors' portfolio share $\hat{\alpha}_p$. The coefficient representation from Step 8 gives

$$p_{D,2}(x, \hat{s}) = a_2(x, \hat{\alpha}_p) + b_2(x, \hat{\alpha}_p) \hat{s},$$

where

$$b_2(x, \hat{\alpha}_p) = \frac{\Lambda_s (1 - 2\bar{s}) - A_1}{(\delta + \kappa_s)^2}$$

and

$$a_2(x, \hat{\alpha}_p) = \frac{1}{\delta} \left[-A_1 a_1 - A_2 p_{D,0} + \Lambda_s B_s b_2 + \mathcal{M}_x - \eta_1 \left(b_1 B_s \sigma_s\right)^\top\right].$$

Recall that

$$A_1 = r_1 + \sigma\eta_1^\top = r_1 + \pi_{Y,1}.$$

The first-order aggregate-price term is independent of $\hat{\alpha}_p$, so

$$\frac{\partial A_1}{\partial \hat{\alpha}_p} = 0.$$

It follows that a_1 and b_2 are also independent of $\hat{\alpha}_p$, except through the wealth-share dynamics term $\mathcal{M}_x$ in a_2. Therefore,

$$\frac{\partial p_{D,2}}{\partial \hat{\alpha}_p} = \frac{\partial a_2}{\partial \hat{\alpha}_p},$$

and

$$\frac{\partial p_{D,2}}{\partial \hat{\alpha}_p} = \frac{1}{\delta}\left[-p_{D,0}\frac{\partial A_2}{\partial \hat{\alpha}_p} + \frac{\partial \mathcal{M}_x}{\partial \hat{\alpha}_p} - \frac{\partial \eta_1}{\partial \hat{\alpha}_p}\left(b_1 B_s \sigma_s\right)^\top\right].$$

Using $p_{D,0} = \bar{s}/\delta$ and $b_1 = 1/(\delta + \kappa_s)$, this can be written as

$$\frac{\partial p_{D,2}}{\partial \hat{\alpha}_p} = -\frac{\bar{s}}{\delta^2}\frac{\partial A_2}{\partial \hat{\alpha}_p} - \frac{B_s}{\delta(\delta+\kappa_s)}\frac{\partial \eta_1}{\partial \hat{\alpha}_p}\sigma_s^\top + \frac{1}{\delta}\frac{\partial \mathcal{M}_x}{\partial \hat{\alpha}_p}.$$

The first term is the aggregate-price-impact channel. The second term is specific to the dividend claim: passive demand changes the first-order market price of risk, and the dividend claim loads on dividend-share risk through $b_1 B_s \sigma_s$. The last term captures the indirect effect operating through first-order wealth-share dynamics. Since

$$\mathcal{M}_x = a_{1,x}\mu_{x,1} + (\sigma - \eta_0)\left[a_{1,x}\sigma_{x,1}\right]^\top,$$

and a_1 is independent of $\hat{\alpha}_p$, this indirect term is

$$\frac{\partial \mathcal{M}_x}{\partial \hat{\alpha}_p} = a_{1,x}\frac{\partial \mu_{x,1}}{\partial \hat{\alpha}_p} + (\sigma - \eta_0)\left[a_{1,x}\frac{\partial \sigma_{x,1}}{\partial \hat{\alpha}_p}\right]^\top.$$

This term is equal to zero. To see this, recall that the first-order diffusion of wealth shares is

$$\sigma_{x_j,1} = x_j \sigma_{j,1},$$

and the first-order drift of wealth shares is

$$\mu_{x_j,1} = x_j\left[(\eta_0 - \sigma)\sigma_{j,1}^\top - c_{j,1} + y_{Y,1}\right].$$

Since $c_{j,1}$ and $y_{Y,1}$ are independent of $\hat{\alpha}_p$, differentiating with respect to $\hat{\alpha}_p$ gives

$$\frac{\partial \mu_{x_j,1}}{\partial \hat{\alpha}_p} = x_j(\eta_0 - \sigma)\frac{\partial \sigma_{j,1}^\top}{\partial \hat{\alpha}_p}, \qquad \frac{\partial \sigma_{x_j,1}}{\partial \hat{\alpha}_p} = x_j \frac{\partial \sigma_{j,1}}{\partial \hat{\alpha}_p}.$$

Therefore,

$$\begin{aligned}\frac{\partial \mathcal{M}_x}{\partial \hat{\alpha}_p} &= \sum_{j=1}^{2} a_{1,x_j} x_j (\eta_0 - \sigma)\frac{\partial \sigma_{j,1}^\top}{\partial \hat{\alpha}_p} + (\sigma - \eta_0)\sum_{j=1}^{2} a_{1,x_j} x_j \frac{\partial \sigma_{j,1}^\top}{\partial \hat{\alpha}_p} \\ &= 0.\end{aligned}$$

Thus the indirect wealth-share-dynamics channel drops out of the dividend-claim price impact. Finally, since

$$p_D(x, s; \epsilon) = p_{D,0} + p_{D,1}(x, \hat{s})\epsilon + p_{D,2}(x, \hat{s})\epsilon^2 + O(\epsilon^3)$$

and $p_{D,1}$ is independent of $\hat{\alpha}_p$, the price impact of passive flows is

$$\frac{\partial p_D(x, s; \epsilon)}{\partial F(x, s; \epsilon)} = \frac{\epsilon}{x_0}\frac{\partial p_{D,2}}{\partial \hat{\alpha}_p} + O(\epsilon^2).$$

Thus, the dividend-claim price impact is the sum of an aggregate-price-impact channel and a direct dividend-risk-exposure channel.

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

$$\begin{aligned}\mu_{\overline{\alpha}_p,t} &= \kappa_p(\hat{\alpha} - \hat{\alpha}_p)\epsilon, \\ \sigma_{\overline{\alpha}_p,t} &= \sigma_p\epsilon.\end{aligned}$$

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 $\overline{\alpha}_p$ satisfy the same chain-rule structure as for s. 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 $\sigma_p = 0_{1\times 3}$.

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

$$\Delta\sigma_{c_j,2} = \frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\sigma_p,$$
$$\Delta\sigma_{y_Y,2} = \frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}}\sigma_p.$$

Therefore,

$$\sigma_{c_j,2} = \sigma^0_{c_j,2} + \frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\sigma_p,$$
$$\sigma_{y_Y,2} = \sigma^0_{y_Y,2} + \frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}}\sigma_p.$$

Similarly, the new drift terms generated by the time variation in $\overline{\alpha}_p$ are

$$\Delta\mu_{c_j,2} = \frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\kappa_p(\hat{\alpha}-\hat{\alpha}_p) + \frac{1}{2}\frac{c_{j,2,\hat{\alpha}_p\hat{\alpha}_p}}{c_{j,0}}\sigma_p\sigma_p^\top,$$
$$\Delta\mu_{y_Y,2} = \frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}}\kappa_p(\hat{\alpha}-\hat{\alpha}_p) + \frac{1}{2}\frac{y_{Y,2,\hat{\alpha}_p\hat{\alpha}_p}}{y_{Y,0}}\sigma_p\sigma_p^\top.$$

Thus the only new second-order terms are proportional to κ_p or σ_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

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

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 σ_p. Thus,

$$\Delta q_{j,2} = \frac{\Delta\eta_2}{\gamma} + \frac{1-\gamma^{-1}}{\psi-1}\Delta\sigma_{c_j,2}.$$

Using Steps 1 and 2,

$$\Delta q_{j,2} = \frac{\Delta\eta_2}{\gamma} + \frac{1-\gamma^{-1}}{\psi-1}\frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\sigma_p.$$

For an unconstrained investor, the multiplier is zero, so the incremental second-order exposure is simply

$$\Delta\sigma_{j,2} = \Delta q_{j,2}.$$

Therefore,

$$\Delta\sigma_{j,2} = \frac{\Delta\eta_2}{\gamma} + \frac{1-\gamma^{-1}}{\psi-1}\frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\sigma_p.$$

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

$$\Delta\sigma_{j,2} = \Delta q_{j,2} - \Delta\hat{\nu}_{j,2}\sigma.$$

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

$$\Delta\hat{\nu}_{j,2} = \frac{\Delta q_{j,2}\sigma^\top}{\sigma\sigma^\top},$$

and therefore

$$\Delta\sigma_{j,2} = \Delta q_{j,2} - \frac{\Delta q_{j,2}\sigma^\top}{\sigma\sigma^\top}\sigma.$$

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 σ_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 σ_p. Second, active investors have additional hedging demand because their consumption-wealth ratios load on σ_p. Third, if the leverage constraint binds for investor 2, the second-order multiplier changes, which changes effective aggregate risk aversion.

From Step 1,

$$\Delta\sigma_{y_Y,2} = \frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}}\sigma_p,$$

and the additional hedging demand is

$$\Delta h_{a,2} = \frac{1-\gamma^{-1}}{\psi-1}\mathbb{E}_a[\Delta\sigma_{c_j,2}] = \frac{1-\gamma^{-1}}{\psi-1}\mathbb{E}_a\left[\frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\right]\sigma_p.$$

The new contribution to effective aggregate risk aversion comes from the second-order multiplier. 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

$$\Delta\eta_2 = \gamma \left[\mathbb{E}_a[\Delta\hat{\nu}_{j,2}]\sigma - \Delta\sigma_{y_Y,2} - \Delta h_{a,2} \right]. \tag{A.11}$$

If no leverage constraint binds, then $\Delta\hat{\nu}_{j,2} = 0$ for all active investors and (A.11) reduces to

$$\Delta\eta_2 = -\gamma \left[\frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}} + \frac{1-\gamma^{-1}}{\psi - 1} \mathbb{E}_a \left[\frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}} \right] \right] \sigma_p.$$

If the leverage constraint binds for investor 2, then $\Delta\hat{\nu}_{1,2} = 0$ and $\Delta\hat{\nu}_{2,2}$ 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 2 is

$$\Delta q_{2,2} = \frac{\Delta\eta_2}{\gamma} + H_2\sigma_p.$$

The binding constraint requires the component of $\Delta\sigma_{2,2}$ in the direction of σ to be zero. Thus,

$$\Delta\hat{\nu}_{2,2} = \frac{\Delta q_{2,2}\sigma^\top}{\sigma\sigma^\top} = \frac{\Delta\eta_2\sigma^\top}{\gamma\sigma\sigma^\top} + H_2\frac{\sigma_p\sigma^\top}{\sigma\sigma^\top}.$$

$$\begin{aligned} \Delta\eta_2 &= \gamma \left[\frac{x_2}{x_a}\Delta\hat{\nu}_{2,2}\sigma - \frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}}\sigma_p - \frac{1-\gamma^{-1}}{\psi-1}\mathbb{E}_a\left[\frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\right]\sigma_p \right], \\ \Delta\hat{\nu}_{2,2} &= \frac{\Delta\eta_2\sigma^\top}{\gamma\sigma\sigma^\top} + H_2\frac{\sigma_p\sigma^\top}{\sigma\sigma^\top}. \end{aligned} \tag{A.12}$$

Equations (A.12) jointly determine the new market-price-of-risk component and the new second-order multiplier when investor 2 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

$$C_p \equiv \frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}} + \frac{1-\gamma^{-1}}{\psi-1}\mathbb{E}_a\left[\frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\right], \qquad \zeta_p \equiv \frac{\sigma_p\sigma^\top}{\sigma\sigma^\top}.$$

Then (A.12) becomes

$$\Delta\eta_2 = \gamma\left[\frac{x_2}{x_a}\Delta\hat{\nu}_{2,2}\sigma - C_p\sigma_p\right],$$
$$\Delta\hat{\nu}_{2,2} = \frac{\Delta\eta_2\sigma^\top}{\gamma\sigma\sigma^\top} + H_2\zeta_p.$$

Projecting the first equation on σ and substituting into the second equation gives

$$\Delta\hat{\nu}_{2,2} = \frac{x_2}{x_a}\Delta\hat{\nu}_{2,2} - C_p\zeta_p + H_2\zeta_p.$$

Since $1 - x_2/x_a = x_1/x_a$, we obtain

$$\Delta\hat{\nu}_{2,2} = \frac{x_a}{x_1}(H_2 - C_p)\zeta_p.$$

Substituting back into the first equation yields

$$\Delta\eta_2 = -\gamma C_p\sigma_p + \gamma\frac{x_2}{x_1}(H_2 - C_p)\zeta_p\sigma.$$

Equivalently, using

$$\mathbb{E}_a[H_j] = \frac{x_1}{x_a}H_1 + \frac{x_2}{x_a}H_2, \qquad C_p = \frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}} + \mathbb{E}_a[H_j],$$

we can write

$$H_2 - C_p = \frac{x_1}{x_a}(H_2 - H_1) - \frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}}.$$

Combining the expressions above, we obtain

$$\Delta\eta_2 = -\gamma\left[\frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}} + \mathbb{E}_a[H_j]\right]\sigma_p - \gamma\frac{x_2}{x_1}\left[\frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}} + \frac{x_1}{x_a}(H_1 - H_2)\right]\frac{\sigma_p\sigma^\top}{\sigma\sigma^\top}\sigma.$$

Thus the constrained investor absorbs the component of the new desired exposure in the direction of σ, while the orthogonal component of σ_p remains priced through $\Delta\eta_2$.

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

$$\Delta\pi_{Y,2} = -\Delta\sigma_{y_Y,2}\eta_0^\top + \sigma\Delta\eta_2^\top. \tag{A.13}$$

Using $\eta_0 = \gamma\sigma$ and Step 1,

$$\Delta\pi_{Y,2} = -\gamma\frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}}\sigma_p\sigma^\top + \sigma\Delta\eta_2^\top.$$

If no leverage constraint binds, substituting the expression for $\Delta\eta_2$ from Step 3 gives

$$\Delta\pi_{Y,2} = -\gamma\left[\frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}}\sigma_p\sigma^\top + \left(\frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}} + \frac{1-\gamma^{-1}}{\psi-1}\mathbb{E}_a\left[\frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\right]\right)\sigma\sigma_p^\top\right].$$

When the leverage constraint binds for investor 2, the same formula (A.13) applies, but $\Delta\eta_2$ is jointly determined with $\Delta\hat{\nu}_{2,2}$ by (A.12).

