Demand Elasticity in Dynamic Asset Pricing*

Zhiguo He
Stanford University and NBER

Péter Kondor
London School of Economics and CEPR

Jessica Shi Li
University of Chicago

First version: June 2025, This version: August 2025

Preliminary; Comments welcome
For latest version click here

Abstract

We show that the standard econometric approach relying on exogenous residual supply shocks systematically underestimates the slope of investors' asset demand curves in dynamic settings. We develop a multi-asset, multi-agent continuous-time model featuring long-lived investors and short-lived hedgers whose demand shifts create exogenous supply changes. Our key finding is that these supply shifts alter the very demand slope that econometricians seek to measure through general equilibrium effects. Namely, supply shifts affect endogenous risk (return volatility and covariance) and change the comovement of returns with future investment opportunities. We compare the "observed slope" measured by econometricians against its conceptual counterpart, the "dynamic slope" derived from investors' first-order condition. Using calibrations based on mutual fund demand shifts and index inclusion events, we find that the dynamic slope is two to seven times larger than the observed measures. The effect through endogenous risk dominates this mismeasurement, while the magnitude depends critically on shock persistence, size, and investor risk aversion.

*He: Graduate School of Business, Stanford University; FTG and NBER; hezhg@stanford.edu. Kondor: London School of Economics and CBER, p.kondor@lse.ac.uk; Li: Booth School of Business, University of Chicago, jessicali@chicagobooth.edu. For helpful comments, we thank Zhiyu Fu, Valentin Haddad, Ben Hebert, Arvind Krishnamurthy, Jiacui Li, Sydney Ludvigson, Ian Martin, and Dimitri Vayanos. All errors are our own.

1 Introduction

Understanding the return slope of investors' asset demand in a dynamic context is essential for a large group of problems faced by today's academics, practitioners, and policymakers alike: How much should the price of GameStop react to a persistent surge in retail demand? How much should the yield curve drop in response to the central bank's government bond purchases if the QE program is expected to be tapered in six months? How much should the stock prices of firms in green industries react to the expectation that green funds will double in size over the next decade?

The standard approach to answering these questions relies on exogenous shifts in the residual supply. Our key finding is that these supply shifts alter the very demand slope that econometricians seek to measure through general equilibrium effects. More precisely, the supply shifts affect the essential determinants of the demand curve in a Merton (1971) type dynamic economy, via endogenous asset risk that includes return volatility and covariance, and the endogenous dynamic hedging component which captures the comovement of returns with future investment opportunities. Putting it differently, in a dynamic economy investors' asset demand depends on the expectation of future asset return dynamics; our effects stem from the fact that the shifts of residual supply curve not only changes today's asset prices but also the entire future paths.

To gauge the magnitude of potential bias we compare the "observed slope" measured by econometricians against its conceptual counterpart, the "dynamic slope" derived from the investor's first order condition. Using calibrations based on mutual fund demand shifts and index inclusion events, we find that the dynamic slope is two to seven times larger than the observed measures. The effect through endogenous risk dominates this mismeasurement, while the magnitude depends critically on shock persistence, size, and investor risk-aversion.

We study a multi-asset-multi-agent dynamic asset pricing model in continuous time. Assets are claims on risky cash flows driven by a multi-dimensional Brownian shock. The economy is populated by two types of agents: hedgers and investors. Investors are long-lived and solve a canonical consumption-saving problem, with optimal portfolio decisions similar to that in Merton (1971). For simplicity, hedgers are assumed to be short-lived. Hedgers receive a stochastic endowment, which is correlated with the cash flows of the asset. This correlation is changing over time driven by a Poisson process. This leads to periodic shifts in hedgers' demand for the assets, that is, shifts in investors' residual supply. In equilibrium, asset prices—and hence their endogenous risks—adjust in response.

Our main question is that whether observing the price and quantity response due to an exogenously shift in residual supply can help recover the slope of the representative investor's demand curve.

Consider a single risky asset. In a static version of our economy, the answer is affirmative.

As in the Markowitz model, the slope of investors' static demand is the ratio of risk-tolerance and return variance. Both are primitives in a static economy. When hedgers demand less of an asset, institutions have to hold more, leading to an increase in its expected equilibrium return. Comparing the changes in investors' demand for the assets and the induced changes in expected returns reveals the slope of their demand curve.

This is in contrast even with the simplest case of our dynamic model where investors have log-utility. In line with Merton (1971), the slope of log-investor's demand curve is simply the inverse of return variance. However, in a dynamic model return volatility is an endogenous object. In fact, when hedgers partially liquidate their positions following the shock, investors have to increase their risk exposure in equilibrium, which gives rise to higher volatility of the asset return. Because the slope is the inverse of the return variance, the shift in residual supply pushes down the slope we wish to recover, leading to a flatter observed slope relative to its conceptual counterpart.

Why does the asset volatility endogenously rise when investors hold more of the asset in equilibrium due to partial liquidation of the asset from hedgers? It is due to a standard amplification effect that prevails in many heterogeneous agent models in macroeconomics and finance.[1] If investors hold more risky assets, it implies that future adverse dividend shocks result in larger adverse shocks to investors' aggregate wealth reducing their risk-bearing capacity increasing risk premium, and implying a larger decrease in price, which amplifies the initial drop in aggregate wealth and, consequently, the price effect of the initial shock.

Allowing for multiple assets adds a covariance channel to this mechanism. As the endogenous volatility of all assets is due to investors' fluctuation of wealth, when investors hold more of a given asset, not only the asset's volatility but also its comovement with every other asset grows. In other words, assets with idiosyncratic dividend risk become more correlated when the pricing kernel, which is a function of the aggregate wealth of the investor sector, becomes more volatile. As a result, in our multi-asset economy case where the slope of investors' demand curve is proportional to the inverse of the covariance matrix, the variance and the covariance channels tend to go into the same direction: both makes the observed slope flatter than its conceptual counterpart. We refer to these effects operating through the endogenous covariance matrix as the *endogenous risk effect*. In our baseline calibration, this effect is responsible for the largest share of the mismeasurement.

When investors' risk-aversion is different from 1, there is an additional effect connected to a general equilibrium version of Merton's intertemporal hedging demand. In our economy, each investor's relative demand for different assets depends on the portfolio of other investors. It is so, because each understands that if others overweight a given asset, whenever that asset peforms well, the groups' risk-bearing capacity improves, risk premia decrease, making

[1]The leading papers are Longstaff and Wang (2012), He and Krishnamurthy (2013), Brunnermeier and Sannikov (2014), and Kondor and Vayanos (2019), among others.

investment opportunities worse. As Merton (1971) explains, when the investor's risk-aversion is larger than 1, she wants to hedge against that, increasing her exposure to the overweighted assets. Therefore, when hedgers are shocked and sell some assets to investors, each investor's demand slope shifts upward due to this force. That is, the residual supply shock affects the demand curve. We refer to this as the *amplified intertemporal hedging effect*.

Finally, investors will also be concerned whether the shift in residual supply is persistent. Each investor understands that if and when households' demand shifts back and they buy back the shocked assets, prices will adjust affecting the wealth of investors who previously overweighted the shocked assets. Again, foreseeing this effect alters the required change in expected returns today for financial institutions to be willing to hold the increased supply of shocked assets. That is, again the slope of their demand curve is affected by the shift in residual supply. We refer to this effect as *hedging against the reversal*.

We formalize our observations by comparing two objects: the *observed slope* and its conceptual counterpart: the *dynamic slope*. The former is simply the equilibrium change in investors' demand per unit change in expected returns before and after the Poisson shock. This is akin to the inverse of the multiplier as referred to in Gabaix and Koijen (2022). The latter instead captures the standard economic concept of the slope of a demand curve: the optimal quantity response of a representative investor per unit change in expected returns keeping all else fixed. Formally, the dynamic slope gives the answer to the question of how the long-term investor would change her portfolio shares under a counterfactual dividend process which would increase the expected returns by an arbitrary vector, until a Poisson shock hits. The difference between these two concepts stems from general equilibrium effects. The observed slope is what the econometrician observes in equilibrium given that all investors react to the shock while the dynamic slope is derived from the investors' first order condition.