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

$$\begin{aligned}\Delta r_2 = &-\psi^{-1}\left(\Delta\mu_{y_Y,2} + \sigma\Delta\sigma_{y_Y,2}^\top\right)\\ &+\psi^{-1}\sum_{j=0}^{2} x_j\left[\Delta\mu_{c_j,2} + (1-\gamma)\Delta\sigma_{c_j,2}\sigma^\top\right]\\ &+(1-\psi^{-1})\gamma\sum_{j=0}^{2} x_j\Delta\sigma_{j,2}\sigma^\top - \Delta\pi_{Y,2}.\end{aligned} \tag{A.14}$$

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

$$\begin{aligned}\Delta\sigma_{c_j,2} &= \frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\sigma_p,\\ \Delta\sigma_{y_Y,2} &= \frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}}\sigma_p,\\ \Delta\mu_{c_j,2} &= \frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\kappa_p(\hat{\alpha}-\hat{\alpha}_p) + \frac{1}{2}\frac{c_{j,2,\hat{\alpha}_p\hat{\alpha}_p}}{c_{j,0}}\sigma_p\sigma_p^\top,\\ \Delta\mu_{y_Y,2} &= \frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}}\kappa_p(\hat{\alpha}-\hat{\alpha}_p) + \frac{1}{2}\frac{y_{Y,2,\hat{\alpha}_p\hat{\alpha}_p}}{y_{Y,0}}\sigma_p\sigma_p^\top.\end{aligned}$$

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.14) gives

$$\begin{aligned}\Delta r_2 = &-\psi^{-1}\Delta\mu_{y_Y,2} - \psi^{-1}\sigma\Delta\sigma_{y_Y,2}^\top\\ &+\psi^{-1}\sum_{j=0}^{2} x_j\left[\Delta\mu_{c_j,2} + (1-\gamma)\Delta\sigma_{c_j,2}\sigma^\top\right]\\ &+(1-\psi^{-1})\gamma\sum_{j=0}^{2} x_j\Delta\sigma_{j,2}\sigma^\top + \Delta\sigma_{y_Y,2}\eta_0^\top - \sigma\Delta\eta_2^\top.\end{aligned}$$

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,

$$\begin{aligned}\Delta r_2 = &-\psi^{-1}\Delta\mu_{y_Y,2} + \psi^{-1}\sum_{j=0}^{2} x_j \Delta\mu_{c_j,2} \\ &+ \psi^{-1}(1-\gamma)\sum_{j=0}^{2} x_j \Delta\sigma_{c_j,2}\sigma^\top + (1-\psi^{-1})\gamma\sum_{j=0}^{2} x_j \Delta\sigma_{j,2}\sigma^\top \\ &+ \left(\gamma - \psi^{-1}\right)\sigma\Delta\sigma_{y_Y,2}^\top - \sigma\Delta\eta_2^\top. \end{aligned} \tag{A.15}$$

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 $\Delta y_{Y,2} = \sum_{j=0}^{2} x_j \Delta c_{j,2}$ implies

$$\sum_{j=0}^{2} x_j \Delta\mu_{c_j,2} = \Delta\mu_{y_Y,2}, \qquad \sum_{j=0}^{2} x_j \Delta\sigma_{c_j,2} = \Delta\sigma_{y_Y,2}.$$

The drift terms in (A.15) 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,

$$\sum_{j=0}^{2} x_j \Delta\sigma_{j,2} = -\Delta\sigma_{y_Y,2},$$

the remaining exposure terms cancel as well. Hence the incremental interest-rate correction reduces to

$$\Delta r_2 = -\sigma\Delta\eta_2^\top. \tag{A.16}$$

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 2, $\Delta\eta_2$ and $\Delta\hat{\nu}_{2,2}$ are jointly determined by (A.12).

Step 6 (time-varying α_p): Dividend yield.

We write the second-order dividend yield as

$$y_{Y,2} = y_{Y,2}^0 + \Delta y_{Y,2},$$

where $y_{Y,2}^0$ 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,

$$y_{Y,2} = r_2 + \pi_{Y,2} + \mu_{y_Y,2} + \sigma\sigma_{y_Y,2}^\top,$$

so the incremental correction satisfies

$$\Delta y_{Y,2} = \Delta r_2 + \Delta\pi_{Y,2} + \Delta\mu_{y_Y,2} + \sigma\Delta\sigma_{y_Y,2}^\top.$$

Using (A.16) and (A.13), we obtain

$$\begin{aligned}\Delta y_{Y,2} &= -\sigma\Delta\eta_2^\top + \left(-\Delta\sigma_{y_Y,2}\eta_0^\top + \sigma\Delta\eta_2^\top\right) + \Delta\mu_{y_Y,2} + \sigma\Delta\sigma_{y_Y,2}^\top \\ &= \Delta\mu_{y_Y,2} + (1-\gamma)\Delta\sigma_{y_Y,2}\sigma^\top,\end{aligned}$$

where the terms involving $\Delta\eta_2$ cancel and we use $\eta_0 = \gamma\sigma$.

Using Step 1, this becomes

$$\Delta y_{Y,2} = \frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}}\kappa_p(\hat{\alpha} - \hat{\alpha}_p) + \frac{1}{2}\frac{y_{Y,2,\hat{\alpha}_p\hat{\alpha}_p}}{y_{Y,0}}\sigma_p\sigma_p^\top + (1-\gamma)\frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}}\sigma_p\sigma^\top.$$

Equivalently, since $y_{Y,2} = y_{Y,2}^0 + \Delta y_{Y,2}$, the incremental dividend-yield correction solves

$$\Delta y_{Y,2} = \frac{(y_{Y,2}^0 + \Delta y_{Y,2})_{\hat{\alpha}_p}}{y_{Y,0}}\left[\kappa_p(\hat{\alpha} - \hat{\alpha}_p) + (1-\gamma)\sigma_p\sigma^\top\right] + \frac{1}{2}\frac{(y_{Y,2}^0 + \Delta y_{Y,2})_{\hat{\alpha}_p\hat{\alpha}_p}}{y_{Y,0}}\sigma_p\sigma_p^\top. \quad \text{(A.17)}$$

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

$$p_{Y,2} = p_{Y,0}^3 y_{Y,1}^2 - p_{Y,0}^2 y_{Y,2},$$

so the incremental term is

$$\Delta p_{Y,2} = -p_{Y,0}^2\Delta y_{Y,2}.$$

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.17). Write the fixed-$\hat{\alpha}_p$ correction as

$$y_{Y,2}^0(x, \hat{\alpha}_p) = a_Y^0(x) + b_Y^0(x)\hat{\alpha}_p + d_Y^0(x)\hat{\alpha}_p^2,$$

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

$$M_p \equiv \kappa_p\hat{\alpha} + (1-\gamma)\sigma_p\sigma^\top, \qquad V_p \equiv \sigma_p\sigma_p^\top.$$

Then (A.17) can be written as

$$\Delta y_{Y,2} = \frac{1}{y_{Y,0}}(b_Y^T + 2d_Y^T\hat{\alpha}_p)\left(M_p - \kappa_p\hat{\alpha}_p\right) + \frac{d_Y^T}{y_{Y,0}}V_p. \tag{A.18}$$

We first match the coefficient on $\hat{\alpha}_p^2$. The only term in (A.18) 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

$$d_Y = -\frac{\frac{2\kappa_p}{y_{Y,0}}}{1 + \frac{2\kappa_p}{y_{Y,0}}}\, d_Y^0. \tag{A.19}$$

Equivalently, the quadratic coefficient of the total second-order correction is

$$d_Y^T = \frac{1}{1 + \frac{2\kappa_p}{y_{Y,0}}}\, d_Y^0. \tag{A.20}$$

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.18) 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

$$b_Y = \frac{2d_Y^T M_p - \kappa_p b_Y^0}{y_{Y,0} + \kappa_p}. \tag{A.21}$$

Substituting (A.20), 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

$$b_Y^T = b_Y^0 + b_Y = \frac{y_{Y,0} b_Y^0 + 2d_Y^T M_p}{y_{Y,0} + \kappa_p}. \tag{A.22}$$

Finally, we solve for the constant coefficient. Matching the constant term in (A.18) gives

$$a_Y = \frac{1}{y_{Y,0}} \left(b_Y^T M_p + d_Y^T V_p \right).$$

Using (A.22) and (A.20), this becomes

$$a_Y = \frac{1}{y_{Y,0}} \left[\frac{y_{Y,0} b_Y^0 + 2d_Y^T M_p}{y_{Y,0} + \kappa_p} M_p + d_Y^T V_p \right], \tag{A.23}$$

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.19), (A.21), and (A.23).

Step 8 (time-varying α_p): Consumption-wealth ratio.

We write the second-order consumption-wealth ratio as

$$c_{j,2} = c^0_{j,2} + \Delta c_{j,2},$$

where $c^0_{j,2}$ 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 $c^0_{j,2}$ and therefore do not appear in $\Delta c_{j,2}$.

Using Step 1, this becomes

$$\begin{aligned} \Delta c_{j,2} = (1-\psi)\left[\Delta r_2 + \Delta\eta_2\sigma^\top\right] \\ + \frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\kappa_p(\hat{\alpha}-\hat{\alpha}_p) + \frac{1}{2}\frac{c_{j,2,\hat{\alpha}_p\hat{\alpha}_p}}{c_{j,0}}\sigma_p\sigma_p^\top \\ + (1-\gamma)\frac{c_{j,2,\hat{\alpha}_p}}{c_{j,0}}\sigma_p\sigma^\top. \end{aligned} \tag{A.24}$$

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.16), (A.11), and (A.24) 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.24), the incremental correction to $c_{j,2}$ satisfies

$$\Delta c_{j,2} = \frac{(c^0_{j,2} + \Delta c_{j,2})_{\hat\alpha_p}}{c_{j,0}} \left[\kappa_p(\hat\alpha - \hat\alpha_p) + (1-\gamma)\sigma_p\sigma^\top\right] + \frac{1}{2}\frac{(c^0_{j,2} + \Delta c_{j,2})_{\hat\alpha_p\hat\alpha_p}}{c_{j,0}}\sigma_p\sigma_p^\top. \tag{A.25}$$

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

$$c^0_{j,2}(x, \hat\alpha_p) = a^0_j(x) + b^0_j(x)\hat\alpha_p + d^0_j(x)\hat\alpha_p^2,$$

and conjecture

$$\Delta c_{j,2}(x, \hat\alpha_p) = a_j(x) + b_j(x)\hat\alpha_p + d_j(x)\hat\alpha_p^2.$$

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

$$M_p \equiv \kappa_p\hat\alpha + (1-\gamma)\sigma_p\sigma^\top, \qquad V_p \equiv \sigma_p\sigma_p^\top.$$

Then (A.25) 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 $\hat\alpha_p^2$ 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

$$d_j = -\frac{\frac{2\kappa_p}{c_{j,0}}}{1 + \frac{2\kappa_p}{c_{j,0}}}\, d^0_j, \tag{A.26}$$

and the quadratic coefficient of the total correction is

$$d_j^T = \frac{1}{1 + \frac{2\kappa_p}{c_{j,0}}} d_j^0. \tag{A.27}$$

Matching the coefficient on $\hat{\alpha}_p$ gives

$$b_j = \frac{1}{c_{j,0}} \left(2 d_j^T M_p - \kappa_p b_j^T \right) = \frac{1}{c_{j,0}} \left(2 d_j^T M_p - \kappa_p (b_j^0 + b_j) \right).$$

Solving for b_j yields

$$b_j = \frac{2 d_j^T M_p - \kappa_p b_j^0}{c_{j,0} + \kappa_p}. \tag{A.28}$$

Equivalently, the linear coefficient of the total correction is

$$b_j^T = b_j^0 + b_j = \frac{c_{j,0} b_j^0 + 2 d_j^T M_p}{c_{j,0} + \kappa_p}. \tag{A.29}$$

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.29) and (A.27), this becomes

$$a_j = \frac{1}{c_{j,0}} \left[\frac{c_{j,0} b_j^0 + 2 d_j^T M_p}{c_{j,0} + \kappa_p} M_p + d_j^T V_p \right]. \tag{A.30}$$

Therefore,

$$\Delta c_{j,2}(x, \hat{\alpha}_p) = a_j(x) + b_j(x) \hat{\alpha}_p + d_j(x) \hat{\alpha}_p^2, \tag{A.31}$$

with d_j, b_j, and a_j given by (A.26), (A.28), and (A.30). Because $c_{j,0} = y_{Y,0}$ in the homogeneous benchmark economy and $y_{Y,2} = \sum_{j=0}^{2} x_j c_{j,2}$, aggregating (A.31) across investors recovers the coefficient formulas for $\Delta y_{Y,2}$ derived above.

Step 10 (time-varying α_p): Pricing the dividend claim.

The dividend claim satisfies the pricing condition

$$\left[r_t + \sigma \eta_t^\top - \mu \right] p_{D,t} = s_t + p_{D,x,t} \mu_{x,t} + p_{D,s,t} \mu_{s,t} + p_{D,\overline{\alpha}_p,t} \mu_{\overline{\alpha}_p,t} + \frac{1}{2} \sum_{k=1}^{3} \Sigma_{D,k,t}^\top p_{D,XX,t} \Sigma_{D,k,t}$$
$$+ (\sigma - \eta_t) \left(p_{D,x,t} \sigma_{x,t} + p_{D,s,t} \sigma_{s,t} + p_{D,\overline{\alpha}_p,t} \sigma_p \epsilon \right)^\top,$$

where now $X = (x, s, \overline{\alpha}_p)$. Relative to the fixed-$\hat{\alpha}_p$ case, the new terms are those involving derivatives with respect to $\overline{\alpha}_p$ and the covariance between s and $\overline{\alpha}_p$.

We write

$$p^T_{D,2}(x, \hat{s}, \hat{\alpha}_p) \equiv p^0_{D,2}(x, \hat{s}, \hat{\alpha}_p) + \Delta p_{D,2}(x, \hat{s}, \hat{\alpha}_p),$$

where $p^0_{D,2}$ is the fixed-$\hat{\alpha}_p$ correction derived above. From the fixed-$\hat{\alpha}_p$ solution,

$$p^0_{D,2}(x, \hat{s}, \hat{\alpha}_p) = a^0_2(x, \hat{\alpha}_p) + b^0_2(x)\hat{s},$$

where b^0_2 is independent of $\hat{\alpha}_p$. We conjecture that the time-varying-α_p correction has the same form,

$$p^T_{D,2}(x, \hat{s}, \hat{\alpha}_p) = a^T_2(x, \hat{\alpha}_p) + b^T_2(x, \hat{\alpha}_p)\hat{s}.$$

The simplification in Step 5 implies

$$\Delta A_2 \equiv \Delta r_2 + \sigma \Delta \eta_2^\top = 0.$$

Thus the aggregate-price term in the dividend-claim pricing equation is unchanged relative to the fixed-$\hat{\alpha}_p$ economy. The only new terms come from the generator associated with the additional state variable $\overline{\alpha}_p$. Using the chain rule for $\overline{\alpha}_p = 1 + \epsilon\hat{\alpha}_p$, the additional second-order operator applied to a function $f(\hat{s}, \hat{\alpha}_p)$ is

$$\mathcal{L}_\alpha f = f_{\hat{\alpha}_p}\left[\kappa_p(\hat{\alpha} - \hat{\alpha}_p) + (\sigma - \eta_0)\sigma_p^\top\right] + \frac{1}{2} f_{\hat{\alpha}_p\hat{\alpha}_p}\sigma_p\sigma_p^\top + B_s f_{\hat{s}\hat{\alpha}_p}\sigma_s\sigma_p^\top, \tag{A.32}$$

where $B_s \equiv \bar{s}(1 - \bar{s})$. Using $\eta_0 = \gamma\sigma$, define

$$M_p \equiv \kappa_p\hat{\alpha} + (1 - \gamma)\sigma_p\sigma^\top, \qquad V_p \equiv \sigma_p\sigma_p^\top.$$

Then

$$\mathcal{L}_\alpha f = f_{\hat{\alpha}_p}(M_p - \kappa_p\hat{\alpha}_p) + \frac{1}{2} f_{\hat{\alpha}_p\hat{\alpha}_p} V_p + B_s f_{\hat{s}\hat{\alpha}_p}\sigma_s\sigma_p^\top.$$

Subtracting the fixed-$\hat{\alpha}_p$ equation from the time-varying-α_p equation gives

$$\delta\Delta p_{D,2} = -\kappa_s\hat{s}\Delta p_{D,2,\hat{s}} + \Lambda_s B_s \Delta p_{D,2,\hat{s}} + \frac{1}{2} B_s^2 \sigma_s\sigma_s^\top \Delta p_{D,2,\hat{s}\hat{s}} + \mathcal{L}_\alpha p^T_{D,2},$$

where $\Lambda_s \equiv (\sigma - \eta_0)\sigma_s^\top$. Because $p^0_{D,2}$ is affine in $\hat{s}$ and b^0_2 is independent of $\hat{\alpha}_p$, the solution remains affine in $\hat{s}$. Moreover, matching the coefficient on $\hat{s}$ gives

$$(\delta + \kappa_s)\Delta b_2 = \mathcal{L}_\alpha b^T_2 = \mathcal{L}_\alpha \Delta b_2,$$

as b_2^0 is independent of $\hat{\alpha}_p$. Then, $\Delta b_2 = 0$ satisfies the equation above, so

$$b_2^T = b_2^0.$$

Thus, time variation in passive demand affects only the constant term in the dividend-claim correction. The cross term $B_s f_{\hat{s}\hat{\alpha}_p} \sigma_s \sigma_p^\top$ in (A.32) drops out because b_2^T is independent of $\hat{\alpha}_p$.