To give a fair chance for the observed slope to coincide with the dynamic slope, we define the latter with respect to a richly parametrized counterfactual dividend process. In fact, we make sure that both objects measure quantity responses to the same change in instantaneous expected return, the same flow share of additional return (versus the lumpsum when the shock reverses), and finally the same persistence of the shock. Hence, our results show that the observed slope mismeasures its conceptual counterpart even if they correspond to the same change in the instantaneous expected return vector, and to similar changes in the follow-up price path.[2]

As we explained above, the observed slope is generally flatter than the dynamic slope. We formally decompose the difference to the endogenous risk effect, the amplified intertemporal

[2]We also note that the simpler concept of *myopic slope*, defined as the product of risk-tolerance and the inverse of the covariance matrix, quantitatively tends to be very close to the dynamic slope. So all our observations on how the observed slope mismeasures the dynamic slope continue to hold if the reader prefers to compare the observed slope with the myopic slope instead.

hedging effect, and the hedging against the reversal effect.

In the final part of the paper we quantify our effects by building two calibrated applications. In our first application, we are interested in our effects in the factor-elasticity context. We calibrate our parameters to Ben-David, Li, Rossi and Song (2021) who measure the price effect of a persistent change in the demand from mutual funds due to a change in Morningstar's rating methodology. In our second application, we are interested in the likely magnitude of "index inclusion" à la Shleifer (1986); hence, this application is in a micro-elasticity context. For this purpose, we study an economy with many fundamentally similar assets, where a small subset of assets have time-varying hedgers' demand akin to index inclusion in practice. More precisely, we assume that with a Poisson shock, a single asset in the index is replaced with one outside; that is, hedgers' demand changes for these two assets only. We compare the observed slope to its conceptual counterpart, namely the dynamic slope.

Although, in line with Gabaix and Koijen (2022), the magnitudes of the observed slope differ across the factor-elasticity and micro-elasticity exercises, we find similar quantitative patterns of mismeasurement across our two calibrations. Namely, in each of our calibrations the dynamic demand slope is close to three times as large as its observed slope based estimate and can be eight times as large when risk-aversion is small. Our decomposition shows that the endogenous risk effect tends to be the bulk of the deviation. The amplified intertemporal hedging effect goes against the endogenous risk effect whenever investors are more risk-averse than the log investor and amplifies it otherwise. Quantitatively, it can be responsible to up to 30% of the mismeasurement when the implied shock in returns is large and persistent. We find that the hedging for reversal effect tends to be small.

Literature Review. To our knowledge,this paper is the first to point out that exogenous shocks to residual supply, due to dynamic general equilibrium effects through changing risk and investment opportunities, alter the very demand slope that econometricians intend to measure based on these shocks. However, our paper naturally builds on several strands of the existing literature.

The closest to our paper is the concurrent work by van Binsbergen, David and Opp (2025) who also highlight that using supply shocks as instruments might violate the exclusion restriction in a dynamic economy. However, they focus on the mismeasurement of price elasticity due to the persistence and predictability of the instrumented price variation. This bias stems from the fact that instrumented price changes are poor proxies for variation in expected returns – the real driver of investors' demand for financial assets. In contrast, we show that the mechanism in our paper would be present even if the econometrician were able to precisely measure the effect of supply shocks on expected instantaneous returns. At the same time, we focus on the effect of one-off supply shocks, hence the bias identified in van Binsbergen et al. (2025)) plays a minimal role in our work.

Second, seminal papers of demand system estimation, (Koijen and Yogo, 2019; Koijen, Richmond and Yogo, 2023; Gabaix and Koijen, 2022) set up dynamic models as the conceptual background for their estimates. However, in these frameworks the possibility of our effects is ruled out by assumption. For instance, in Koijen and Yogo (2019) the form of the covariance matrix is determined by an assumption under which it is independent of the residual supply shocks. This rules out the endogenous risk effect. Similarly, the main model in Gabaix and Koijen (2022) (as its generalized version in Gabaix, Koijen, Mainardi, Oh and Yogo (2025)) assumes that investors' response to changes in expected returns is driven by an exogenous, constant coefficient. In contrast, our model allows for endogenous risk deriving the equilibrium covariance matrix from first principles.

Two recent papers, Fuchs, Fukuda and Neuhann (2025) and Haddad, He, Huebner, Kondor and Loualiche (2025) (HHHKL, from now on) focus on problems stemming from the fact that investors' demand curve for a given asset potentially depends the prices of all assets in the investment opportunity set. HHHKL work out a flexible method to estimate each coefficient of this demand curve, using the cross-section for relative elasticity between assets with same observables, and the time-series for substitution across assets with different obervables. As these papers focus on a static economy where demand curve coefficients are exogenously given by primitives, while at the same time we allow the econometrician to observe the effect of the shock to all prices, our work is orthogonal to this literature. That is, all our results are above and beyond the ones highlighted in these two papers.

Having said that, there are a number of papers that have highlighted issues that we are touching on. For instance, just as we do, Gabaix and Koijen (2022) emphasize that the persistence of the shock should matter for the response in quantities. Davis, Kargar and Li (2023), in line with our approach, explain that focusing on price elasticity instead of return elasticity might be problematic, because "price pass-through"—a concept closely related to the persistence of the shock—strongly influences this mapping. An and Zheng (2025) build a dynamic asset pricing model with a role for the predictability of flow shocks that are similar to our residual supply shocks. Somewhat relatedly, we also focus on non-trivial effects of these shocks to prices and quantities.

We think of Haddad, Moreira and Muir (2024) as the closest to provide empirical evidence for our effect by showing that the Fed's quantitative easing program (a series of residual supply shocks from investors' point of view) changed the market perception of the future return distribution of government bonds, hence changing investors' demand for them.

As we mentioned, the endogenous risk effect is present in a wide range of models with heterogeneous agents with and without explicit financial frictions (e.g. Longstaff and Wang, 2012; He and Krishnamurthy, 2013; Brunnermeier and Sannikov, 2014; Kondor and Vayanos, 2019). Thanks to its simple structure, we build our model on Kondor and Vayanos (2019). None of these papers focus on the problem of recovering demand sensitivities using exogenous

residual supply shocks. We are also unaware of any other models that emphasize the amplified intertemporal hedging effect.

Finally, in the companion paper He, Kondor and Li (2025), we take a step back from the view that asset markets can be seen as a demand system with well-defined supply and demand shocks. Instead, we focus on the quantitative effect of an exogenous supply change on equilibrium returns and prices. As we explain in this paper, this shock affects the demand curves of each group, hence the multiplier does not correspond to the inverse of the demand elasticity for any group. Still, we provide sophisticated benchmarks for the literature on what "canonical" models imply on this aggregate multiplier and how sensitive it is to changes of parameters in the economy.

2 The Tilt and the Shift: A Minimal Demand System

To set the stage, we start with an illustrative example. We present a demand system with the minimal reduced-form properties required for our main effects to be present. In the dynamic model, these properties arise endogenously.

Minimal demand system. Consider the following minimal demand system with a single asset, which is given as follows:

$$\begin{aligned} \text{hedger's demand:} \quad & \hat{x} = c\left[\mathbb{E}(R) - r\right] + \varepsilon_x, \\ \text{investor's demand:} \quad & \hat{y} = a\left(y\right) + b(y)\left[\mathbb{E}(R) - r\right] + \varepsilon_y, \\ \text{marker clearing:} \quad & S = x + y. \end{aligned} \tag{1}$$

Here, r denotes the exogenous risk-free rate while $\mathbb{E}(R)$ denotes the asset's *endogenous* expected return. The individual demand of any representative hedger and investor, denoted by $\hat{x}$ and $\hat{y}$ respectively, depends on the expected excess return $\mathbb{E}(R) - r$, with $\varepsilon_x, \varepsilon_y$ being orthogonal error terms. The positive constant S is the asset's aggregate supply, which is fixed in our asset pricing context. As x, y are the aggregate counterparts of demand, the last equation represents market clearing.

The only non-standard element of the minimal asset demand system (1) is on the investor's side: both the slope, $b(y)$, and the intercept, $a(y)$, of the investor's demand curve depend on the equilibrium allocation of risk represented by y. We will explain why this would be the case shortly.

Demand slope and Observed slope. We think of the slope of an agent's demand as her optimal quantity response to unit change in expected returns. In the demand system (1), the slope of hedger's demand is captured by a positive constant c, which says that each

individual hedger will increase c units of this asset per unit of increase in the asset's excess expected return $\mathbb{E}(R) - r$.