The constant term satisfies the scalar equation

$$\Delta a_2 = \frac{1}{\delta}\left[a_{2,\hat{\alpha}_p}^T (M_p - \kappa_p \hat{\alpha}_p) + \frac{1}{2} a_{2,\hat{\alpha}_p\hat{\alpha}_p}^T V_p\right]. \tag{A.33}$$

This equation is identical in form to the coefficient equation for the aggregate dividend yield. Indeed, in the benchmark economy $y_{Y,0} = \delta$, where $\delta \equiv r_0 + \sigma\eta_0^\top - \mu$. Write

$$a_2^0(x, \hat{\alpha}_p) = a_D^0(x) + b_D^0(x)\hat{\alpha}_p + d_D^0(x)\hat{\alpha}_p^2,$$

and

$$\Delta a_2(x, \hat{\alpha}_p) = a_D(x) + b_D(x)\hat{\alpha}_p + d_D(x)\hat{\alpha}_p^2.$$

Let $b_D^T \equiv b_D^0 + b_D$ and $d_D^T \equiv d_D^0 + d_D$ denote the linear and quadratic coefficients of $a_2^T = a_2^0 + \Delta a_2$. Matching coefficients in (A.33) gives

$$\begin{aligned}
d_D^T &= \frac{1}{1 + \frac{2\kappa_p}{\delta}} d_D^0, \\
b_D^T &= \frac{\delta b_D^0 + 2 d_D^T M_p}{\delta + \kappa_p}, \\
a_D &= \frac{1}{\delta}\left(b_D^T M_p + d_D^T V_p\right).
\end{aligned}$$

Therefore,

$$p_{D,2}^T(x, \hat{s}, \hat{\alpha}_p) = a_D^0(x) + a_D(x) + b_D^T(x)\hat{\alpha}_p + d_D^T(x)\hat{\alpha}_p^2 + b_2^0(x)\hat{s}.$$

Finally, since $p_{D,1}$ is independent of $\hat{\alpha}_p$, the price impact of passive flows for the dividend claim is

$$\frac{\partial p_D(x, s, \overline{\alpha}_p; \epsilon)}{\partial F(x, s, \overline{\alpha}_p; \epsilon)} = \frac{\epsilon}{x_0}\left(b_D^T + 2 d_D^T \hat{\alpha}_p\right) + O(\epsilon^2).$$

We now relate this expression explicitly to the price impact of the aggregate claim. Let

$$G_Y^T \equiv \frac{\partial y_{Y,2}^T}{\partial \hat{\alpha}_p} = b_Y^T + 2 d_Y^T \hat{\alpha}_p,$$

where $y^T_{Y,2} \equiv y^0_{Y,2} + \Delta y_{Y,2}$. Because $p_Y = 1/y_Y$ and $y_{Y,1}$ is independent of $\hat{\alpha}_p$,

$$\frac{\partial p_Y(x, \overline{\alpha}_p; \epsilon)}{\partial F(x, \overline{\alpha}_p; \epsilon)} = -\frac{\epsilon}{x_0}\frac{1}{\delta^2} G^T_Y + O(\epsilon^2). \tag{A.34}$$

The fixed-$\hat{\alpha}_p$ calculation gives the derivative of the dividend-claim correction as

$$\frac{\partial p^0_{D,2}}{\partial \hat{\alpha}_p} = -\frac{\bar{s}}{\delta^2}\frac{\partial y^0_{Y,2}}{\partial \hat{\alpha}_p} - \frac{B_s}{\delta(\delta + \kappa_s)}\frac{\partial \eta_1}{\partial \hat{\alpha}_p}\sigma_s^\top.$$

The first term is the aggregate-claim channel, while the second term is the dividend-risk-premium channel. Equivalently, define the fixed-$\hat{\alpha}_p$ excess slope

$$G^{\text{ex},0}_D \equiv \frac{\partial p^0_{D,2}}{\partial \hat{\alpha}_p} + \frac{\bar{s}}{\delta^2}\frac{\partial y^0_{Y,2}}{\partial \hat{\alpha}_p} = -\frac{B_s}{\delta(\delta + \kappa_s)}\frac{\partial \eta_1}{\partial \hat{\alpha}_p}\sigma_s^\top. \tag{A.35}$$

This excess slope is independent of $\hat{\alpha}_p$. Therefore, when passive portfolios are time-varying, the same coefficient equation attenuates this slope by the factor $\delta/(\delta + \kappa_p)$, but it does not create an additional M_p term in the dividend-specific component:

$$G^{\text{ex},T}_D = \frac{\delta}{\delta + \kappa_p} G^{\text{ex},0}_D = -\frac{\delta}{\delta + \kappa_p}\frac{B_s}{\delta(\delta + \kappa_s)}\frac{\partial \eta_1}{\partial \hat{\alpha}_p}\sigma_s^\top. \tag{A.36}$$

Combining (A.34) and (A.36), the dividend-claim price impact is

$$\frac{\partial p_D}{\partial F} = \bar{s}\frac{\partial p_Y}{\partial F} - \frac{\epsilon}{x_0}\frac{\delta}{\delta + \kappa_p}\frac{B_s}{\delta(\delta + \kappa_s)}\frac{\partial \eta_1}{\partial \hat{\alpha}_p}\sigma_s^\top + O(\epsilon^2). \tag{A.37}$$

Equation (A.37) is the time-varying-passive-share analogue of the fixed-$\hat{\alpha}_p$ decomposition. There are two effects of portfolio dynamics. First, $\partial p_Y/\partial F$ is itself modified by mean reversion and passive-demand risk through the coefficients b^T_Y and d^T_Y. Second, the dividend-specific risk-premium channel is attenuated by $\delta/(\delta + \kappa_p)$. There is no separate third dividend-specific term in this benchmark specification because the fixed-$\hat{\alpha}_p$ excess slope in (A.35) is constant in $\hat{\alpha}_p$. A separate M_p term would appear only if the fixed-$\hat{\alpha}_p$ excess component had a nonzero quadratic coefficient in $\hat{\alpha}_p$.

B Deriving the aggregate market elasticity

B.1 No preference heterogeneity or leverage constraints

In this section, we consider the special case of no preference heterogeneity or leverage constraints, that is, $\hat{\gamma}_j = 0$ for $j = 0, 1, 2$ and $\hat{\nu}_2 = 0$. This will allow us to derive the cleaner expressions for all equilibrium objects. For simplicity, we focus on the case where $\hat{\alpha} = 0$.

The next proposition summarizes the results for this special case. It encompasses the results of Proposition 4 and Proposition 5.

Proposition 8 (Time-varying passive demand without heterogeneity or constraints)**.** *Consider the case with no preference heterogeneity and no leverage constraints, that is, $\hat{\gamma}_j = 0$ for $j = 0, 1, 2$ and $\hat{\nu}_2 = 0$. Let*

$$M_p \equiv (1-\gamma)\sigma_p\sigma^\top, \qquad d_Y^T \equiv \frac{1}{1+2\kappa_p/y_{Y,0}}\frac{1}{2}(1-\psi^{-1})\gamma\|\sigma\|^2\frac{x_0}{x_a}.$$

Then the second-order correction to the dividend yield is

$$y_{Y,2}^T = \frac{2d_Y^T}{y_{Y,0}}\left[\frac{M_p^2}{y_{Y,0}+\kappa_p} + \frac{\sigma_p\sigma_p^\top}{2} + \frac{y_{Y,0}}{y_{Y,0}+\kappa_p}M_p\hat{\alpha}_p + \frac{y_{Y,0}}{2}\hat{\alpha}_p^2\right].$$

The inverse aggregate market elasticity is

$$\varepsilon_{Y,t}^{-1} = \left(1-\psi^{-1}\right)\frac{\gamma\|\sigma\|^2}{y_{Y,0}+2\kappa_p}\frac{1}{x_a}\left[1-\overline{\alpha}_p + \frac{(\gamma-1)\sigma_p\epsilon\sigma^\top}{y_{Y,0}+\kappa_p}\right] + O(\epsilon^2).$$

The return volatility satisfies

$$\sigma_{R_Y} = \sigma + x_0\epsilon_M^{-1}\sigma_p\epsilon + O(\epsilon^3).$$

The market price of risk is

$$\eta = \gamma\sigma\left(\frac{1-x_0\overline{\alpha}_p}{x_a}\right) + \left[1-H_a\right]\gamma x_0\epsilon_M^{-1}\sigma_p\epsilon + O(\epsilon^3),$$

where

$$H_a \equiv \frac{1-\gamma^{-1}}{\psi-1}\left(\frac{\psi}{x_a}-1\right).$$

The risk premium is

$$\pi_Y = \gamma\|\sigma\|^2\left(\frac{1-x_0\overline{\alpha}_p}{x_a}\right) + (2-H_a)\gamma x_0\epsilon_M^{-1}\sigma\sigma_p^\top\epsilon + O(\epsilon^3).$$

The incremental interest-rate correction is

$$\Delta r_2 = (1 - H_a)\gamma \frac{G_p}{y_{Y,0}} \sigma \sigma_p^\top,$$

where

$$G_p \equiv \frac{\partial y_{Y,2}^T}{\partial \hat{\alpha}_p} = 2 d_Y^T \left(\hat{\alpha}_p + \frac{M_p}{y_{Y,0} + \kappa_p} \right).$$

Finally, the normalized price impact of the dividend claim satisfies

$$\frac{1}{p_D} \frac{\partial p_D}{\partial F} = \frac{1}{p_Y} \frac{\partial p_Y}{\partial F} + \frac{\delta \gamma (1 - \bar{s})}{(\delta + \kappa_p)(\delta + \kappa_s)} \frac{\epsilon}{x_a} \sigma_s \sigma^\top + O(\epsilon^2).$$

Proof. To derive the results, we rely on the general coefficient formulas derived in Appendix A.3.

Dividend yield. In the absence of preference heterogeneity and leverage constraints,

$$a_Y^0 = b_Y^0 = 0, \qquad d_Y^0 = \frac{1}{2}(1 - \psi^{-1})\gamma \|\sigma\|^2 \frac{x_0}{x_a}.$$

Applying the general coefficient formulas from (A.20)– (A.23), the total second-order dividend-yield correction is

$$y_{Y,2}^T = a_Y^T + b_Y^T \hat{\alpha}_p + d_Y^T \hat{\alpha}_p^2,$$

where

$$d_Y^T = \frac{1}{1 + 2\kappa_p / y_{Y,0}} d_Y^0,$$

$$b_Y^T = \frac{2 d_Y^T M_p}{y_{Y,0} + \kappa_p},$$

$$a_Y^T = \frac{1}{y_{Y,0}} \left(b_Y^T M_p + d_Y^T V_p \right).$$

Thus, substituting the coefficients above,

$$y_{Y,2}^T = \frac{2 d_Y^T}{y_{Y,0}} \left[\frac{M_p^2}{y_{Y,0} + \kappa_p} + \frac{V_p}{2} + \frac{y_{Y,0}}{y_{Y,0} + \kappa_p} M_p \hat{\alpha}_p + \frac{y_{Y,0}}{2} \hat{\alpha}_p^2 \right].$$

Using the expression for d_Y^T, this becomes

$$y_{Y,2}^T = \left(1 - \psi^{-1}\right) \frac{\gamma \|\sigma\|^2}{y_{Y,0}} \frac{x_0}{x_a} \frac{1}{1 + 2\kappa_p / y_{Y,0}} \left[\frac{M_p^2}{y_{Y,0} + \kappa_p} + \frac{V_p}{2} + \frac{y_{Y,0}}{y_{Y,0} + \kappa_p} M_p \hat{\alpha}_p + \frac{y_{Y,0}}{2} \hat{\alpha}_p^2 \right].$$

Aggregate market elasticity. The aggregate market elasticity is given by

$$\epsilon_M^{-1} = \frac{1}{p(X,\epsilon)} \frac{\partial p(X,\epsilon)}{\partial F(X,\epsilon)} = -\frac{1}{y_{Y,0}} \frac{\partial y_{Y,2}^T}{\partial \hat{\alpha}_p} \frac{\epsilon}{x_0} + O(\epsilon^2).$$

Using the closed-form slope

$$G_p(x, \hat{\alpha}_p) \equiv \frac{\partial y_{Y,2}^T}{\partial \hat{\alpha}_p} = 2 d_Y^T \left(\hat{\alpha}_p + \frac{M_p}{y_{Y,0} + \kappa_p} \right),$$

we obtain

$$\epsilon_M^{-1} = -\frac{G_p}{y_{Y,0}} \frac{\epsilon}{x_0} + O(\epsilon^2).$$

Substituting the expression for G_p and using

$$d_Y^T = \frac{1}{1 + 2\kappa_p / y_{Y,0}} d_Y^0, \qquad d_Y^0 = \frac{1}{2}(1 - \psi^{-1}) \gamma \|\sigma\|^2 \frac{x_0}{x_a},$$

gives

$$\epsilon_M^{-1} = -\left(1 - \psi^{-1}\right) \frac{\gamma \|\sigma\|^2}{y_{Y,0}} \frac{1}{x_a} \frac{1}{1 + 2\kappa_p / y_{Y,0}} \left(\hat{\alpha}_p + \frac{M_p}{y_{Y,0} + \kappa_p} \right) \epsilon + O(\epsilon^2).$$

Finally, using

$$\overline{\alpha}_p = 1 + \hat{\alpha}_p \epsilon, \qquad M_p = (1 - \gamma) \sigma_p \sigma^\top,$$

where the second equality follows from the assumption $\hat{\alpha} = 0$, we obtain

$$\epsilon_M^{-1} = \left(1 - \psi^{-1}\right) \frac{\gamma \|\sigma\|^2}{y_{Y,0} + 2\kappa_p} \frac{1}{x_a} \left[1 - \overline{\alpha}_p + \frac{(\gamma - 1) \sigma_p \epsilon \sigma^\top}{y_{Y,0} + \kappa_p} \right] + O(\epsilon^2).$$

Thus, mean reversion in passive demand attenuates the aggregate market elasticity through the factor $(y_{Y,0} + 2\kappa_p)^{-1}$, while passive-demand risk introduces an additional correction proportional to $(\gamma - 1)\sigma_p \sigma^\top$.