Our focus is on the slope of the investors' asset demand. Suppose that an investor is asked how much she would increase her demand for the asset if $\mathbb{E}(R)$ would increase to $\mathbb{E}(R) + \delta$ where δ is an arbitrary constant. In the demand system in (1), this gives the *conceptual* slope of investors' demand curve as

$$\mathcal{C}^{\text{minimal}} \equiv \frac{\hat{y}^h - \hat{y}}{\delta} = b(y),$$

where $\hat{y}^h$ is the demand under the hypothetical increase in expected excess return.

Now suppose instead that we observe the equilibrium under two independent realizations of $\varepsilon_x = \{\varepsilon_x^n, \varepsilon_x^s\}$. That is, by only shocking the x-demand-curve, this gives us the best possibility to identify the slope of y-demand-curve. Then, the observed equilibrium relationship between changes in quantities and returns, that is, the *observed slope* is

$$\begin{aligned}\mathcal{O}^{\text{minimal}} &\equiv \frac{y^s - y^n}{\mathbb{E}(R^s) - \mathbb{E}(R^n)} \\ &= \underbrace{b(y^n)}_{\text{slope of demand}} + \underbrace{\frac{(b(y^s) - b(y^n))[\mathbb{E}(R^s) - r]}{\mathbb{E}(R^s) - \mathbb{E}(R^n)}}_{\text{endogenous risk effect}} + \underbrace{\frac{(a(y^s) - a(y^n))}{\mathbb{E}(R^s) - \mathbb{E}(R^n)}}_{\text{amplified hedging effects}}. \end{aligned} \tag{2}$$

where the superscripts refer to the equilibrium values under the two realizations. Just as in the text-book case, if the intercept and the slope of investors' demand were constants, then the observed slope would recover the slope of investors' demand. When this is not the case, the second and third terms will bias the measurement.

We refer to the second term as the endogenous risk effect because, as we argue in the remainder of the paper, in a dynamic asset pricing model, the tilt $b(y^s) - b(y^n)$ is related to the change in endogenous return volatility and co-movement due to the change in investors' aggregate risk-exposure. As we will explain, this is akin to the amplification effect in standard heterogeneous agent models in macroeconomics and finance (e.g. Longstaff and Wang, 2012; He and Krishnamurthy, 2013; Brunnermeier and Sannikov, 2014; Kondor and Vayanos, 2019). We will also explain how the shift, $a(y^s) - a(y^n)$, is related to intertemporal hedging amplified by general equilibrium forces.

Importantly, we will demonstrate that even if the shock to hedgers is small, implying in equilibrium only a small tilt $b(y^s) - b(y^n)$ or a small shift $a(y^s) - a(y^n)$, the second and third components still might be sizable. This is so, because the denominator is often the same order as the numerators in the last two terms of (2) in equilibrium. Intuitively, when the supply shock changes the return covariance only marginally, it also tends to change returns at the same order of magnitude.

3 A Model of Dynamic Asset Pricing

We build a dynamic asset pricing model with multiple assets and heterogeneous investors. Our aim is to study how returns and holdings of a particular group of investors change when an exogenous shock shifts their residual supply curve. To this goal, in this section we first set up the model, and then present our main objects of interest in the static benchmark.

3.1 The Dynamic Setup and Definitions

Time is continuous and infinite, indexed by $t \in [0, \infty)$.

Assets. There is a risk-free asset that yields a constant instantaneous rate of return $r > 0$, which is taken as exogenous in the model; that is, we do not clear the goods market.[3]

There are I long-lived risky assets with cumulative cash flows D_t following

$$dD_t = \bar{D}dt + \sigma^\top dB_t, \tag{3}$$

where $\bar{D}$ is a constant $I \times 1$ vector, σ is a constant and invertible $I \times I$ matrix, and $\top$ denotes transpose. Uncertainty of cash flows is described by the I-dimensional Brownian motion B_t. Each risky asset is in fixed supply, and we denote by S the $I \times 1$ vector consisting of asset supplies measured in terms of number of shares. We let $\Sigma \equiv \sigma^\top \sigma$.

Agents. The economy is populated by two types of agents: hedgers and investors. Without loss of generality, we assume that there is a unit measure continuum of each type of agents.

For simplicity, we assume that hedgers are short-lived and maximize a mean-variance objective over their instantaneous changes in wealth. Also for simplicity, we assume that each generation leaves their wealth to the next generation without consumption.[4] That is, a hedger maximizes

$$\mathbb{E}_t[dv_t] - \frac{\alpha}{2}\mathrm{Var}_t(dv_t),$$

where v_t is the hedger's time t wealth. $\alpha > 0$ is hedgers' absolute risk aversion coefficient over wealth.

On the other hand, investors are long-lived and maximize their discounted utility over

[3] By fixing interest rate to be exogenously given as in the tradition of the asset demand system literature, our analysis represents the minimal departure from Koijen and Yogo (2019); Gabaix and Koijen (2022) and therefore sharpens our economic message.

[4] As we explain below, in our model short-lived hedgers' asset demand just serves as shifter to identify the demand slope for the long-lived investors. As a result, the exact modeling of hedgers is less important for our purpose. For instance, the alternative assumption that each generation of hedgers consume their wealth and the next generation arrives at the economy with some exogenous constant wealth would not change the analysis significantly. Kondor and Vayanos (2019) solves a version of this model with long-lived hedgers.

intertemporal consumption. That is, an investor maximizes

$$\mathbb{E}_t\Big[\int_t^\infty e^{-\rho(\tau-t)}u(\hat{c}_\tau)d\tau\Big],$$

where $\hat{c}_\tau$ is the individual investor's time τ consumption, and ρ is investors' time discount rate. We assume that $\rho > r$ to ensure that investors do not accumulate infinite wealth over time. We denote by w_t investors' aggregate time-t wealth, which is one of the state variables in this economy.

Our paper focuses on Constant Relative Risk Aversion (CRRA) investors, with the following two subcases.[5]

Case 1 (Log investors)**.** *Investors' instantaneous utility* $u(\hat{c}_\tau) = \log \hat{c}_\tau$.

Case 2 (General CRRA investors)**.** *Investors' instantaneous utility* $u(\hat{c}_\tau) = \frac{\hat{c}_\tau^{1-\gamma}}{1-\gamma}$ *for* $\gamma \neq 1$.

Hedging demand and states. Each hedger receives a random endowment $u_t^\top dD_t$ from time t to $t + dt$, where u_t is an $I \times 1$ vector and parametrizes the background risk in the economy. In our main model, u_t is modeled to follow a Markov switching process. The economy is either in a normal state or a shock state. We use $\xi \in \{n, s\}$ to indicate the state of the economy, with n denoting the normal state and s denoting the shock state. Then $u_t \in \{u^n, u^s\}$ depending on the state of the economy. Both u^n and u^s are $I \times 1$ vectors. For expositional reasons only, we assume $u^s \geq u^n$.

The state of the economy switches from the normal state to the shock state with Poisson intensity λ^n, and from the shock state to the normal state with intensity λ^s.

Equilibrium definition. Throughout we focus on symmetric equilibrium where economic agents are taking the same policy within each group (hedgers or investors). Denote by X_t^ξ and y_t^ξ, each $I \times 1$ vectors, the aggregate asset demand from hedgers and investors. The endogenous price process is P_t^ξ, also with dimension of $I \times 1$. Market clearing requires that the induced asset demand from both types of agents clear the exogenous asset supply at any time, that is,

$$y_t^\xi w_t^\xi = \mathrm{diag}(P_t^\xi)S - X_t^\xi. \tag{4}$$

It is useful to think about the right hand side as the residual supply curve from investors' point of view. Note that investors' aggregate demand y_t^ξ is defined as a share of their aggregate wealth, while X_t^ξ is the value of hedger's aggregate demand in state ξ and time t.

[5]In Appendix C, we also work out the case of CARA investors with $u(\hat{c}_\tau) = -e^{-\frac{\gamma}{r}\hat{c}_\tau}$. That specification gives analytical characterization for various of our objects of interest.

Prices and returns. We conjecture and later verify that the structure of the equilibrium price vector P_t^ξ is

$$P_t^\xi(w_t^\xi) = \frac{\bar{D}}{r} - \pi^\xi(w_t^\xi), \tag{5}$$

where $\pi^\xi(w_t^\xi)$ is an $I \times 1$ vector describing the time-varying risk premium when investors' aggregate wealth is w_t^ξ, with $\xi \in \{s, n\}$. That is, the conjecture implies that the aggregate wealth of the investors, w_t^ξ, and the shock state for hedgers, ξ, are the only two state variables of the model. As the jumps in hedging demand shocks affect equilibrium prices, we conjecture that for the endogenously determined drift μ_{Pt}^ξ, volatility σ_{Pt}^ξ, and jump $J_t^\xi \equiv P_t^{-\xi} - P_t^\xi$ we can write

$$dP_t^\xi = \mu_{Pt}^\xi dt + \sigma_{Pt}^{\xi\top} dB_t + J_t^\xi dN_t^\xi, \tag{6}$$

where $\{N_t^\xi : t \geq 0\}$ is a Poisson counting process with intensity λ^ξ. It will be useful to map the change in the price vector, dP_t^ξ to the (percentage) asset return vector defined as

$$dR_t^\xi \equiv \operatorname{diag}(P_t^\xi)^{-1}\left(dD_t + dP_t^\xi\right). \tag{7}$$

Then the endogenous (percentage) asset return vector dR_t^ξ has a drift $\mu_{Rt}^\xi \equiv \operatorname{diag}(P_t^\xi)^{-1}(\bar{D} + \mu_{Pt}^\xi)$ and volatility $\sigma_{Rt}^\xi \equiv (\sigma_{Pt}^\xi + \sigma)\operatorname{diag}(P_t^\xi)^{-1}$ defining the return covariance matrix, $\Omega_t^\xi \equiv (\sigma_{Rt}^\xi)^\top \sigma_{Rt}^\xi$ and a jump component $J_{Rt}^\xi \equiv \operatorname{diag}(P_t^\xi)^{-1}(P_t^{-\xi} - P_t^\xi)$. We restrict the parameter space to ensure that P_t^ξ is strictly positive, and hence dR_t^ξ is well defined. With a slight abuse of notation, we shall write

$$\mathbb{E}_t(R^\xi) \equiv \frac{\mathbb{E}_t(dR_t^\xi)}{dt} \quad \text{and} \quad \Delta\mathbb{E}_t(R) \equiv \frac{\mathbb{E}_t(dR_t^s)}{dt} - \frac{\mathbb{E}_t(dR_t^n)}{dt} \tag{8}$$

for the instantaneous expected return in state ξ and the change in instantaneous expected return when the shock hits.

Hedgers' demand curve and the residual supply shock Given the dynamics of returns, hedgers' wealth follows

$$\begin{aligned} dv_t^\xi = &\Big[rv_t^\xi + X_t^{\xi\top}\left(dR_t^\xi - r\mathbf{1}dt\right) + \underbrace{u^{\xi\top}\bar{D}}_{\text{expected endowment dividends}}\Big]dt + \\ &\Big[(\sigma_{Rt}^\xi)X_t^\xi + \underbrace{\sigma u^\xi}_{\text{endowment shocks}}\Big]^\top dB_t + X_t^{\xi\top} J_{Rt}^\xi dN_t^\xi, \end{aligned} \tag{9}$$

where $\mathbf{1}$ is an $I \times 1$ vector of ones. Given the hedger's wealth and their mean-variance utility, the first-order condition implies

$$X_t^\xi = \left[\Omega_t^\xi + \lambda^\xi J_{Rt}^\xi J_{Rt}^{\xi\top}\right]^{-1}\left[\frac{\mathbb{E}_t(R^\xi) - r\mathbf{1}}{\alpha} - (\sigma_{Rt}^\xi)^\top \sigma u^\xi\right]. \tag{10}$$

This is hedgers' demand. As it is apparent, when u^n switches to $u^s \geq u^n$, hedgers demand curve shifts downward. This translates to an upward shift in investors' residual supply curve $\text{diag}(P_t^\xi)S - X_t^\xi$, and we expect that the state s corresponds to positive flows from hedgers to investors in equilibrium.

A one-off residual supply shock. Instead of the general case with arbitrary λ^s, λ^n, sometimes it is convenient to focus on the case of $\lambda^n \to 0, \lambda^s = \lambda > 0$. Intuitively, this corresponds to an economy with a *one-off shock* where the shock state has a persistence of $\frac{1}{\lambda}$.

The observed slope at a one-off shock. There is a long tradition in economics to recover the slope of a demand curve of some agents by observing how quantities and prices are affected by shocks to the (residual) supply. We define a corresponding measure in our economy as follows.

Definition 1. *Consider a dynamic equilibrium with a one-off shock with persistence $\frac{1}{\lambda}$. Observed slope is defined as*

$$\mathcal{O}\left(w_t^n; \lambda, u^n, u^s\right) \equiv \frac{[y_t^s - y_t^n]_i}{[\Delta \mathbb{E}_t\left(R\right)]_i}. \tag{11}$$

We define observed slope with respect to a one-off shift in residual supply induced by the change of the state from normal to shocked. $\frac{1}{\lambda}$, captures the persistence of this shift, and once the economy shifts back to the normal state, no further shifts are expected.

Road map. The main topic of our paper is to understand whether the observed slope recovers the slope of investor's demand curve. Conceptually, we think of the latter as the representative investor's optimal quantity response per a unit change in expected returns. The analysis in the most general case of our model requires a general formal definition of the slope of investor's demand which we will provide in Section 5.2. Before that, we build up the intuition of why and when it is mismeasured by the observed slope through a series of examples and special cases.

3.2 The Static, Multi-Asset Benchmark

Just as in the dynamic model, in this static benchmark we have I assets and two groups of agents. However, following Campbell and Viceira (2002), we assume that the risky assets pay log-normally distributed dividends $\log D \sim N(\bar{D}, \Sigma)$ where S, $\bar{D}$ and D are $I \times 1$ while Σ is $I \times I$. For simplicity, we keep hedgers' demand general. That is, we assume they demand $x = c(P, u)$ share of wealth where $c(P, u) : \mathbb{R}^I \times \mathbb{R}^I \to \mathbb{R}^I$ is an arbitrary function, P is the price vector and u is a vector of demand shocks. Then, following Campbell and Viceira (2002) and Koijen and Yogo (2019) in relying on the log-linearized approximation, we write investors' demand curve in terms of portfolio shares as

$$\hat{y} = \frac{1}{\gamma}\Sigma^{-1}\left(\mathbb{E}(R) - r\mathbf{1} + \frac{1}{2}\operatorname{diag}(\Sigma)\right) \tag{12}$$

where $\mathbb{E}(R) = \bar{D} - \log P$ is the expected log gross return vector and γ is their constant relative risk-aversion.

Comparing (12) to (1), one could map $b(y)$ to the matrix $\frac{1}{\gamma}\Sigma^{-1}$, and $a(y)$ to the matrix $\frac{1}{2\gamma}\Sigma^{-1}\operatorname{diag}(\Sigma)$. Importantly, in this static benchmark, both of these depend on deep parameters only, hence they are not sensitive to investors' aggregate position in equilibrium.

Static demand slope $\mathcal{C}^{\text{static}}$. However, for didactic purposes, we map the vector-to-vector slope $\frac{1}{\gamma}\Sigma^{-1}$, into an asset specific slope as follows. Suppose an investor is asked how much she would increase her demand for asset i if the vector $\mathbb{E}(R)$ would change to $\mathbb{E}(R) + \delta$ where δ is an arbitrary vector. The static slope will measure the response by comparing the change in the investor's portfolio share to the change in the implied return, that is

$$\mathcal{C}^{\text{static}}(\delta) \equiv \frac{\left[\hat{y}^h - \hat{y}\right]_i}{[\delta]_i} = \frac{1}{\gamma}\left[\Sigma^{-1}\right]_{ii} + \frac{1}{\gamma}\frac{\sum_{j \neq i}\Sigma^{-1}_{ij}\,[\delta]_j}{[\delta]_i}, \tag{13}$$

where $\hat{y}^h$ is the investor's portfolio share under the hypothetical.[6]

The first element in (13) is the own-slope characterizing the investor's response to a change in the return of the same asset only. The second element is due to the cross-sensitivities $\frac{1}{\gamma}\Sigma^{-1}_{ij}$ characterizing the response due to changes in the returns of other assets. If only δ_i were different from 0, or if cash-flows were uncorrelated, or if there were a single asset, the inverse of return variance (scaled by risk-tolerance) would describe the slope of the demand curve. However, for a complete characterization of static slope the full slope matrix of $\frac{1}{\gamma}\Sigma^{-1}$ is required.