Return volatility. The return volatility is given by

$$\sigma_{R_Y} = \sigma - \sigma_{y_Y},$$

where

$$\sigma_{y_Y} = \frac{y_{Y,x,t}}{y_{Y,t}} \sigma_{x,t} + \frac{y_{Y,\overline{\alpha}_p,t}}{y_{Y,t}} \sigma_p = \frac{y_{Y,2,\hat{\alpha}_p}}{y_{Y,0}} \sigma_p \epsilon^2 + O(\epsilon^3).$$

Using the expression for the aggregate market elasticity,

$$\epsilon_M^{-1} = -\frac{1}{y_{Y,0}} \frac{\partial y_{Y,2}^T}{\partial \hat{\alpha}_p} \frac{\epsilon}{x_0},$$

we obtain

$$\sigma_{R_Y} = \sigma + x_0 \epsilon_M^{-1} \sigma_p \epsilon + O(\epsilon^3).$$

Risk exposure. The risk exposure of passive investors is given by

$$\sigma_0 = \overline{\alpha}_p \sigma_{R_Y} = \sigma + \sigma \hat{\alpha}_p \epsilon - \sigma_{y_Y,2} \epsilon^2 + O(\epsilon^3),$$

From the expression for $\sigma_{y_Y,2}$, we have

$$\sigma_0 = \sigma + \hat{\alpha}_p \sigma \epsilon + x_0 \epsilon_M^{-1} \sigma_p \epsilon + O(\epsilon^3),$$

Using the fact that $\sum_{j=0}^{2} x_j \sigma_j = \sigma_{R_Y}$, we can write the risk exposure of active investor $j = 1, 2$ as

$$\sigma_j = \sigma - \frac{x_0 \hat{\alpha}_p}{x_a} \sigma \epsilon + x_0 \epsilon_M^{-1} \sigma_p \epsilon + O(\epsilon^3), \qquad j = 1, 2.$$

We then write the first-order correction to the risk exposure as $\sigma_{j,1} = m_j \sigma$, where

$$m_a = -\frac{x_0 \hat{\alpha}_p}{x_a}, \qquad m_0 = \hat{\alpha}_p,$$

where $m_1 = m_2 = m_a$.

Consumption-wealth ratio. The consumption-wealth ratio of investor $j = 0, 1, 2$ can be written as

$$c_j = c_j^0 + \Delta c_j,$$

where c_j^0 is the fixed-$\hat{\alpha}_p$ consumption-wealth ratio and Δc_j is the time-varying-α_p correction.

The consumption-wealth ratio in the fixed-$\hat{\alpha}_p$ economy is given by

$$c_j^0 = c_{j,0} + c_{j,2}^0 \epsilon^2 + O(\epsilon^3),$$

where

$$c_{j,2}^0 = (1 - \psi) \left[y_{Y,2}^0 + \left(m_j m_a - \frac{1}{2} m_j^2 \right) \gamma \|\sigma\|^2 \right].$$

We can then write $c^0_{j,2} = d^0_j \hat{\alpha}_p^2$, where

$$d^0_a = (1-\psi)\frac{\gamma\|\sigma\|^2}{2}\left[(1-\psi^{-1})\frac{x_0}{x_a} + \left(\frac{x_0}{x_a}\right)^2\right], \qquad d^0_0 = (1-\psi)\gamma\|\sigma\|^2\left[\frac{1-\psi^{-1}}{2}\frac{x_0}{x_a} - \left(\frac{x_0}{x_a}+\frac{1}{2}\right)\right].$$

The individual consumption-wealth ratios solve the same coefficient equation as the aggregate dividend yield, but investor by investor. In the no-heterogeneity, no-constraint case, active investors are identical, so write

$$c^T_{0,2}(x,\hat{\alpha}_p) = a^T_0 + b^T_0\hat{\alpha}_p + d^T_0\hat{\alpha}_p^2, \qquad c^T_{a,2}(x,\hat{\alpha}_p) = a^T_a + b^T_a\hat{\alpha}_p + d^T_a\hat{\alpha}_p^2,$$

where $c^T_{a,2}$ denotes the common correction for active investors $j = 1, 2$. The coefficients are obtained from the fixed-$\hat{\alpha}_p$ coefficients using

$$\begin{aligned} d^T_j &= \frac{1}{1+2\kappa_p/y_{Y,0}}d^0_j, \\ b^T_j &= \frac{2d^T_j M_p}{y_{Y,0}+\kappa_p}, \\ a^T_j &= \frac{1}{y_{Y,0}}\left(b^T_j M_p + d^T_j V_p\right), \qquad j \in \{0,a\}, \end{aligned}$$

where we use $c_{j,0} = y_{Y,0}$ in the homogeneous benchmark economy.

Consumption-wealth ratio drift and diffusion. The slope of each individual consumption-wealth ratio with respect to the passive portfolio share is

$$G_j(x,\hat{\alpha}_p) \equiv \frac{\partial c^T_{j,2}}{\partial\hat{\alpha}_p} = 2d^T_j\left(\hat{\alpha}_p + \frac{M_p}{y_{Y,0}+\kappa_p}\right), \qquad j \in \{0,a\}.$$

Hence the diffusion of the consumption-wealth ratio of investor j is

$$\sigma_{c_j,t} = \frac{G_j}{y_{Y,0}}\sigma_p\epsilon^2 + O(\epsilon^3), \qquad j \in \{0,a\}.$$

The drift of the consumption-wealth ratio is computed in the same way. Since

$$c^T_{j,2,\hat{\alpha}_p} = G_j, \qquad c^T_{j,2,\hat{\alpha}_p\hat{\alpha}_p} = 2d^T_j,$$

we have

$$\mu_{c_j,t} = \left[\frac{G_j}{y_{Y,0}} \kappa_p (\hat{\alpha} - \hat{\alpha}_p) + \frac{d_j^T}{y_{Y,0}} V_p \right] \epsilon^2 + O(\epsilon^3), \qquad j \in \{0, a\}.$$

Under the maintained normalization $\hat{\alpha} = 0$, this becomes

$$\mu_{c_j,t} = \left[-\frac{G_j}{y_{Y,0}} \kappa_p \hat{\alpha}_p + \frac{d_j^T}{y_{Y,0}} V_p \right] \epsilon^2 + O(\epsilon^3), \qquad j \in \{0, a\}.$$

By contrast, the diffusion of the aggregate dividend yield is

$$\sigma_{y_Y,t} = \frac{G_p}{y_{Y,0}} \sigma_p \epsilon^2 + O(\epsilon^3), \qquad G_p = x_0 G_0 + x_a G_a.$$

Thus, aggregation implies

$$x_0 \sigma_{c_0,t} + x_a \sigma_{c_a,t} = \sigma_{y_Y,t},$$

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

$$\eta_t = \eta_0 + \eta_1 \epsilon + \eta_2 \epsilon^2 + O(\epsilon^3).$$

At the benchmark,

$$\eta_0 = \gamma \sigma.$$

At first order, active investors absorb the residual portfolio exposure created by passive investors. Since

$$m_a = -\frac{x_0}{x_a} \hat{\alpha}_p,$$

we have

$$\eta_1 = \gamma m_a \sigma = -\gamma \sigma \frac{x_0}{x_a} \hat{\alpha}_p.$$

We next derive the second-order term. Because active investors are unconstrained and identical, their optimal exposure satisfies

$$\sigma_{a,t} = \frac{\eta_t}{\gamma} + \frac{1 - \gamma^{-1}}{\psi - 1} \sigma_{c_a,t}.$$

Therefore,

$$\eta_2 = \gamma \left[\sigma_{a,2} - \frac{1 - \gamma^{-1}}{\psi - 1} \sigma_{c_a,2} \right]. \tag{B.1}$$

where

$$\sigma_{a,2} = -\sigma_{y_Y,2}.$$

Substituting into (B.1) gives

$$\eta_2 = -\gamma \left[\sigma_{y_Y,2} + \frac{1-\gamma^{-1}}{\psi - 1} \sigma_{c_a,2} \right].$$

Using the diffusion formulas above,

$$\sigma_{y_Y,2} = \frac{G_p}{y_{Y,0}} \sigma_p, \qquad \sigma_{c_a,2} = \frac{G_a}{y_{Y,0}} \sigma_p,$$

we obtain the closed-form expression

$$\eta_2 = -\frac{\gamma}{y_{Y,0}} \left[G_p + \frac{1-\gamma^{-1}}{\psi - 1} G_a \right] \sigma_p. \tag{B.2}$$

Equivalently, using the explicit slopes

$$G_p = 2 d_Y^T \left(\hat{\alpha}_p + \frac{M_p}{y_{Y,0} + \kappa_p} \right), \qquad G_a = 2 d_a^T \left(\hat{\alpha}_p + \frac{M_p}{y_{Y,0} + \kappa_p} \right),$$

we can write

$$\eta_2 = -\frac{2\gamma}{y_{Y,0}} \left[d_Y^T + \frac{1-\gamma^{-1}}{\psi - 1} d_a^T \right] \left(\hat{\alpha}_p + \frac{M_p}{y_{Y,0} + \kappa_p} \right) \sigma_p.$$

It is useful to express this result in terms of the aggregate market elasticity. Because d_a^T and d_Y^T are attenuated by the same factor, we have

$$\frac{G_a}{G_p} = \frac{d_a^T}{d_Y^T} = \frac{d_a^0}{d_Y^0}.$$

Using the expressions for d_a^0 and d_Y^0, this ratio is

$$\chi_a \equiv \frac{G_a}{G_p} = \frac{d_a^0}{d_Y^0} = (1 - \psi) \left[1 + \frac{x_0/x_a}{1 - \psi^{-1}} \right].$$

Therefore, (B.2) can be written as

$$\eta_2 = -\frac{\gamma G_p}{y_{Y,0}} \left[1 + \frac{1-\gamma^{-1}}{\psi - 1} \chi_a \right] \sigma_p.$$

Using the expression for χ_a,

$$\frac{1-\gamma^{-1}}{\psi - 1} \chi_a = -\frac{1-\gamma^{-1}}{1-\psi^{-1}} \left(\frac{1}{x_a} - \psi^{-1} \right)$$

Finally, using the aggregate market elasticity relation

$$\epsilon_M^{-1} = -\frac{G_p}{y_{Y,0}}\frac{\epsilon}{x_0} + O(\epsilon^2),$$

we obtain

$$\eta = \gamma\sigma\left(\frac{1 - x_0\overline{\alpha}_p}{x_a}\right) + \left[1 - \frac{1-\gamma^{-1}}{1-\psi^{-1}}\left(\frac{1}{x_a} - \psi^{-1}\right)\right]\gamma x_0 \epsilon_M^{-1}\sigma_p\epsilon + O(\epsilon^3).$$

This expression shows that the second-order market price of risk is proportional to the aggregate market elasticity, with an adjustment for the active investors' hedging demand.

Risk premium. The risk premium on the aggregate claim is

$$\pi_Y = \sigma_{R_Y}\eta^\top.$$

Using the expressions derived above,

$$\sigma_{R_Y} = \sigma + x_0\epsilon_M^{-1}\sigma_p\epsilon + O(\epsilon^3),$$

and

$$\eta = \gamma\sigma\left(\frac{1 - x_0\overline{\alpha}_p}{x_a}\right) + (1 - H_a)\gamma x_0\epsilon_M^{-1}\sigma_p\epsilon + O(\epsilon^3),$$

where

$$H_a \equiv \frac{1-\gamma^{-1}}{1-\psi^{-1}}\left(\frac{1}{x_a} - \psi^{-1}\right).$$

Therefore,

$$\pi_Y = \gamma\|\sigma\|^2\left(\frac{1 - x_0\overline{\alpha}_p}{x_a}\right) + x_0\epsilon_M^{-1}\epsilon\left[2 - H_a\right]\gamma\sigma\sigma_p^\top + O(\epsilon^3).$$

The term H_a captures the hedging-demand response of active investors. Under the usual assumptions $\gamma > 1$ and $\psi > 1$, $H_a > 0$, so hedging demand dampens the effect of passive-demand risk on the aggregate risk premium.

Interest rate. The interest rate follows from

$$r = y_Y - \pi_Y - \mu_{y_Y} - \sigma\sigma_{y_Y}^\top.$$

Using the expressions above,

$$r = y_{Y,0} - \gamma\|\sigma\|^2\left(\frac{1 - x_0\overline{\alpha}_p}{x_a}\right) + y_{Y,2}^T\epsilon^2 - \mu_{y_Y} - \left[(2 - H_a)\gamma - 1\right]x_0\epsilon_M^{-1}\sigma\sigma_p^\top\epsilon + O(\epsilon^3),$$

where

$$\mu_{y_Y} = \left[-\frac{G_p}{y_{Y,0}}\kappa_p\hat{\alpha}_p + \frac{d_Y^T}{y_{Y,0}}V_p\right]\epsilon^2 + O(\epsilon^3).$$

Equivalently, isolating only the incremental effect of time-varying passive demand, Step 5 gives

$$\Delta r_2 = -\sigma\Delta\eta_2^\top.$$

Since

$$\Delta\eta_2\epsilon^2 = (1 - H_a)\gamma x_0\epsilon_M^{-1}\sigma_p\epsilon + O(\epsilon^3),$$

we obtain

$$\Delta r_2\epsilon^2 = -(1 - H_a)\gamma x_0\epsilon_M^{-1}\sigma\sigma_p^\top\epsilon + O(\epsilon^3).$$

Equivalently,

$$\Delta r_2 = (1 - H_a)\gamma\frac{G_p}{y_{Y,0}}\sigma\sigma_p^\top.$$

Price of dividend claim. We now price the dividend claim, whose payoff is $D_t = s_tY_t$. Write

$$s_t = \bar{s} + \epsilon\hat{s}_t,$$

and conjecture the expansion

$$p_D(x, s, \overline{\alpha}_p; \epsilon) = p_{D,0} + p_{D,1}(\hat{s})\epsilon + p_{D,2}^T(\hat{s}, \hat{\alpha}_p)\epsilon^2 + O(\epsilon^3).$$

At the benchmark,

$$p_{D,0} = \frac{\bar{s}}{\delta}, \qquad \delta \equiv y_{Y,0}.$$

The first-order term is the present value of the transitory dividend-share component:

$$p_{D,1}(\hat{s}) = \frac{\hat{s}}{\delta + \kappa_s}.$$

To obtain the second-order term, it is useful to decompose the dividend claim into an aggregate-claim component and a dividend-specific component. The aggregate component is simply $\bar{s}$ times

the aggregate claim. Because

$$p_Y = \frac{1}{y_Y} = \frac{1}{\delta} - \frac{y^T_{Y,2}}{\delta^2}\epsilon^2 + O(\epsilon^3),$$

this component contributes

$$-\frac{\bar{s}}{\delta^2} y^T_{Y,2}$$

to $p^T_{D,2}$.

The dividend-specific component comes from the risk premium on the dividend-share shock. The fixed-$\hat{\alpha}_p$ calculation gives

$$p^0_{D,2}(\hat{s}, \hat{\alpha}_p) = -\frac{\bar{s}}{\delta^2} y^0_{Y,2}(\hat{\alpha}_p) - \frac{B_s}{\delta(\delta + \kappa_s)} \eta_1 \sigma_s^\top + b^0_2 \hat{s},$$

where $B_s \equiv \bar{s}(1 - \bar{s})$ and b^0_2 is independent of $\hat{\alpha}_p$. In the present special case,

$$\eta_1 = -\gamma\sigma \frac{x_0}{x_a} \hat{\alpha}_p,$$

so the fixed-$\hat{\alpha}_p$ excess component is linear in $\hat{\alpha}_p$:

$$e^0_D(\hat{\alpha}_p) \equiv -\frac{B_s}{\delta(\delta + \kappa_s)} \eta_1 \sigma_s^\top = \frac{\gamma B_s}{\delta(\delta + \kappa_s)} \frac{x_0}{x_a} \sigma\sigma_s^\top \hat{\alpha}_p.$$

With time-varying passive portfolios, this excess component solves the same scalar coefficient equation as the aggregate dividend-yield correction. Because e^0_D is linear in $\hat{\alpha}_p$, there is no quadratic term and hence no new M_p term in its slope. Mean reversion attenuates the linear coefficient by $\delta/(\delta + \kappa_p)$. Thus the total excess component is

$$e^T_D(\hat{\alpha}_p) = \frac{\delta}{\delta + \kappa_p} e^0_D(\hat{\alpha}_p) + \frac{1}{\delta + \kappa_p} \frac{\gamma B_s}{\delta(\delta + \kappa_s)} \frac{x_0}{x_a} \sigma\sigma_s^\top M_p,$$

where the second term is the constant term generated by the drift and risk-adjustment operator. Equivalently,

$$e^T_D(\hat{\alpha}_p) = \frac{\gamma B_s}{(\delta + \kappa_p)(\delta + \kappa_s)} \frac{x_0}{x_a} \sigma\sigma_s^\top \left(\hat{\alpha}_p + \frac{M_p}{\delta} \right).$$

Combining the aggregate and dividend-specific components, the second-order price of the dividend claim is

$$p^T_{D,2}(\hat{s}, \hat{\alpha}_p) = -\frac{\bar{s}}{\delta^2} y^T_{Y,2}(\hat{\alpha}_p) + e^T_D(\hat{\alpha}_p) + b^0_2 \hat{s}.$$

Therefore,

$$p_D(x, s, \overline{\alpha}_p; \epsilon) = \frac{\bar{s}}{\delta} + \frac{\hat{s}}{\delta + \kappa_s}\epsilon + \left[-\frac{\bar{s}}{\delta^2} y^T_{Y,2}(\hat{\alpha}_p) + e^T_D(\hat{\alpha}_p) + b^0_2 \hat{s}\right] \epsilon^2 + O(\epsilon^3).$$

Differentiating with respect to passive flows gives the price impact of passive demand on the dividend claim:

$$\begin{aligned}\frac{\partial p_D}{\partial F} &= \bar{s}\frac{\partial p_Y}{\partial F} + \frac{\epsilon}{x_0}\frac{\partial e^T_D}{\partial \hat{\alpha}_p} + O(\epsilon^2) \\ &= \bar{s}\frac{\partial p_Y}{\partial F} + \frac{\epsilon}{x_0}\frac{\gamma B_s}{(\delta + \kappa_p)(\delta + \kappa_s)}\frac{x_0}{x_a}\sigma\sigma_s^\top + O(\epsilon^2).\end{aligned}$$

Equivalently, using $p_{D,0} = \bar{s}/\delta$ and $p_{Y,0} = 1/\delta$, the normalized price impact of the dividend claim can be related to the normalized price impact of the aggregate claim as

$$\frac{1}{p_D}\frac{\partial p_D}{\partial F} = \frac{1}{p_Y}\frac{\partial p_Y}{\partial F} + \frac{\delta}{\bar{s}}\frac{\epsilon}{x_0}\frac{\gamma B_s}{(\delta + \kappa_p)(\delta + \kappa_s)}\frac{x_0}{x_a}\sigma\sigma_s^\top + O(\epsilon^2).$$

Since $B_s = \bar{s}(1 - \bar{s})$, this simplifies to

$$\frac{1}{p_D}\frac{\partial p_D}{\partial F} = \frac{1}{p_Y}\frac{\partial p_Y}{\partial F} + \frac{1}{\bar{s}}\frac{\delta\gamma\bar{s}(1 - \bar{s})\sigma_s\sigma^\top}{(\delta + \kappa_p)(\delta + \kappa_s)}\frac{\epsilon}{x_a} + O(\epsilon^2).$$

Thus, the dividend-claim price impact is the sum of two terms. The first term is the mechanical exposure to the aggregate claim. The second term is the dividend-specific risk-premium channel, attenuated by passive-demand mean reversion through the factor $(\delta + \kappa_p)^{-1}$. □

B.2 Preference heterogeneity and leverage constraints

We consider next the aggregate market elasticity in the case with preference heterogeneity and leverage constraints.

First, consider the case with no preference heterogeneity and no leverage constraints, that is, $\hat{\gamma}_j = 0$ for $j = 0, 1, 2$ and $\hat{\nu}_2 = 0$. In this case, the dividend yield is given by

$$y^0_{Y,2} = a^0_y + b^0_y\hat{\alpha}_p + d^0_y\hat{\alpha}^2_p.$$

The inverse aggregate market elasticity is given by

$$(\epsilon^0_{Y,t})^{-1} = -\frac{1}{x_{0,t} y_{Y,0}} \frac{\partial y^0_{Y,2}}{\partial \hat{\alpha}_p} \epsilon + O(\epsilon^2).$$
$$= -\frac{1}{x_{0,t} y_{Y,0}} \left(2 d^0_y \hat{\alpha}_p + b^0_y \right) \epsilon + O(\epsilon^2).$$

In the case with preference heterogeneity and leverage constraints, the inverse aggregate market elasticity is given by

$$(\epsilon^0_{Y,t})^{-1} = (1 - \psi^{-1}) \frac{\gamma \|\sigma\|^2}{y_{Y,0}} \frac{\overline{\alpha}^{opt}_p - 1}{x_a - x_c} (1 - x_c) + (1 - \psi^{-1}) \frac{\gamma \|\sigma\|^2}{y_{Y,0}} \frac{1 - \overline{\alpha}_p}{x_a - x_c} (1 - x_c) + O(\epsilon^2).$$

This implies that the coefficients are given by

$$\frac{2 d^0_Y}{x_{0,t} y_{Y,0}} = (1 - \psi^{-1}) \frac{\gamma \|\sigma\|^2}{y_{Y,0}} \frac{(1 - x_c)}{x_a - x_c},$$
$$-\frac{b^0_Y}{x_{0,t} y_{Y,0}} \epsilon = \frac{2 d^0_Y}{x_{0,t} y_{Y,0}} (\overline{\alpha}^{opt}_p - 1).$$

The second-order correction for the dividend yield in the case with passive portfolio dynamics is given by

$$y^T_{Y,2} = a^T_y + b^T_y \hat{\alpha}_p + d^T_y \hat{\alpha}^2_p,$$

where

$$d^T_Y = \frac{y_{Y,0}}{y_{Y,0} + 2\kappa_p} d^0_Y,$$
$$b^T_Y = \frac{y_{Y,0}}{y_{Y,0} + \kappa_p} b^0_Y + \frac{2 d^T_Y}{y_{Y,0} + \kappa_p} M_p,$$

where $M_p = \kappa_p \hat{\alpha} + (1 - \gamma) \sigma_p \sigma^\top$.

The inverse aggregate market elasticity is given by

$$(\epsilon^T_{Y,t})^{-1} = -\frac{1}{x_{0,t} y_{Y,0}} \left(2 d^T_Y \hat{\alpha}_p + b^T_Y \right) \epsilon + O(\epsilon^2).$$

The coefficient on $1 - \overline{\alpha}_p$ can be written as

$$\frac{2 d^T_Y}{x_{0,t} y_{Y,0}} = \frac{y_{Y,0}}{y_{Y,0} + 2\kappa_p} \frac{2 d^0_Y}{x_{0,t} y_{Y,0}} = (1 - \psi^{-1}) \frac{\gamma \|\sigma\|^2}{y_{Y,0} + 2\kappa_p} \frac{(1 - x_c)}{x_a - x_c},$$

and the intercept term is given by

$$\begin{aligned}
-\frac{b_Y^T}{x_{0,t}y_{Y,0}} &= -\frac{y_{Y,0}}{y_{Y,0}+\kappa_p}\frac{b_Y^0}{x_{0,t}y_{Y,0}} - \frac{M_p}{y_{Y,0}+\kappa_p}\frac{2d_Y^T}{x_{0,t}y_{Y,0}} \\
&= \frac{y_{Y,0}}{y_{Y,0}+\kappa_p}\frac{2d_Y^0}{x_{0,t}y_{Y,0}}\hat{\alpha}_p^{opt} - \frac{M_p}{y_{Y,0}+\kappa_p}\frac{y_{Y,0}}{y_{Y,0}+2\kappa_p}\frac{2d_Y^0}{x_{0,t}y_{Y,0}} \\
&= \frac{y_{Y,0}}{y_{Y,0}+2\kappa_p}\frac{2d_Y^0}{x_{0,t}y_{Y,0}}\left[\frac{y_{Y,0}+2\kappa_p}{y_{Y,0}+\kappa_p}\hat{\alpha}_p^{opt} - \frac{M_p}{y_{Y,0}+\kappa_p}\right],
\end{aligned}$$

where $\overline{\alpha}_p^{opt}\epsilon \equiv \overline{\alpha}_p^{opt} - 1$.

Combining the expressions above, we obtain

$$(\epsilon_{Y,t}^T)^{-1} = (1-\psi^{-1})\frac{\gamma\|\sigma\|^2}{y_{Y,0}+2\kappa_p}\frac{(1-x_c)}{x_a-x_c}\left[\overline{\alpha}_{p,t}^{opt} - \overline{\alpha}_{p,t} + \frac{\kappa_p(\overline{\alpha}_{p,t}^{opt}-\overline{\alpha})}{y_{Y,0}+\kappa_p} + \frac{(\gamma-1)\sigma_p\sigma^\top}{y_{Y,0}+\kappa_p}\right] + O(\epsilon^2),$$

using the fact that $\overline{\alpha} = 1 + \hat{\alpha}\epsilon$.

B.3 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_0 denote the wealth share of passive investors and $x_a = 1 - x_0$ the wealth share of active investors. Because passive investors hold the portfolio share $\overline{\alpha}_p$, their wealth-share diffusion is

$$\sigma_{x_0,t} = x_{0,t}\left(\sigma_{0,t} - \sigma_{R_Y,t}\right) = x_{0,t}\left(\overline{\alpha}_{p,t} - 1\right)\sigma_{R_Y,t}.$$

Using $\overline{\alpha}_{p,t} = 1 + \hat{\alpha}_{p,t}\epsilon$ and $\sigma_{R_Y,t} = \sigma + O(\epsilon^2)$, this gives

$$\sigma_{x_0,t} = x_{0,t}\hat{\alpha}_{p,t}\sigma\epsilon + O(\epsilon^3). \tag{B.3}$$

Thus, the wealth-distribution state moves in response to the same aggregate return shock that redistributes wealth between passive and active investors.

The aggregate dividend yield satisfies

$$y_Y(x_0, \hat{\alpha}_p; \epsilon) = y_{Y,0} + y_{Y,2}^T(x_0, \hat{\alpha}_p)\epsilon^2 + O(\epsilon^3).$$

Therefore, allowing both x_0 and $\hat{\alpha}_p$ to move, its diffusion is

$$\sigma_{y_Y,t} = \frac{1}{y_{Y,0}} \left[y^T_{Y,2,\hat{\alpha}_p} \sigma_p + y^T_{Y,2,x_0} \sigma_{x_0,t} \right] \epsilon^2 + O(\epsilon^4).$$

The first term is the direct portfolio-flow channel studied above. The second term is the wealth-distribution channel. Using (B.3),

$$\sigma_{y_Y,t} = \frac{G_p}{y_{Y,0}} \sigma_p \epsilon^2 + \frac{x_0 \hat{\alpha}_p}{y_{Y,0}} y^T_{Y,2,x_0} \sigma \epsilon^3 + O(\epsilon^4),$$

where

$$G_p \equiv y^T_{Y,2,\hat{\alpha}_p}.$$

Consequently, the price-dividend-ratio diffusion induced by a portfolio-flow shock is

$$\sigma_{p_Y,t} = -\sigma_{y_Y,t} = -\frac{G_p}{y_{Y,0}} \sigma_p \epsilon^2 - \frac{x_0 \hat{\alpha}_p}{y_{Y,0}} y^T_{Y,2,x_0} \sigma \epsilon^3 + O(\epsilon^4).$$

Using the aggregate market elasticity relation

$$\epsilon_M^{-1} = -\frac{G_p}{y_{Y,0}} \frac{\epsilon}{x_0} + O(\epsilon^2),$$

we can write

$$\sigma_{p_Y,t} = x_0 \epsilon_M^{-1} \sigma_p \epsilon - \frac{x_0 \hat{\alpha}_p}{y_{Y,0}} y^T_{Y,2,x_0} \sigma \epsilon^3 + O(\epsilon^4). \tag{B.4}$$

Thus, to the order at which the aggregate market elasticity is computed, the equilibrium price response to a portfolio-flow shock is the direct partial-equilibrium effect holding the wealth distribution fixed. The endogenous movement in the wealth distribution contributes only at the next order.

In the closed-form no-heterogeneity solution, $y^T_{Y,2}$ depends on the wealth distribution only through x_0/x_a. Hence

$$y^T_{Y,2,x_0} = \frac{y^T_{Y,2}}{x_0 x_a}. \tag{B.5}$$

Substituting (B.5) into (B.4) gives the explicit wealth-distribution correction

$$\sigma_{p_Y,t} = x_0 \epsilon_M^{-1} \sigma_p \epsilon - \frac{\hat{\alpha}_p}{x_a} \frac{y^T_{Y,2}}{y_{Y,0}} \sigma \epsilon^3 + O(\epsilon^4). \tag{B.6}$$

Equation (B.6) separates the two effects of a portfolio-flow shock. The first term is the direct demand effect captured by the aggregate market elasticity. The second term is the equilibrium

wealth-distribution effect: the price response to the change in x_0 caused by the shock itself. Because this second term is of order ϵ^3 in price diffusion, it affects the flow-price impact only at order ϵ^2 and does not change the first-order aggregate elasticity formula above.

C Derivation of the Market Elasticity

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

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

Let

$$F(X,\epsilon) \equiv \omega_0 \frac{W_0(1+\hat{\alpha}_p\epsilon) - W_0}{P} = \hat{\alpha}_p \epsilon x_0$$

denote the flow into the risky asset relative to the benchmark economy. Note that $F = O(\epsilon)$.

In this section, we consider the case where the portfolio of passive investors is constant.

Proof. Using (C.1), the derivative of the price-dividend ratio with respect to flows can be written as

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

From Proposition 3, $y_1(X)$ does not depend on $\hat{\alpha}_p$, so

$$\frac{\partial y_1(X)}{\partial \hat{\alpha}_p} = 0.$$

Therefore,

$$\frac{\partial p(X,\epsilon)}{\partial F(X,\epsilon)} = 0 + O(\epsilon),$$

which implies

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

Thus, the aggregate market elasticity is infinite to first order.

To understand this result, note that from the pricing condition

$$y_t = r_t + \pi_t - \mu_{P,t},$$

we can write the first-order term as

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

Since $y_1(X)$ is constant, we must have

$$\mu_{y,1}(X) = 0, \qquad \sigma_{y,1}(X) = 0,$$

which implies $\sigma_{R,1}(X) = 0$ and therefore $\mu_{P,1}(X) = 0$.