[6]Note that for this static benchmark, whether the source of the change in return is a drop in prices or a hike in expected dividends is immaterial.

One important take-away from the static benchmark demand slope in (13), is that while $\mathcal{C}^{\text{static}}(\delta)$ depends on the direction of the change in returns δ, the demand curve as a function of δ is independent of the demand from the other agents (i.e., hedgers).

Note that we define the slope as the quantity response to a unit change of expected return, not to a unit change of price. In a static setting, these two definitions would coincide, because with exogenous liquidating dividends a unit drop of today's price translates one to one to a unit increase of expected return by holding one share of asset today. However, as we will see, in a dynamic setting, the persistence of shocks and future risk matters, hence this translation no longer holds. Because researchers in financial economics think of demand for an asset as a function of its return characteristics, we believe this distinction is conceptually important.

Also, in contrast with the empirical literature, we do not normalize our measure to transform the slope to an elasticity. As we explain in HHHKL, this is so, because *natural units* differ across modeling approaches. For instance, in the logit framework of Koijen and Yogo (2019), a percentage change in demand as a function of a percentage change in return gives the most canonical form for the demand curve, while in our CRRA framework the portfolio share as a function of percentage (or log) returns does.

Static observed slope $\mathcal{O}^{\text{static}}$. Now we can compare $\mathcal{C}^{\text{static}}(\delta)$ to the equilibrium effect of a shift in residual demand on the position of investors, the observed slope. Demand shocks change the equilibrium through the market clearing condition

$$y + x = \text{diag}(P)S,$$

where we normalized the wealth of each group to 1. Hence, $\Delta\mathbb{E}_t(R) \equiv (\bar{D} - \log P^s) - (\bar{D} - \log P^n)$ is the change in expected (log) return due to the shock.

As a shift in u leaves $-\frac{(1+r)}{\gamma}\Sigma^{-1}$ unchanged, using (12), we have

$$\mathcal{O}^{\text{static}}(u^n, u^s) \equiv \frac{[y^s - y^n]_i}{\Delta\mathbb{E}(R)_i} = \frac{1}{\gamma}\left(\left[\Sigma^{-1}\right]_{ii} + \frac{\sum_{j\neq i}\Sigma^{-1}_{ij}\left[\Delta\mathbb{E}(R)\right]_j}{\left[\Delta\mathbb{E}(R)\right]_i}\right). \tag{14}$$

Comparing (13) and (14) shows that the observed slope is identical to the static slope with

respect to a carefully chosen δ.[7] Indeed,

$$\mathcal{O}^{\text{static}}\left(u^n, u^s\right) = \mathcal{C}^{\text{static}}(\delta)|_{\delta=\Delta\mathbb{E}(R^s)}. \tag{15}$$

Therefore a shift in residual supply recovers the slope of static demand.

The main insight of the paper is that this principle does not hold in a general dynamic framework.[8]

4 Log Investors: The Anatomy of the Tilt

To introduce our effects in steps, we start the analysis of our dynamic model in some special cases. We consider log investors, a one-off shock in residual supply, and two extreme cases: when the shocked state is fully persistent, $\lambda \to 0$, and when it is fully transitory, $\lambda \to \infty$. In the former, after the shift in hedgers' demand, investors expect to remain in the shock state indefinitely. In the latter, investors expect the shock state to last only for an instant and return to the normal state almost immediately. These cases not only provide simplified intuition for the tilt but also allows for some analytical results.

4.1 Equilibrium with Log Investors

Before we study the limiting economy with one-off shock, we derive the equilibrium under general λ^n, λ^s, following standard arguments. The details are in Appendix A.1.1, here we provide a draft only. We need to distinguish between an investor's individual wealth $\hat{w}_t$ and investors' aggregate wealth w_t. Each individual controls $\hat{w}_t$, which coincides with the investor sector's aggregate wealth w_t only in equilibrium. In contrast, w_t serves as a state variable in this model, which each individual takes as given. The conjectured value function, which

[7] Fuchs et al. (2025) refers to the second element in the bracket of expression (14) as the spill-over. One of the main insight from their paper is that in general $\mathcal{C}^{\text{static}}(\delta) \neq \mathcal{O}^{\text{static}}$ unless $\delta = k \cdot \mathbf{1}_i$, which is proportional to i^{th} element of unit vector (k is a scalar). That is, observed elasticity is not own price elasticity because supply shocks affect all expected returns. Our requirement of $\delta = \Delta\mathbb{E}\left(R^s\right)$ in (15) represents the implicit assumption that the econometrician understands this point, hence all the effects we discuss in the rest of the paper are over and above of those originating from this spillover.

[8] Note that most empirical studies would not attempt to identify each element of the slope matrix $\frac{1}{\gamma}\Sigma^{-1}$. Instead, as HHHKL explain, there are implicit or explicit assumptions on the structure of the slope matrix helping to identify some combination of elements of interest. For instance, studies based on cross-sectional variation typically identify relative elasticity, the response in demand for one asset relative to another one with the same observables to a change in the relative price of these assets. HHHKL also explain how the rest of the matrix might be recovered using time-series variation. In this paper, we abstract away from the implementation problems arising from the dimensionality of the slope matrix and focus only on the higher level question: what if this matrix is not invariant to supply shocks?

depends on both $\hat{w}_t^\xi$ and w_t^ξ, is

$$V^\xi(\hat{w}_t^\xi, w_t^\xi) = \frac{1}{\rho}\log \hat{w}_t^\xi + \bar{q}^\xi(w_t^\xi). \tag{16}$$

As it is conventional with power utility, we define an investor's demand in terms of her chosen portfolio share, $\hat{y}_t^\xi$. Let y_t^ξ be the aggregate counterpart of $\hat{y}_t^\xi$, and the corresponding consumption variables are c_t^ξ and $\hat{c}_t^\xi$, respectively. Then each investor's wealth dynamics is

$$\begin{aligned} d\hat{w}_t^\xi &= r\hat{w}_t^\xi dt + \hat{w}_t^\xi \hat{y}_t^{\xi\top}\Big(dR_t^\xi - r\mathbf{1}dt\Big) - \hat{c}_t^\xi dt \\ &= \Big[(1-\hat{y}_t^{\xi\top}\mathbf{1})r\hat{w}_t^\xi - \hat{c}_t^\xi + \hat{w}_t^\xi \hat{y}_t^{\xi\top}\mu_{Rt}^\xi\Big]dt + \hat{w}_t^\xi \hat{y}_t^{\xi\top}(\sigma_{Rt}^\xi)^\top dB_t + \hat{w}_t^\xi \hat{y}_t^{\xi\top} J_{Rt}^\xi dN_t^\xi. \end{aligned} \tag{17}$$

From (17), the total wealth dynamic for investors aggregates to

$$dw_t^\xi = \Big[(1-y_t^{\xi\top}\mathbf{1})rw_t^\xi - c_t^\xi + w_t^\xi y_t^{\xi\top}\mu_{Rt}^\xi\Big]dt + w_t^\xi y_t^{\xi\top}(\sigma_{Rt}^\xi)^\top dB_t + w_t^\xi y_t^{\xi\top} J_{Rt}^\xi dN_t^\xi, \tag{18}$$

which gives the drift and diffusion of aggregate wealth, $\mu_t^{w,\xi}$ and $\sigma_t^{w,\xi}$. Then, substituting the conjectured value function into the representative investor's HJB equation and deriving the first-order conditions give the standard consumption rule

$$\hat{c}_t^\xi = \rho\hat{w}_t^\xi, \tag{19}$$

and the demand function for risky assets. The next proposition spells out the latter for the following two cases of residual supply shock: i) one-off and fully persistent with $\lambda \to 0$; and ii) fully transitory with $\lambda \to \infty$.[9]

Proposition 1. *When investors have log utility and there is a one-off residual supply shock, both in the limits $\lambda \to 0$ and $\lambda \to \infty$, the optimal portfolio policy of investors has the form of*

$$\hat{y}_t^\xi = (\Omega_t^\xi)^{-1}(\mathbb{E}_t(R^\xi) - r\mathbf{1}). \tag{20}$$

Note the simple form of investors' demand function (20). Similarly as we noted for our static benchmark, the slope of investors' demand is captured by the inverse of the covariance matrix of returns. Indeed, analogously to the static case we can define the slope of demand of the log investors as follows. Suppose in state n, the investor is asked how much she would increase her demand for asset i if the vector $\mathbb{E}_t(R)$ would change to $\mathbb{E}_t(R) + \delta$ where δ is an arbitrary vector. The log slope will measure the response by comparing the change in the