Hence,

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

From Proposition 3,

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

These two effects exactly cancel. Thus, at first order, portfolio flows raise the risk premium but lower the interest rate by the same amount, leaving prices unchanged.

We now turn to the second-order effect. From (C.1), we have

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

Because $F = O(\epsilon)$, second-order changes in the dividend yield generate a first-order response in prices.

At second order, the cancellation between interest rates and risk premia breaks down due to endogenous reallocation of risk across investors. In particular, from the expression for $y_2(X)$,

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

where $\sigma_{j,1} = m_{j,1}\sigma$.

Thus, $\partial y_2 / \partial \hat{\alpha}_p \neq 0$, implying a non-zero price impact and a finite aggregate elasticity.

□

C.1 Proof of Proposition 4

Proof. Consider the case with no preference heterogeneity and no leverage constraints. That is, $\hat{\gamma}_j = 0$ for $j = 0, 1, 2$ and $\nu_{2,t} = 0$. In this case, the first-order exposure coefficients are

$$m_{0,1} = \hat{\alpha}_p, \qquad m_{j,1} = -\frac{x_0}{x_a}\hat{\alpha}_p, \qquad j = 1, 2,$$

where $x_a \equiv 1 - x_0$ is the wealth share of active investors and $\sigma_{j,1} = m_{j,1}\sigma$. The second-order correction to the dividend yield therefore simplifies to

$$\begin{aligned} y_{Y,2}(X) &= (1-\psi^{-1})\gamma\|\sigma\|^2 \sum_{j=0}^{2} x_j \frac{m_{j,1}^2}{2} \\ &= (1-\psi^{-1})\gamma\|\sigma\|^2 \frac{1}{2}\left[x_0\hat{\alpha}_p^2 + x_a\left(\frac{x_0}{x_a}\hat{\alpha}_p\right)^2\right] \\ &= (1-\psi^{-1})\gamma\|\sigma\|^2 \frac{x_0}{2x_a}\hat{\alpha}_p^2. \end{aligned}$$

Hence,

$$\frac{\partial y_{Y,2}(X)}{\partial \hat{\alpha}_p} = (1-\psi^{-1})\gamma\|\sigma\|^2 \frac{x_0}{x_a}\hat{\alpha}_p.$$

Using the change of variables from the previous subsection,

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

Substituting the derivative above gives

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

Finally, since $\overline{\alpha}_p = 1 + \hat{\alpha}_p\epsilon$, we obtain

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

This proves the result. □

C.2 Proof of Proposition 6

Proof. Consider the case in which investors have heterogeneous risk aversions but the leverage constraints are slack. Then $\hat{\nu}_{j,1} = 0$ for $j = 1, 2$. Let

$$\bar{\gamma}_a \equiv \sum_{j=1}^{2} \frac{x_j}{x_a} \hat{\gamma}_j, \qquad x_a \equiv x_1 + x_2 = 1 - x_0.$$

From the first-order solution, the scalar exposure coefficients defined by $\sigma_{j,1} = m_{j,1}\sigma$ are

$$m_{0,1} = \hat{\alpha}_p, \qquad m_{j,1} = \bar{\gamma}_a - \frac{x_0}{x_a}\hat{\alpha}_p - \hat{\gamma}_j, \qquad j = 1, 2.$$

The second-order correction to the dividend yield is

$$y_{Y,2} = (1 - \psi^{-1})\gamma \|\sigma\|^2 \sum_{j=0}^{2} x_j \left(\frac{1}{2} m_{j,1}^2 + \hat{\gamma}_j m_{j,1} \right).$$

We differentiate this expression with respect to $\hat{\alpha}_p$. Since

$$\frac{\partial m_{0,1}}{\partial \hat{\alpha}_p} = 1, \qquad \frac{\partial m_{j,1}}{\partial \hat{\alpha}_p} = -\frac{x_0}{x_a}, \qquad j = 1, 2,$$

we obtain

$$\begin{aligned} \frac{\partial y_{Y,2}}{\partial \hat{\alpha}_p} &= (1 - \psi^{-1})\gamma \|\sigma\|^2 \sum_{j=0}^{2} x_j (m_{j,1} + \hat{\gamma}_j) \frac{\partial m_{j,1}}{\partial \hat{\alpha}_p} \\ &= (1 - \psi^{-1})\gamma \|\sigma\|^2 x_0 \left[\hat{\alpha}_p + \hat{\gamma}_0 - \sum_{j=1}^{2} \frac{x_j}{x_a} (m_{j,1} + \hat{\gamma}_j) \right]. \end{aligned}$$

For active investors,

$$m_{j,1} + \hat{\gamma}_j = \bar{\gamma}_a - \frac{x_0}{x_a}\hat{\alpha}_p.$$

Therefore,

$$\frac{\partial y_{Y,2}}{\partial \hat{\alpha}_p} = (1 - \psi^{-1})\gamma \|\sigma\|^2 x_0 \left[\frac{\hat{\alpha}_p}{x_a} + \hat{\gamma}_0 - \bar{\gamma}_a \right].$$

Using the change of variables from the market-elasticity derivation,

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

Substituting the derivative above gives

$$\frac{1}{p(X,\epsilon)}\frac{\partial p(X,\epsilon)}{\partial F(X,\epsilon)} = -(1-\psi^{-1})\frac{\gamma\|\sigma\|^2}{y_0(X)}\left[\frac{\hat{\alpha}_p\epsilon}{x_a} + (\hat{\gamma}_0 - \bar{\gamma}_a)\epsilon\right] + O(\epsilon^2).$$

Finally, using $\overline{\alpha}_p = 1 + \hat{\alpha}_p\epsilon$ and $\gamma_j = \gamma(1+\hat{\gamma}_j\epsilon)$, we can write the expression in primitives as

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

where

$$\mathbb{E}_a[\gamma_j] \equiv \sum_{j=1}^{2}\frac{x_j}{x_a}\gamma_j.$$

Equivalently, this can be written in the form of Proposition 6 by setting $x_c = 0$, $x_u = x_a$, and

$$\overline{\alpha}_p^{\text{opt}} = 1 - x_a\frac{\gamma_0 - \mathbb{E}_a[\gamma_j]}{\gamma} = 1 + \left(\sum_{j=0}^{2} x_k\hat{\gamma}_k - \hat{\gamma}_0\right)\epsilon,$$

where the expression above coincides with first-order correction for the passive investor's optimal portfolio share, that is, the portfolio they would choose if they could freely adjust their portfolio share.

Then

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

This proves the version of the result without leverage constraints.

We next derive the corresponding expression when investor 2 is constrained and investor 1 is unconstrained. In this case, the first-order exposure coefficients are

$$m_{0,1} = \hat{\alpha}_p, \qquad m_{2,1} = \hat{\sigma}, \qquad m_{1,1} = -\frac{x_0\hat{\alpha}_p + x_2\hat{\sigma}}{x_1},$$

where the expression for $m_{1,1}$ follows from market clearing,

$$x_0 m_{0,1} + x_1 m_{1,1} + x_2 m_{2,1} = 0.$$

Using again

$$y_{Y,2} = (1-\psi^{-1})\gamma\|\sigma\|^2\sum_{j=0}^{2} x_j\left(\frac{1}{2}m_{j,1}^2 + \hat{\gamma}_j m_{j,1}\right),$$

we differentiate with respect to $\hat{\alpha}_p$. Since

$$\frac{\partial m_{0,1}}{\partial \hat{\alpha}_p} = 1, \qquad \frac{\partial m_{1,1}}{\partial \hat{\alpha}_p} = -\frac{x_0}{x_1}, \qquad \frac{\partial m_{2,1}}{\partial \hat{\alpha}_p} = 0,$$

we obtain

$$\begin{aligned}\frac{\partial y_{Y,2}}{\partial \hat{\alpha}_p} &= (1-\psi^{-1})\gamma\|\sigma\|^2\left[x_0(m_{0,1}+\hat{\gamma}_0) - x_0(m_{1,1}+\hat{\gamma}_1)\right] \\ &= (1-\psi^{-1})\gamma\|\sigma\|^2 x_0\left[\frac{x_0+x_1}{x_1}\hat{\alpha}_p + \frac{x_2}{x_1}\hat{\sigma} + \hat{\gamma}_0 - \hat{\gamma}_1\right].\end{aligned}$$

Using the change of variables from the market-elasticity derivation,

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

we get

$$\frac{1}{p(X,\epsilon)}\frac{\partial p(X,\epsilon)}{\partial F(X,\epsilon)} = -(1-\psi^{-1})\frac{\gamma\|\sigma\|^2}{y_0(X)}\left[\frac{x_0+x_1}{x_1}\hat{\alpha}_p\epsilon + \frac{x_2}{x_1}\hat{\sigma}\epsilon + (\hat{\gamma}_0-\hat{\gamma}_1)\epsilon\right] + O(\epsilon^2).$$

Finally, using $\overline{\alpha}_p = 1 + \hat{\alpha}_p\epsilon$, $\overline{\sigma} = \|\sigma\|(1+\hat{\sigma}\epsilon)$, and $\gamma_j = \gamma(1+\hat{\gamma}_j\epsilon)$, this becomes

$$\frac{1}{p(X,\epsilon)}\frac{\partial p(X,\epsilon)}{\partial F(X,\epsilon)} = (1-\psi^{-1})\frac{\gamma\|\sigma\|^2}{y_0(X)}\left[\frac{x_0+x_1}{x_1}(1-\overline{\alpha}_p) - \frac{x_2}{x_1}\left(\frac{\overline{\sigma}}{\|\sigma\|}-1\right) - \frac{\gamma_0-\gamma_1}{\gamma}\right] + O(\epsilon^2).$$

Define $x_c \equiv x_2$, $x_u \equiv x_1 = x_a - x_c$, and $\mathbb{E}^u[\gamma_j] \equiv \gamma_1$. Also define

$$\overline{\alpha}_p^{\text{opt}} = 1 - \frac{x_u}{1-x_c}\frac{\gamma_0 - \mathbb{E}^u[\gamma_j]}{\gamma} - \frac{x_c}{1-x_c}\left(\frac{\overline{\sigma}}{\|\sigma\|}-1\right) = 1 + \left(\sum_{j=0}^{1}\frac{x_k}{1-x_c}\hat{\gamma}_k - \hat{\gamma}_0 - \frac{x_c}{1-x_c}\hat{\sigma}\right)\epsilon,$$

where the expression above coincides with first-order correction for the passive investor's optimal portfolio share when the leverage constraint of investor 2 binds.

Since $1 - x_c = x_0 + x_1$, the previous display is equivalent to

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

This proves the version of the result in the regime in which the leverage constraint of investor 2 binds.

□

D Production and Monetary Policy

This appendix describes an extension of the baseline model described in Section 3 to allow for production and monetary policy. The response of interest rates played a crucial role in the determination of aggregate market elasticity, but, in practice, the interest rate is under control of the central bank. In this extension, we show that a central bank following an inflation-targeting policy would choose the same interest rate as the one implied by the baseline model.

We describe next a version of the model with production and monetary policy. We focus only on the main changes relative to the baseline model. We first describe the version of the model with production and flexible prices. We then describe the version of the model with sticky prices and monetary policy.

D.1 The flexible-price version

Firm's problem. As it is standard in the literature, production is carried by a continuum of monopolistically competitive firms indexed by $i \in [0, 1]$. Firm i's production function is given by

$$\tilde{Y}_{i,t} = A_t L_{i,t}^{1-\alpha}, \tag{D.1}$$

where A_t is the productivity level and $L_{i,t}$ is the labor input.

Aggregate productivity follows a geometric Brownian motion:

$$\frac{dA_t}{A_t} = \mu dt + \sigma dZ_t,$$

where μ is the drift and $\sigma \in \mathbb{R}^{1\times 3}$ is the volatility. As in the baseline model, we assume that only the first entry of σ is non-zero.

Using the production function (D.1) to eliminate $L_{i,t}$, we can write the firm's problem as follows:

$$\max_{P_{i,t}} P_{i,t}\tilde{Y}_{i,t} - \mathcal{W}_{i,t}\left(\frac{\tilde{Y}_{i,t}}{A_t}\right)^{\frac{1}{1-\alpha}},$$

subject to the demand function:

$$\tilde{Y}_{i,t} = \left(\frac{P_{i,t}}{P_t}\right)^{-\phi}\tilde{Y}_t, \tag{D.2}$$

where $\mathcal{W}_t$ is the nominal wage, $P_t = \left[\int_{i=0}^{1} P_{i,t}^{1-\phi} di\right]^{\frac{1}{1-\phi}}$ is the aggregate price level and ϕ is the elasticity of substitution between goods.

The first-order condition gives the usual pricing condition, where the price equals a markup

over the marginal cost:

$$P_{i,t} = \frac{\phi}{\phi - 1} \frac{\mathcal{W}_{i,t}}{1-\alpha} \frac{\tilde{Y}_{i,t}^{\frac{\alpha}{1-\alpha}}}{A_t^{\frac{1}{1-\alpha}}}.$$

In a symmetric equilibrium, all firms charge the same price, so $P_{i,t} = P_t$ for all i. We can then write the pricing condition as:

$$(1 - \phi^{-1})(1-\alpha) A_t L_{i,t}^{-\alpha} = \frac{\mathcal{W}_t}{P_t}.$$

Household's problem. As in the baseline model, we have three types of households, $j = 0, 1, 2$, which differ in their risk aversion and portfolio constraints. Following Bigio, Silva, and Zilberman (2025), we assume that households have Epstein-Zin preferences over a GHH composite of consumption and labor.

Household j chooses consumption $\tilde{C}_{j,t}$, labor supply $L_{j,t}$, and portfolio allocations $\alpha_{j,t}$ to solve the problem

$$V_{j,t} = \max_{\tilde{C}_{j,t}, L_{j,t}, \alpha_{j,t}} \mathbb{E}_t \left[\int_t^\infty f_j(\tilde{C}_{j,s} - \xi_s L_{j,s}, V_{j,s})\, ds \right], \qquad s.t.$$

$$\frac{dW_{j,t}}{W_{j,t}} = \left[r_t + \pi_t^\top \alpha_{j,t} - \frac{\tilde{C}_{j,t} - \mathcal{W}_t / P_t L_{j,t}}{W_{j,t}} \right] dt + \sigma_{j,t}\, dZ_t,$$

with $W_{j,t} \geq 0$ and $\alpha_{j,t} \in \Omega_{j,t}$, where $\sigma_{j,t} = \alpha_{j,t}^\top \Sigma_t$ is the household's risk exposure, $W_{j,t}$ denotes the household's wealth, $\Omega_{j,t}$ is the investment set for type-j household, and ξ_t controls the disutility of labor supply.

Define *net consumption* as $C_{j,t} = \tilde{C}_{j,t} - \xi_s L_{j,s}$. We can then rewrite the budget constraint as

$$\frac{dW_{j,t}}{W_{j,t}} = \left[r_t + \pi_t^\top \alpha_{j,t} - \frac{C_{j,t}}{W_{j,t}} + \frac{\mathcal{W}_t / P_t - \xi_t}{W_{j,t}} L_{j,t} \right] dt + \sigma_{j,t}\, dZ_t.$$

In an interior solution, the real wage rate is given by

$$\frac{\mathcal{W}_t}{P_t} = \xi_t.$$

We assume that labor disutility proportional to aggregate producitivity:

$$\xi_t = \xi A_t.$$

This acts as an income effect on labor supply that ensures that labor supply is stationary despite the growth of aggregate producitivity.

Central bank. We consider a cashless economy with a central bank that sets the nominal interest rate i_t. We assume that prices are predetermined, consistent with the behavior under sticky prices, so the nominal interest rate is given by $i_t = r_t + \mu_{P,t}$, where $\mu_{P,t}$ is the drift of the aggregate price level. We consider a central bank that implements zero inflation at all periods, so $\mu_{P,t} = 0$, which requires setting the nominal interest rate to the (natural) risk-free rate.

Equilibrium in labor markets and production. The market clearing condition for labor is given by

$$L_t = \int_{j=0}^{2} \omega_j L_{j,t}.$$

By equating labor demand and supply, we obtain the equilibrium labor supply:

$$L_t = \left(\frac{(1-\alpha)(1-\phi^{-1})}{\xi} \right)^{\frac{1}{\alpha}} \equiv \overline{L}.$$

The market clearing condition for goods is given by

$$\sum_{j=0}^{2} \omega_j \tilde{C}_{j,t} = \tilde{Y}_t.$$

Equivalently, we can write the market clearing condition for goods in terms of net consumption and net output:

$$\sum_{j=0}^{2} \omega_j C_{j,t} = Y_t.$$

where $Y_t = \tilde{Y}_t - \xi_t L_t$.

Net output is given by

$$Y_t = \alpha A_t \overline{L}^{\alpha}.$$

This implies that the aggregate (net) output follows a geometric Brownian motion:

$$\frac{dY_t}{Y_t} = \mu dt + \sigma dZ_t.$$

Hence, the equilibrium of the flexible-price economy is identical to the equilibrium of the baseline endowment economy considered in Section 3, once the nominal interest rate is given by $i_t = r_t$ and the central bank implements zero inflation at all periods.

D.2 The sticky-price version

Suppose now that firms are subject to Rotemberg adjustment costs in pricing. In this case, the firm's problem is given by

$$Q_{i,t}(P_{i,t}) = \max_{[\pi_{i,s}]_{s\geq t}} \mathbb{E}_t \left[\int_t^\infty \frac{\Lambda_s}{\Lambda_t} \left(\frac{P_{i,s}}{P_s} \tilde{Y}_{i,s} - \frac{\mathcal{W}_s}{P_s} \left(\frac{\tilde{Y}_{i,s}}{A_s} \right)^{\frac{1}{1-\alpha}} - \frac{\varphi}{2} \pi_{i,s}^2 \right) ds \right].$$

subject to the demand function (D.2) and $\tilde{P}_{i,t} = \pi_{i,t} P_{i,t}$, given the initial price $P_{i,t}$.

We assume that the adjustment cost is rebated back to shareholders, so they affect the firm's decision, but it does not represent a real resource cost. This assumption is not essential for the results, but it simplifies the analysis.

The next lemma provides a characterization of the equilibrium inflation rate in this economy.

Lemma D.1. *In a symmetric equilibrium where $P_{i,0} = P_0$, the equilibrium inflation rate $\pi_t \equiv \frac{\dot{P}_t}{P_t}$ is given by*

$$\pi_t = \mathbb{E}_t \left[\int_t^\infty \frac{\Lambda_s}{\Lambda_t} \frac{\phi - 1}{\varphi} \frac{\tilde{Y}_t}{P_t} \left(\frac{\phi}{\phi - 1} \frac{\mathcal{W}_t}{P_t} \frac{\tilde{Y}_t^{\frac{\alpha}{1-\alpha}}}{(1-\alpha) A_s^{\frac{1}{1-\alpha}}} - 1 \right) ds \right], \tag{D.3}$$

where we imposed that $P_{i,t} = P_t$, so we obtain a symmetric solution for all firms.

Proof. The value function $Q_{i,t}(P_{i,t}) = Q_i(P_{i,t}; \hat{X}_t)$ is a function of the aggregate state variable $\hat{X}_t$.

The HJB equation is given by

$$0 = \left[\frac{P_{i,t}}{P_t} \tilde{Y}_{i,t} - \frac{\mathcal{W}_t}{P_t} \left(\frac{\tilde{Y}_{i,t}}{A_t} \right)^{\frac{1}{1-\alpha}} - \frac{\varphi}{2} \pi_{i,t}^2 \right] dt + \frac{\mathbb{E}_t[d(\Lambda_t Q_{i,t})]}{\Lambda_t}.$$

More explicitly, we can write the HJB equation as

$$\begin{aligned} 0 = {} & \frac{P_{i,t}}{P_t} \tilde{Y}_{i,t} - \frac{\mathcal{W}_t}{P_t} \left(\frac{\tilde{Y}_{i,t}}{A_t} \right)^{\frac{1}{1-\alpha}} - \frac{\varphi}{2} \pi_{i,t}^2 + \frac{\partial Q_{i,t}}{\partial P_i} \pi_{i,t} P_{i,t} + \frac{\partial Q_{i,t}}{\partial \hat{X}_t^\top} \mu_{\hat{X},t} + \frac{1}{2} \sum_{k=1}^{d} \sigma_{\hat{X},K,t}^\top \frac{\partial^2 Q_{i,t}}{\partial \hat{X}_t \partial \hat{X}_t^\top} \sigma_{\hat{X},K,t} \\ & - r_t Q_{i,t} - \eta_t \sigma_{Q_i,t}^\top. \end{aligned}$$

The first-order condition gives the following pricing condition:

$$\varphi \pi_{i,t} = \frac{\partial Q_{i,t}(P_i)}{\partial P_i}$$

The envelope condition with respect to $P_{i,t}$ is given by

$$0 = (1-\phi)\frac{\tilde{Y}_{i,t}}{P_t}\left[1 - \frac{\phi}{\phi-1}\frac{\mathcal{W}_t}{P_{i,t}}\frac{\tilde{Y}_{i,t}^{\frac{\alpha}{1-\alpha}}}{(1-\alpha)A_t^{\frac{1}{1-\alpha}}}\right]dt + \frac{\mathbb{E}_t[d(\Lambda_t\frac{\partial Q_{i,t}}{\partial P_{i,t}})]}{\Lambda_t}.$$

Multiplying both sides by φ^{-1} and using the first-order condition, we obtain The envelope condition with respect to $P_{i,t}$ is given by

$$0 = \varphi^{-1}(1-\phi)\frac{\tilde{Y}_{i,t}}{P_t}\left[1 - \frac{\phi}{\phi-1}\frac{\mathcal{W}_t}{P_{i,t}}\frac{\tilde{Y}_{i,t}^{\frac{\alpha}{1-\alpha}}}{(1-\alpha)A_t^{\frac{1}{1-\alpha}}}\right]dt + \frac{\mathbb{E}_t[d(\Lambda_t\pi_{i,t})]}{\Lambda_t}.$$

Equivalently, we can write the inflation rate as follows:

$$\pi_t = \mathbb{E}_t\left[\int_t^\infty \frac{\Lambda_s}{\Lambda_t}\frac{\phi-1}{\varphi}\frac{\tilde{Y}_t}{P_t}\left(\frac{\phi}{\phi-1}\frac{\mathcal{W}_t}{P_t}\frac{\tilde{Y}_t^{\frac{\alpha}{1-\alpha}}}{(1-\alpha)A_s^{\frac{1}{1-\alpha}}} - 1\right)ds\right],$$

where we imposed that $P_{i,t} = P_t$, so we obtain a symmetric solution for all firms. □

Equation (D.3) corresponds to the non-linear version of the Phillips curve in our sticky-price model. Crucially, it depends on how much the real marginal cost is expected to deviate from its flexible-price level.

Central bank and the zero-inflation equilibrium. We assume that the central bank adopts the following Taylor rule:

$$i_t = \varrho_t + \phi_\pi \pi_t,$$

where $\phi_\pi > 1$ is the Taylor coefficient on inflation and ϱ_t is a time-varying intercept.

Let r_t^n be the natural rate of interest, that is, the real interest rate in the flexible-price economy. We assume that the time-varying intercept coincides with the natural rate of interest: $\varrho_t = r_t^n$. This implies that the inflation rate is equal to zero in our sticky-price economy.

To see this is the case, notice that the zero-inflation equilibrium requires the following condition to hold at all periods:

$$\frac{\phi}{\phi-1}\frac{\mathcal{W}_t}{P_t}\frac{\tilde{Y}_t^{\frac{\alpha}{1-\alpha}}}{(1-\alpha)A_s^{\frac{1}{1-\alpha}}} - 1 = 0.$$

This coincides with the pricing condition for the flexible-price economy, so the equilibrium labor supply is given by $L_t = \overline{L}$ and aggregate net output is given by $Y_t = \alpha A_t \overline{L}^{\alpha}$. This implies that

Y_t follows the process:

$$\frac{dY_t}{Y_t} = \mu dt + \sigma dZ_t.$$

The interest rate in this economy will then coincide with the natural rate of interest: $i_t = r_t^n$, which requires the intercept ϱ_t to be equal to the natural rate of interest: $\varrho_t = r_t^n$.

E The Second-Order Demand System

Consider next the second-order approximation of the demand system. We focus on the case where there is no preference heterogeneity and no leverage constraints.

Diffusion and drift of price-dividend ratio. The drift of $p_{Y,t}$ is given by

$$\sigma_{p_Y} = \frac{p_{Y,\overline{\alpha}_p}}{p_Y}\sigma_p + \frac{p_{Y,x}}{p_Y}\sigma_x.$$

The price-dividend ratio is given by

$$p_Y(x, \overline{\alpha}_p) = p_Y^* + p_{Y,2}(x, \hat{\alpha}_p)\epsilon^2 + O(\epsilon^3),$$

using the fact that $p_{Y,1} = 0$ in the absence of preference heterogeneity.

The expression above implies that $p_{Y,x} = O(\epsilon^2)$. As $\sigma_x = O(\epsilon)$, we have that $\frac{p_{Y,x}}{p_Y}\sigma_x = O(\epsilon^3)$. We can then ignore this term in the diffusion of $p_{Y,t}$ up to second order.

Taking the derivative of p_Y with respect to $\hat{\alpha}_p$, we obtain

$$\frac{p_{Y,\overline{\alpha}_p}}{p_Y}\frac{\partial \overline{\alpha}_p}{\partial \hat{\alpha}_p} = \frac{p_{Y,2,\hat{\alpha}_p}}{p_{Y,0}}\epsilon^2 + O(\epsilon^3) \Rightarrow \frac{p_{Y,\overline{\alpha}_p}}{p_Y} = \frac{p_{Y,2,\hat{\alpha}_p}}{p_{Y,0}}\epsilon + O(\epsilon^2),$$

as $\frac{\partial \overline{\alpha}_p}{\partial \hat{\alpha}_p} = \epsilon$.

Hence, the diffusion of $p_{Y,t}$ can be written as

$$\sigma_{p_Y} = \frac{p_{Y,2,\hat{\alpha}_p}}{p_{Y,0}}\hat{\sigma}_p\epsilon^2 + O(\epsilon^3).$$

The price-dividend ratio takes the form

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

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

The derivative of the price-dividend ratio with respect to $\overline{\alpha}_p$ is given by

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

The diffusion of the price-dividend ratio is given by

$$\sigma_{p_Y} = \left[a_Y(x_t)(\overline{\alpha}_{p,t} - 1) + b_Y(x_t)\right] \sigma_p + O(\epsilon^3).$$

The diffusion of the price-dividend ratio is given by

$$\mu_{p_Y} = \frac{p_{Y,\overline{\alpha}_p}}{p_Y}\mu_{\overline{\alpha}_p} + \frac{1}{2}\frac{p_{Y,\overline{\alpha}_p\overline{\alpha}_p}}{p_Y}\sigma_p\sigma_p^\top + O(\epsilon^3).$$

We can then write the drift of the price-dividend ratio as follows:

$$\mu_{p_Y} = [a_Y(x)(\overline{\alpha}_{p,t} - 1) + b_Y(x)]\kappa_p(1 - \overline{\alpha}_p) + \frac{1}{2}a_Y(x)\sigma_p\sigma_p^\top + O(\epsilon^3).$$

The demand for the risky asset. The demand for the risky asset is given by

$$q_{j,t} = \alpha_{j,Y,t} - 1.$$

The portfolio share of the aggregate claim is given by

$$\alpha_{j,Y} = \frac{\sigma_j\sigma_{R_Y}^\top}{\sigma_{R_Y}\sigma_{R_Y}^\top} = \frac{(\sigma + \sigma_{j,1}\epsilon + \sigma_{j,2}\epsilon^2)(\sigma + \sigma_{p_Y,2}\epsilon^2)^\top}{(\sigma + \sigma_{p_Y,2}\epsilon^2)(\sigma + \sigma_{p_Y,2}\epsilon^2)^\top} = 1 + \frac{\sigma_{j,1}\sigma^\top}{\sigma\sigma^\top}\epsilon + \frac{(\sigma_{j,2} - \sigma_{p_Y,2})\sigma^\top}{\sigma\sigma^\top}\epsilon^2 + O(\epsilon^3).$$

The risk exposure of active investors is given by

$$\sigma_{j,t} = \frac{\eta_0}{\gamma} + \frac{\eta_1}{\gamma}\epsilon + \left[\frac{\eta_2}{\gamma} + \frac{1 - \gamma^{-1}}{\psi - 1}\sigma_{c_j,2}\right]\epsilon^2 + O(\epsilon^3),$$

and the risk exposure of the passive investor is given by

$$\sigma_{0,t} = \sigma + \hat{\alpha}_p\sigma\epsilon + \sigma_{p_Y,2}\epsilon^2 + O(\epsilon^3).$$

Combining the expressions above, we obtain the demand for the risky asset for active investors:

$$q_{j,t} = \frac{\eta_t\sigma^\top}{\gamma\|\sigma\|^2} + \left[\frac{1 - \gamma^{-1}}{\psi - 1}\sigma_{c_j,t} - \sigma_{p_Y,t}\right]\frac{\sigma^\top}{\|\sigma\|^2} + O(\epsilon^3).$$

The demand for the risky asset for passive investors is given by

$$q_{0,t} = \overline{\alpha}_{p,t} - 1 + O(\epsilon^3).$$

Next, we eliminate the market price of risk from the demand for the risky asset. From the pricing condition for the risky asset, we have

$$\frac{1}{p_Y} = r + \eta(\sigma + \sigma_{p_Y})^\top - \left(\mu + \mu_{p_Y} + \sigma\sigma_{p_Y}^\top\right).$$

Rearranging the expression above, we obtain

$$\eta\sigma^\top = \frac{1}{p_Y} - r + \mu + \mu_{p_Y} + \sigma\sigma_{p_Y}^\top - \eta\sigma_{p_Y}^\top.$$

Expanding the expression above, we obtain

$$\left[\eta_1\epsilon + \eta_2\epsilon^2\right]\sigma^\top = \left[-\frac{1}{p_{Y,0}^2}p_{Y,1} - r_1\right]\epsilon + \left[-\frac{p_{Y,2}}{p_{Y,0}^2} + \frac{p_{Y,1}^2}{p_{Y,0}^3} - r_2 + \mu_{p_Y,2} + (1-\gamma)\sigma\sigma_{p_Y,2}^\top\right]\epsilon^2.$$