[9]The case of $\lambda \to 0$ is easy to understand since in this limit it is as if the economy never reverts back to the normal state from the shocked state and vice-versa. The case of $\lambda \to \infty$ is a bit more subtle, as it involves the equilibrium property of the price path; for some intuition, see footnote 14.

investor's portfolio share to the change in the implied return, that is

$$\mathcal{C}^{\log}(w_t^n;\delta) \equiv \frac{\left[\hat{y}_t^h - \hat{y}_t^n\right]_i}{[\delta]_i} = \left[(\Omega_t^n)^{-1}\right]_{ii} + \frac{\sum_{j\neq i}\left[(\Omega_t^n)^{-1}\right]_{ij}[\delta]_j}{[\delta]_i}, \tag{21}$$

where $\hat{y}_t^h$ is the investor's portfolio share under the hypothetical scenario with a higher expected return $\mathbb{E}_t(R)+\delta$. An inconsequential difference between (13) and (21) is that log investors have a constant relative risk-aversion of 1, hence the γ parameter can be omitted. A consequential difference is that the exogenous Σ, which is the covariance matrix of the logarithm of date-1 dividends as in the setting of Campbell and Viceira (2002), is replaced by the endogenous Ω_t^{ξ}, the covariance of equilibrium asset returns. To see the implication, let us write the observed slope as

$$\begin{aligned}\mathcal{O}^{\log}(w_t^n;u^s,u^n) &\equiv \frac{[y_t^s - y_t^n]_i}{[\Delta\mathbb{E}_t(R)]_i} \\ &= \mathcal{C}^{\log}(w_t^n;\delta)|_{\delta=\Delta\mathbb{E}_t(R)} + \frac{\left[\left((\Omega_t^s)^{-1} - (\Omega_t^n)^{-1}\right)\left(\mathbb{E}_t(R^s) - r\mathbf{1}\right)\right]_i}{[\Delta\mathbb{E}_t(R)]_i},\end{aligned} \tag{22}$$

where we have used (20).[10] Unlike in the static case, in our dynamic equilibrium with log investors (either perfectly persistent shocks $\lambda \to 0$ or transitory shocks $\lambda \to \infty$), shifts in residual supply does not recover the slope of investors' demand, $\mathcal{C}^{\log}$. The difference stems from the fact that when hedgers dump assets on investors, the persistent change in the allocation of risk changes the covariance matrix of returns. Hence, the change in residual supply tilts investors' demand.[11] Just as in equation (2), the resulting second term in (22) is the endogenous risk effect.

To build more intuition on the sources and direction of the tilt, we consider two examples. First, we consider the case of a single risky asset. In the second example, we consider multiple assets out of which the shift in residual supply affects only two, but in exactly opposite ways. As it turns out, for the first case the $\lambda \to \infty$ limit, while for the second case, the $\lambda \to 0$ limit admits sharp analytical characterization of the tilt.

[10] In the Appendix A.2 we show that $\lim_{\lambda\to\infty}\Delta\mathbb{E}_t(R) > 0$ hence the second term in (22) is bounded in this limit.

[11] Note that our analysis on log investors provides a sharp comparison to Koijen and Yogo (2019) who also consider (myopic) log investors. The difference stems from the fact that their Assumption 1 assumes away the possibility that the residual supply shock affects the volatility of assets. In contrast, we show that general equilibrium forces imply that this is likely to be the case.

4.2 Single Risky Asset

With a single risky asset and log investors, the slope of demand $\mathcal{C}^{\log}$ simplifies to inverse of return variance, $\frac{1}{(\sigma^{\xi}_{Rt})^2}$. In state s, when hedgers demand less of the risky asset for given expected return, investors hold a riskier portfolio. Therefore, we expect the volatility of their wealth to increase, which, through (5), is expected to increase the volatility of the risky asset. This should imply that the observed slope, $\mathcal{O}^{\log}$ to be smaller than $\mathcal{C}^{\log}$.

It is tempting to conjecture that at least when the hedging shock is extremely short-lived, so that $\lambda \to \infty$, this effect diminishes. The following lemma proves that this is never the case: $\mathcal{O}^{\log}$ is strictly smaller than $\mathcal{C}^{\log}$ even in this limit.

Proposition 2. *Consider an economy with log investors, single risky asset $I = 1$, and the demand shift shock is transitory $\lambda \to \infty$. In this economy, the variance of the asset increases, $\sigma^n_{Rt} < \sigma^s_{Rt}$, hence the endogenous risk effect is negative implying $\mathcal{O}^{log}(w^n_t; u^n, u^s) < \mathcal{C}^{log}(w^n_t; \delta)|_{\delta=\Delta\mathbb{E}_t(R)}$. Furthermore, the endogenous risk effect is bounded away from zero even if the residual supply shock is diminishingly small:*

$$\lim_{u^n \to u^s} \left(\mathcal{O}^{log}(w^n_t; u^n, u^s) - \mathcal{C}^{log}(w^n_t; \delta)|_{\delta=\Delta\mathbb{E}_t(R)} \right) < 0.$$

Proposition 2 states that in our economy, as hedgers liquidate some of their positions due to the shock, the endogenous return volatility goes up. Hence, the endogenous risk effect is negative and the observed slope is flatter than its conceptual counterpart. The economic intuition is closely related to the well-known amplification effect in various heterogeneous agent models (e.g. Longstaff and Wang, 2012; He and Krishnamurthy, 2013; Brunnermeier and Sannikov, 2014; Kondor and Vayanos, 2019).[12] Intuitively, if investors hold more of the risky asset in equilibrium due to hedgers' shock, $y^s_t > y^n_t$, it implies that a future adverse shock, $dD_\tau < 0$ results in a larger drop in investors' aggregate wealth w_τ, increasing risk premium more; and this leads to a larger decrease in price, which amplifies the initial drop in aggregate wealth. As this amplification effect is stronger when $y^s_t > y^n_t$, the shock implies a volatility hike: $\sigma^n_{Rt} < \sigma^s_{Rt}$.

To see the argument formally, from return volatility definition (equations (6)-(7)) we have $\sigma^{\xi}_{Rt} = (\sigma + \sigma^{\xi}_{Pt})/P^{\xi}_t$, where the first (second) term captures the exogenous (endogenous) volatility of Brownian dividends (asset price). The amplification mentioned above applies to the second component regarding the endogenous price volatility σ^{ξ}_{Pt} of the risky asset, which

[12] In He and Krishnamurthy (2013) and Brunnermeier and Sannikov (2014) intermediaries play the role of investors. In He and Krishnamurthy (2013) the amplification is due to the endogenous equity constraint induced by moral hazard (of intermediaries). To see the parallel between Brunnermeier and Sannikov (2014) and our set up, consider the $\gamma \to 0$ limit in our case. Then, in both models intermediaries/investors are the efficient holders of risk. However, as they foresee that they cannot operate when their wealth is 0, they are willing to hold less and less risk as their capital shrinks. (For a larger γ this effect is stronger by risk-aversion.) This is why, the two models imply similar patterns of endogenous risk and amplification.

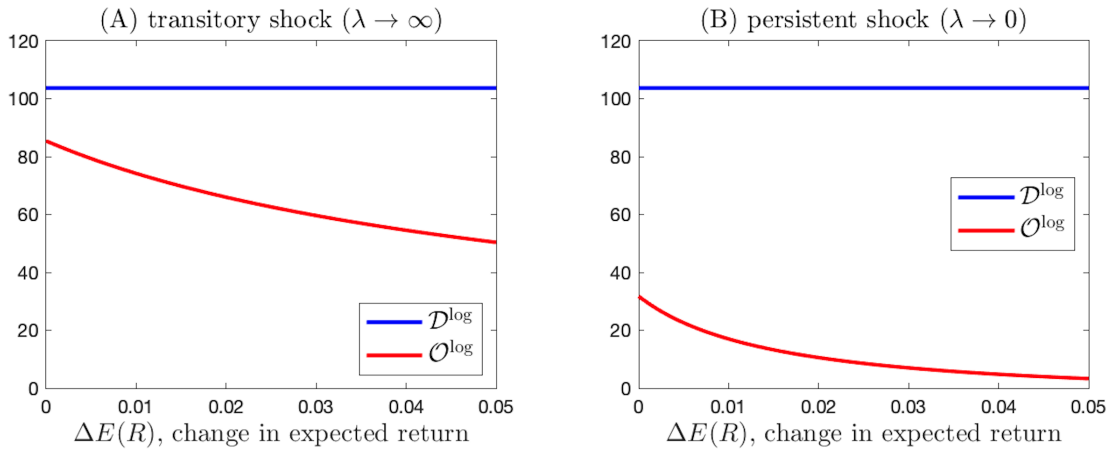


Figure 1: Demand slope of a log investor, $\mathcal{C}^{\log}$, and the corresponding observed slope, $\mathcal{C}^{\log}(\delta)|_{\delta=\Delta\mathbb{E}_t(R)}$ with one risky asset, and log investors. Panel A plots $\mathcal{C}^{\log}$ and $\mathcal{O}^{\log}$ for a transitory shock ($\lambda \to \infty$), and Panel B plots the same objects for a persistent shock ($\lambda \to 0$). We change the u^s shock and plot our measures against the implying particular change in percentage return $\Delta\mathbb{E}_t(R)$.

is determined in equilibrium by (using Ito's lemma on (5) and matching volatility terms):

$$\underbrace{\sigma_{Pt}^{\xi}}_{\text{price volatility}} = \underbrace{P^{\xi\prime}(w_t)}_{>0\text{, price loading on wealth}} \times \underbrace{\overbrace{\text{Investor Holdings}_t}^{\text{increasing in } y_t^{\xi}} \cdot \left(\sigma + \sigma_{Pt}^{\xi}\right)}_{\text{wealth } w_t\text{'s exposure to risky asset}}. \tag{23}$$

This immediately implies that the price volatility of the risky asset increases when investors increase their holdings. Finally, because the endogenous risky asset holdings of investors across two states are determined by the nature of u shocks and differ even when $\lambda \to \infty$, while risky asset prices converge in this limiting case, the return volatility of the risky asset remains strictly higher in the shock state than in the normal state. This delivers our desired negative endogenous risk effect.[13]

We finish this section with an illustration of the difference due to this effect in the fully persistent and fully transitory cases. Panel A and B of Figure 1 show $\mathcal{O}^{\log}$ and $\mathcal{C}^{\log}(\delta)|_{\delta=\Delta\mathbb{E}_t(R)}$ for various values of u^s (putting the resulting $\Delta\mathbb{E}_t(R)$ on the x-axis) for these two cases. While the difference between $\mathcal{O}^{\log}$ and $\mathcal{C}^{\log}$ is significantly lower when the shock is transitory, in line with Proposition 2, it is always negative and clearly bounded away from 0 even for small shocks.

[13] We show in the Appendix that when $\lambda \to \infty$, we have $P^s(\cdot) \to P^n(\cdot)$ and $w^n \to w^s$; intuitively, price path should be continuous under immediate (and almost certain reversal). However, the endogenous risky asset holdings of investors differ across two states even when $\lambda \to \infty$. Because $P^{n\prime}(w) \to P^{s\prime}(w) > 0$, the return volatility of the risky asset is strictly higher in the shock state than in the normal state.

4.3 Multiple Assets: Index Inclusion

Next, we consider a multi-asset example, to see how moving from return variance to the covariance matrix affects our intuition.

We picture a market with a large number of fundamentally identical assets separated into two groups: red and blue. The only difference between these assets is in the corresponding component of hedgers' demand shifter, u^{ξ}; that is, hedgers hold red assets in larger proportion due to the difference in their background risk. We imagine that when the state switches from n to s, hedgers treat the first asset $i = 1$ (which was red) as they used to treat the last asset $i = I$ (which was blue) and vice-versa. The shock does not affect any other asset. Essentially, in state s hedgers demand decreases for the first asset but increases for the last asset.

More precisely, suppose that $\Sigma = \sigma^2 \mathbf{I}$, $S = s\mathbf{1}$, $\bar{D} = \bar{d}\mathbf{1}$ where $\mathbf{I}$ is the identity matrix, and s, σ^2, and $\bar{d}$ are arbitrary scalars. The demand shifter for hedgers are modeled as

$$[u^n]_i = \begin{cases} \upsilon & \text{for } i < I_1, \\ \upsilon + \Delta\upsilon & \text{for } i \geq I_1, \end{cases}$$

and

$$[u^s]_i = \begin{cases} \upsilon + \Delta\upsilon & \text{for } i = 1, \\ \upsilon & \text{for } i = I, \\ [u^n]_i & \text{otherwise,} \end{cases}$$

with υ and $\Delta\upsilon$ being some positive constants. Here, the first I_1 assets are initially blue.

To sharpen our theoretical results, we consider the limit when the difference in hedgers' exposure to blue and red assets, and consequently, the change in their demand captured by $\Delta\upsilon$ in (4.3), is diminishingly small. One might expect that the negligible effect of this small local shock on the equilibrium outcome implies that the informativeness of shifts to residual supply should perhaps be restored. We show that this is not the case. The next lemma formally defines the example and states the result.

Proposition 3. *Consider the limiting economy when $\lambda \to 0$. In this economy, for sufficiently small $\Delta\upsilon$ the variance of the first asset and the covariance between the first asset and any other asset increases due to the shock increases as*

$$\lim_{\Delta\upsilon \to 0} \frac{[\Omega_t^s - \Omega_t^n]_{1i}}{\Delta\upsilon} = \bar{F}(w_t)\,\sigma^2 \begin{cases} 2\frac{\bar{D}}{r}\frac{(S+\upsilon)^2}{(P_t)^3} & \textit{for } i = 1, \\ \frac{\bar{D}}{r}\frac{s+\upsilon}{(P_t)^3} & \textit{for } I > i > 1, \\ 0 & \textit{for } i = I. \end{cases}$$

where $\bar{F}(\cdot)$ is a scalar function defined in the Appendix. The endogenous risk effect is negative

so that $\mathcal{O}^{log}(w_t^n; u^n, u^s) < \mathcal{C}^{log}(w_t^n; \delta)|_{\delta=\Delta\mathbb{E}_t(R)}$*, and the effect is strictly bounded away from zero.*

The first notable result of Lemma 3 is that the diminishingly small shock does not lead to diminishingly small difference in $\mathcal{C}^{\log} - \mathcal{O}^{\log}$; this is because both the denominator and the numerator of the second term in (22) is of order $\Delta\upsilon$. That is, while small shocks lead to small changes in equilibrium objects as expected, the mismeasurement depends on the ratio of the effect on the covariance and on expected returns.

We also state that just as in the single asset case, the observed slope is flatter than its conceptual counterpart. It is because as hedgers liquidate their positions in the first asset, both its endogenous return volatility and its endogenous covariance with all the unaffected assets go up. The intuition is similar to the single asset case, because in our model the endogenous volatility of all assets is due to investors' wealth shocks. As a result, when the first asset becomes blue, hence investors hold more of it, not only its volatility increases but also its comovement with every other asset.

5 The General Case: Intertemporal Hedging and Persistence

Now we turn to the general model with CRRA utility and residual supply shocks with arbitrary persistence, $1/\lambda$. First, we summarize the main properties of the equilibrium. Then we show how the investor's first-order condition defines a general concept of the slope of investors' demand, the *dynamic slope*. Finally, we show that when the residual supply shock hits, investor's demand curve not only tilts due to the endogenous risk effect, but also shifts due to additional effects stemming from general equilibrium forces connected to intertemporal hedging and the persistence of the shock.

5.1 Equilibrium in the General Case

We derive the equilibrium in our multi-asset dynamic model. As main ingredients, we already defined the conjectured price process, hedgers' demand, and investor's individual and aggregate wealth in equations (6) and (10) and (17)-(18).