Using the fact that $p_{Y,1} = 0$ in the absence of preference heterogeneity, we obtain

$$q_j = -\frac{p_{Y,2}\epsilon^2}{p_{Y,0}^2\gamma\|\sigma\|^2} - \frac{r_1\epsilon + r_2\epsilon^2}{\gamma\|\sigma\|^2} + \left[\mu_{p_Y,2} + (1-2\gamma)\sigma_{p_Y,2}\sigma^\top + \frac{\gamma-1}{\psi-1}\sigma_{c_j,2}\sigma^\top\right]\frac{\epsilon^2}{\gamma\|\sigma\|^2} + O(\epsilon^3).$$

Using the fact that $\hat{p}_{Y,t} = \frac{p_{Y,t}-p_{Y,0}}{p_{Y,0}}$, we can write the demand for the risky asset for active investors as follows:

$$q_{j,t} = -\zeta_{j,p}\hat{p}_{Y,t} - \zeta_{j,r}\hat{r}_t + f_{j,t} + O(\epsilon^3),$$

where $\zeta_{j,p} = \frac{1}{p_{Y,0}}\frac{1}{\gamma\|\sigma\|^2}$, $\zeta_{j,r} = \frac{1}{\gamma\|\sigma\|^2}$, and the shifter $f_{j,t}$ is given by

$$f_{j,t} = \underbrace{\zeta_{j,r}(\mu_{p_Y,t} + \sigma_{p_Y,t}\sigma^\top)}_{\text{expected price appreciation}} - \underbrace{2\frac{\sigma_{p_Y,t}\sigma^\top}{\sigma\sigma^\top}}_{\text{endogenous volatility}} + \underbrace{\frac{1-\gamma^{-1}}{\psi-1}\frac{\sigma_{c_j,t}\sigma^\top}{\sigma\sigma^\top}}_{\text{hedging demand}}.$$

The demand for the risky asset for passive investors takes the same form, but with $\zeta_{0,p} = \zeta_{0,r} = 0$, and the shifter $f_{0,t}$ given by

$$f_{0,t} = \hat{\alpha}_{p,t}\epsilon.$$

Consumption-wealth ratio. The aggregate consumption-wealth ratio $c_t \equiv \sum_{j=0}^{2} x_{j,t} c_{j,t}$ is given by

$$c_t = \psi\rho + (1-\psi)\left[r_t + \eta_t \sigma_{R_Y,t}^{\top} - \frac{\gamma}{2}\sum_{j=0}^{2} x_{j,t}\|\sigma_{j,t}\|^2\right] + \sum_{j=0}^{2} x_{j,t}\left[\mu_{c_j,t} + (1-\gamma)\sigma_{c_j,t}\sigma_{R_Y,t}^{\top} + \frac{\psi-\gamma}{1-\psi}\frac{\|\sigma_{c_j,t}\|^2}{2}\right].$$

The risk exposure of the passive investor is given by $\sigma_{0,t} = \overline{\alpha}_{p,t}\sigma_{R_Y,t}$. From market clearing, the risk exposure of active investors is given by $\sigma_{j,t} = \frac{1-x_{0,t}\overline{\alpha}_{p,t}}{1-x_{0,t}}\sigma_{R_Y,t}$. Hence, the risk exposure of both group of investors is proportional to $\sigma_{R_Y,t}$. We can then write the consumption-wealth ratio as follows:

$$c_t = \psi\rho + (1-\psi)\left[r_t + \eta_t \sigma_{R_Y,t}^{\top} - \frac{\gamma\|\sigma_{R_Y,t}\|^2}{2}\sum_{j=0}^{2} x_{j,t}\alpha_{j,Y,t}^2\right] + \sum_{j=0}^{2} x_{j,t}\left[\mu_{c_j,t} + (1-\gamma)\sigma_{c_j,t}\sigma_{R_Y,t}^{\top} + \frac{\psi-\gamma}{1-\psi}\frac{\|\sigma_{c_j,t}\|^2}{2}\right].$$

Expanding the expression above up to second order, we obtain

$$c_2 = (1-\psi)\left[r_2 + \eta_2\sigma^{\top} - \frac{\gamma\|\sigma\|^2}{2}\frac{x_0}{x_a}(\overline{\alpha}_{p,t}-1)^2\right] - \left[\mu_{p_Y,2} + (1-\gamma)\sigma_{p_Y,2}\sigma^{\top}\right].$$

From the pricing condition for the risky asset, we have

$$r_2 + \eta_2\sigma^{\top} = -\frac{p_{Y,2}}{p_{Y,0}^2} + \mu_{p_Y,2} + (1-\gamma)\sigma_{p_Y,2}\sigma^{\top}.$$

Combining the expressions above, we obtain

$$c_2 = (1-\psi)\left[-\frac{p_{Y,2}}{p_{Y,0}^2} + \mu_{p_Y,2} + (1-\gamma)\sigma_{p_Y,2}\sigma^{\top} - \frac{\gamma\|\sigma\|^2}{2}\frac{x_0}{x_a}(\overline{\alpha}_{p,t}-1)^2\right] - \left[\mu_{p_Y,2} + (1-\gamma)\sigma_{p_Y,2}\sigma^{\top}\right].$$

The consumption-wealth ratio takes the form

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

where

$$\chi_{0,t} = (\psi-1)\frac{\gamma\|\sigma\|^2}{2}\frac{x_0}{x_a}(\overline{\alpha}_{p,t}-1)^2 - \psi\left[\mu_{p_Y,2} + (1-\gamma)\sigma_{p_Y,2}\sigma^{\top}\right].$$

The aggregate market elasticity. Using the market clearing condition for goods, we have

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

The left-hand side of the expression above is given by

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

The right-hand side of the expression above is given by

$$\hat{p}_{Y,t} = -\left[(1-\psi^{-1})\gamma\|\sigma\|^2\frac{x_0}{x_a} + 2\kappa_p a_Y(x)\right]\frac{(\overline{\alpha}_{p,t}-1)^2}{2y^*} - \frac{1}{y^*}\left[b_Y(x)\kappa_p + (\gamma-1)a_Y(x_t)\sigma_p\sigma^\top\right](\overline{\alpha}_{p,t}-1)$$

$$+ \frac{1}{y^*}\left[\frac{1}{2}a_Y(x)\sigma_p\sigma_p^\top + (1-\gamma)b_Y(x_t)\sigma_p\sigma^\top\right]$$

Matching coefficients, we obtain a condition determining $a_Y(x)$:

$$y^* a_Y(x) = -\left[(1-\psi^{-1})\gamma\|\sigma\|^2\frac{x_0}{x_a} + 2\kappa_p a_Y(x)\right].$$

Rearranging the expression above, we obtain

$$a_Y(x) = -(1-\psi^{-1})\frac{\gamma\|\sigma\|^2}{y^* + 2\kappa_p}\frac{x_0}{x_a}.$$

Similarly, we can determine $b_Y(x)$:

$$y^* b_Y(x) = -\left[b_Y(x)\kappa_p + (\gamma-1)a_Y(x_t)\sigma_p\sigma^\top\right].$$

Rearranging the expression above, we obtain

$$b_Y(x) = (1-\gamma)\frac{\sigma_p\sigma^\top}{y^* + \kappa_p}a_Y(x_t).$$

The derivative of $\hat{p}_{Y,t}$ with respect to $\overline{\alpha}_p$ is given by

$$\frac{\partial\hat{p}_{Y,t}}{\partial\overline{\alpha}_p} = a_Y(x)(\overline{\alpha}_{p,t} - 1) + b_Y(x).$$

Using the expression above, we obtain

$$\epsilon_{Y,t}^{-1} \equiv \frac{1}{x_{0,t}}\frac{\partial\hat{p}_{Y,t}}{\partial\overline{\alpha}_p} = \frac{1-\psi^{-1}}{y^* + 2\kappa_p}\frac{\gamma\|\sigma\|^2}{x_a}\left[1 - \overline{\alpha}_{p,t} + \frac{\gamma-1}{y^* + \kappa_p}\sigma_p\sigma^\top\right] + O(\epsilon^2).$$

Endogenous volatility and hedging demand. The volatility of the price-dividend ratio is given by

$$\sigma_{p_Y,t} = x_{0,t}\epsilon_{Y,t}^{-1}\sigma_p + O(\epsilon^3).$$

Let $c_{j,t}^0$ denote the consumption-wealth ratio in the fixed-$\hat{\alpha}_p$ economy. In this case, we can write $y_{Y,2}^0 = d_Y\hat{\alpha}_p^2$, and the second-order correction to the consumption-wealth ratio of active investor $j = 1, 2$ is given by

$$c_{j,2}^0 = \left[1 - \frac{1}{x_a}\psi\right] d_Y^0\hat{\alpha}_p^2,$$

and the second-order correction to the consumption-wealth ratio of passive investor is given by

$$c_{0,2}^0 = \left[1 + \frac{1}{x_0}\psi\right] d_Y^0\hat{\alpha}_p^2.$$

This implies that the diffusion of the consumption-wealth ratio of active investor $j = 1, 2$ is given by

$$\sigma_{c_j,t} = -\left[1 - \frac{1}{x_a}\psi\right] x_{0,t}\epsilon_{Y,t}^{-1}\sigma_p + O(\epsilon^3),$$

and the diffusion of the consumption-wealth ratio for passive investor is given by

$$\sigma_{c_0,t} = -\left[1 + \frac{1}{x_0}\psi\right] x_{0,t}\epsilon_{Y,t}^{-1}\sigma_p + O(\epsilon^3).$$

We can then write the hedging demand for active investor $j = 1, 2$ as follows:

$$h_{j,t} = \frac{1 - \gamma_j^{-1}}{\psi - 1}\sigma_{c_j,t} = H_{a,t}x_{0,t}\epsilon_{Y,t}^{-1}\sigma_p + O(\epsilon^3),$$

where $H_{a,t} = \frac{1-\gamma^{-1}}{\psi-1}\left[\frac{\psi}{x_{a,t}} - 1\right]$.

Market price of risk and risk premium. The market price of risk is given by

$$\eta_t = \gamma\sigma\chi_{a,t} + \left[1 - H_{a,t}\right]\gamma x_{0,t}\epsilon_{Y,t}^{-1}\sigma_p + O(\epsilon^3),$$

where $\chi_{a,t} = \frac{1-x_{0,t}\overline{\alpha}_{p,t}}{1-x_{0,t}}$ is the aggregate risky exposure that must be absorbed by active investors per unit of active wealth, and $h_{a,t}$ is the average hedging demand of active investors, adjusted for leverage constraints.

The risk premium is given by

$$\pi_{Y,t} = \gamma\|\sigma\|^2\chi_{a,t} + \left[2 - H_{a,t}\right]\gamma x_{0,t}\epsilon_{Y,t}^{-1}\sigma_p\sigma^\top + O(\epsilon^3).$$

Interest rate. The interest rate follows from the pricing condition for the aggregate claim,

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

Thus,

$$r_t = y_{Y,t} + \mu + \mu_{p_Y,t} + \sigma\sigma_{p_Y,t}^\top - \pi_{Y,t},$$

where $y_{Y,t} = 1/p_{Y,t}$ and $\pi_{Y,t} = \eta_t \sigma_{R_Y,t}^\top$.

From the market clearing condition for goods, we have

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

which can be written as

$$(1-\psi^{-1})\frac{\gamma\|\sigma\|^2}{2}\frac{x_0}{x_a}(\overline{\alpha}_{p,t} - 1)^2 = \left[\mu_{p_Y,t} + (1-\gamma)\sigma_{p_Y,t}\sigma^\top - y^*\hat{p}_{Y,t}\right] + O(\epsilon^3).$$

We can then write the interest rate as follows:

$$\hat{r}_t = (1-\psi^{-1})\frac{\gamma\|\sigma\|^2}{2}\frac{x_0}{x_a}(\overline{\alpha}_{p,t} - 1)^2 - \hat{\eta}_t\sigma^\top + O(\epsilon^3).$$

F Details of Section 5

F.1 Flow of Funds aggregation

This appendix describes the procedure used in Section 5.1 to aggregate Flow of Funds (FoF) sectors into the three ultimate sectors of the model: passive, constrained active, and unconstrained 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), *constrained* (hedge funds,

broker-dealers, and other short-horizon intermediaries), and *unconstrained* (insurance, pensions, and rest-of-the-world net of hedge funds).

F.2 Constructing quarterly equity flows

This appendix describes how we construct the quarterly aggregate equity flow series plotted in Figure 7 and used to discipline the passive-demand process in Section 5.1. 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 b_j via time-series regressions of:

$$\log(V_{j,t}/V_{j,t-1}) = a_j + b_j \cdot \text{Mkt}_t + \epsilon_{j,t}$$

where Mkt_t is the CRSP market return factor. Then, we impute the "active" dollar holdings change as:

$$\Delta F^E_{j,t} = V_{j,t} - V_{j,t-1} \cdot (1 + \beta_j \cdot \text{Mkt}_t)$$

Note that this adjustment removes mechanical revaluation effects. Let ΔF^{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 F.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 the cross-sectional average of absolute scaled flows,

$$y^{\text{Gross}}_t = \frac{\sum_j \left|\Delta F^E_{j,t}\right| + \left|\Delta F^{\text{Firm}}_t\right|}{2\, E_{t-1}},$$

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 the left panel of Figure 7 and is the empirical analogue of the volatility of passive demand σ_p in the model.

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 σ_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. This indirect-inference target is robust to the level of equity risk premia and aggregate volatility, since both numerator and denominator scale with the size of the equity market. For the correlation parameter ν_p, we regress the quarterly aggregate flow series on contemporaneous market returns and use the slope to discipline the loading of passive-demand shocks on the aggregate fundamental shock.

F.3 Numerical methodology

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

F.3.1 Simplified economy

The state variables in the simplified economy are

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

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

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

with neural networks parameterized by Θ:

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

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

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

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

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

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

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

their first and second derivatives are known. The problem reduces to minimizing the following loss functions

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

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

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

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

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

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

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

The resulting loss is

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

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

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

The full algorithm is given in Algorithm (1).

F.3.2 Valuation block

After solving the simplified economy, we price the dividend claim as a function of the state variables X. We approximate the price-dividend ratio by a additional neural network $\hat{p}_D(\Theta_D)$ parameterized by Θ_D. At this stage the first-stage simplified equilibrium is held fixed. For each state X, the first-stage provides the risk-free rate $r(X)$, the market price of risk $\eta(X)$, the state drift $\mu_X(X)$, the state diffusion matrix $\Sigma_X(X)$, and the aggregate-claim return volatility. Given a

Algorithm 1 First-stage training.

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

candidate price-dividend ratio $\hat{p}_D(X)$, we compute its first and second derivatives with respect to X using automatic differentiation, leading to the residual $\mathcal{L}_D^{\text{val}}(X; \Theta_D)$. The training objective is simply the empirical mean squared PDE residual,

$$\widehat{\mathcal{L}}_D^{\text{val}}(\Theta_D) = \frac{1}{N} \sum_{n=1}^{N} \big(\mathcal{L}_D^{\text{val}}(X^n; \Theta_D)\big)^2.$$

The full algorithm is given in Algorithm (2). The total time to run the baseline model on a single-threaded computer with 128GB RAM and NVIDIA Ada 4500 GPU is 2 hours. The HJB loss in the first-stage training problem converges to a mean-squared loss of 3.7×10^{-7}. The corresponding price-dividend ratio training problem converge to a normalized mean-squared training loss of 1.4×10^{-6}. These losses are computed from the batch objectives in (F.1) and in the valuation objective, respectively.

Algorithm 2 Valuation step.

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