Investors' demand. Given the dynamics of the individual and aggregate wealth of investors, an investor's value function in state $\xi \in \{n, s\}$ satisfies the Hamilton-Jacobi-Bellman (HJB) equation which we write in a general form:

$$\begin{aligned}\rho V^{\xi} = \max_{\hat{c}_t^{\xi}, \hat{y}_t^{\xi}} : \Big\{ & u(\hat{c}_t^{\xi}) + V_{\hat{w}}^{\xi} \mu_t^{\hat{w},\xi} + \frac{1}{2} V_{\hat{w}\hat{w}}^{\xi} (\sigma_t^{\hat{w},\xi})^{\top} \sigma_t^{\hat{w},\xi} + V_w^{\xi} \mu_t^{w,\xi} + \frac{1}{2} V_{ww}^{\xi} (\sigma_t^{w,\xi})^{\top} \sigma_t^{w,\xi} \\ & + V_{\hat{w}w}^{\xi} (\sigma_t^{\hat{w},\xi})^{\top} \sigma_t^{w,\xi} + \lambda^{\xi} (V^{-\xi,+} - V^{\xi}) \Big\}.\end{aligned} \tag{24}$$

Here, V^{ξ} stands for the value function $V^{\xi}(\hat{w}_t^{\xi}, w_t^{\xi})$; $V^{-\xi,+}$ is the value function evaluated after the jump, with the superscript + denoting post-jump while $-\xi$ denoting the state after the jump:

$$V^{-\xi,+} \equiv V^{-\xi}\left(\hat{w}_t^{\xi}\left(1+\hat{y}_t^{\xi\top}J_{Rt}^{\xi}\right), w_t^{\xi}\left(1+y_t^{\xi\top}J_{Rt}^{\xi}\right)\right). \tag{25}$$

Denote $\Omega_t^{\xi} \equiv (\sigma_{Rt}^{\xi})^{\top}(\sigma_{Rt}^{\xi})$ to be the covariance matrix of returns. The first-order condition in $\hat{y}_t^{\xi}$ gives the familiar components of asset demand as in Merton (1971):

$$\hat{y}_t^{\xi} = \underbrace{\overbrace{\frac{V_{\hat{w}}^{\xi}}{\hat{w}_t^{\xi}(-V_{\hat{w}\hat{w}}^{\xi})}}^{\text{risk tolerance}}(\Omega_t^{\xi})^{-1}\left(\mathbb{E}_t(R^{\xi}) - r\mathbf{1}\right)}_{\text{myopic}} + \underbrace{\frac{V_{\hat{w}w}^{\xi}}{\hat{w}_t^{\xi}(-V_{\hat{w}\hat{w}}^{\xi})}w_t^{\xi}y_t^{\xi}}_{\text{hedging of Brownian,}} + \underbrace{\frac{1}{\hat{w}_t^{\xi}}\frac{V_{\hat{w}}^{-\xi,+}-V_{\hat{w}}^{\xi}}{(-V_{\hat{w}\hat{w}}^{\xi})}(\Omega_t^{\xi})^{-1}\lambda^{\xi}J_{Rt}^{\xi}}_{\text{hedging for reversal, } J_{Rt}}. \tag{26}$$

The first part, which is myopic demand, is analogous to the static demand curve: the instantaneous expected return scaled by the product of the inverse of the covariance matrix and risk-tolerance. One could also map the term $b(x)$ in (1) in our example to $\frac{V_{\hat{w}}^{\xi}}{\hat{w}_t^{\xi}(-V_{\hat{w}\hat{w}}^{\xi})}(\Omega_t^{\xi})^{-1}$. As we will see, the shock to residual supply affects this term through the endogenous covariance matrix. This endogenous risk effect is responsible for the tilt of the demand curve.

The second part is the intertemporal hedging demand with respect to the diffusion part of the aggregate state variable, w_t. When w_t affects the marginal value of a dollar, its correlation with the investor's individual wealth $\hat{w}_t$ affects her demand. The last part is a similar term but it represents hedging demand with respect to the reversal. When the residual supply curve reverts, prices jump, affecting investor's marginal value of a dollar and their individual wealth at the same time. Hedging this affects her demand. One could map the term $a(y)$ in (1) to these two hedging terms. As we will see, these terms are affected by the residual supply shock in a complex way, resulting in a potential shift of the demand curve.

Solving for equilibrium. We solve for the equilibrium prices after imposing that individual positions equal the aggregate position $\hat{y}_t^{\xi} = y_t$ and market clearing for each instant

$$X_t^{\xi} + w_t^{\xi}y_t^{\xi} = \operatorname{diag}(P_t^{\xi})S. \tag{27}$$

The next proposition spells out the resulting dynamic demand function and the structure of equilibrium for the cases of log investors and CRRA investors with arbitrary state switching intensities $\{\lambda^s, \lambda^n\}$.

Proposition 4. *In Case 1 (log investors), the value function has the form of*

$$V^{\xi}(\hat{w}_t^{\xi}, w_t^{\xi}) = \frac{1}{\rho}\log \hat{w}_t^{\xi} + \bar{q}^{\xi}(w_t^{\xi}). \tag{28}$$

The optimal consumption-portfolio policies are

$$\hat{c}_t^{\xi} = \rho \hat{w}_t^{\xi}, \tag{29}$$

$$\hat{y}_t^{\xi} = (\Omega_t^{\xi})^{-1}\Big(\mathbb{E}_t(R^{\xi}) - r\mathbf{1}\Big) + \lambda^{\xi}(\Omega_t^{\xi})^{-1}\left(\frac{1}{1+\hat{y}_t^{\xi\top} J_{Rt}^{\xi}} - 1\right) J_{Rt}^{\xi}. \tag{30}$$

The ODE that determines $\bar{q}^{\xi}(w_t^{\xi})$ is in the Appendix.

In Case 2 (CRRA investors), the value function has the form of

$$V^{\xi}(\hat{w}_t^{\xi}, w_t^{\xi}) = \frac{(\hat{w}_t^{\xi})^{1-\gamma}}{1-\gamma} q^{\xi}(w_t^{\xi}), \tag{31}$$

where $q^{\xi}(w_t^{\xi})$ is the marginal value of wealth in state w_t^{ξ}. The optimal consumption-portfolio policies are

$$\hat{c}_t^{\xi} = q^{\xi}(w_t^{\xi})^{-\frac{1}{\gamma}} \hat{w}_t^{\xi}; \tag{32}$$

$$\hat{y}_t^{\xi} = \frac{1}{\gamma}(\Omega_t^{\xi})^{-1}\left(\mathbb{E}_t\left(R^{\xi}\right) - r\mathbf{1}\right) + \frac{1}{\gamma}\frac{q^{\xi\prime}(w_t^{\xi})}{q^{\xi}(w_t^{\xi})} w_t^{\xi} y_t^{\xi} + \frac{\lambda^{\xi}}{\gamma}(\Omega_t^{\xi})^{-1}\left(\frac{q^{-\xi}\left(w_t^{\xi}\left(1+y_t^{\xi\top} J_{Rt}^{\xi}\right)\right)}{q^{\xi}(w_t^{\xi})[1+\hat{y}_t^{\xi\top} J_{Rt}^{\xi}]^{\gamma}} - 1\right) J_{Rt}^{\xi}. \tag{33}$$

The ODEs determining $q^{\xi}(w_t^{\xi})$ are given in the Appendix.

Note that investors' demand in each of our special cases follows the general structure of (26). However, (30) illustrates that in the simpler log case there is no intertemporal hedging term for the Brownian shock. This is a well-known property pointed out by Merton (1971). However, the intertemporal hedging term for the reversal remains. [14]

Importantly, because of the form of the equilibrium pricing vector (5), investors' portfolio choice changes not only with the state of the residual supply curve, ξ, but also with their aggregate wealth w. Poorer investors are willing to absorb less risk. As a result, aggregate investors' wealth affects the risk premium.

[14] Equation (30) also echos the simple optimal myopic portfolio decision in Proposition 1; the subtlety of $\lambda \to \infty$ case is that in Appendix XXX we prove the associated return jump shrinks to zero, so that J_{Rt} $\lim_{\lambda\to\infty} \lambda J_{Rt}$ remains finite and the term inside the bracket (with the order of $\hat{y}_t^{\xi\top} J_{Rt}^{\xi}$) converges to zero.

5.2 Dynamic Slope and Myopic Slope

How should we think about the slope of our investors' demand function in this general dynamic environment? In a static benchmark with a single asset, the slope, which measures the response in the asset holding to a unit change in its return, is a one dimensional object. As (12) illustrates, with multiple assets, a matrix describes how each component of the portfolio changes as the returns of assets do.

The general dynamic environment is more complex. As the investment opportunity set can predictably fluctuate, the investor will consider all current instantaneous (expected) returns as well as the distribution of future return paths when making her portfolio decision. Conceptually, one should measure the response in $\hat{y}_t$ for any marginal change at any point of these return paths. This is an infinite-dimension object.

The power of the Merton approach comes from summarizing the joint-distribution of all future price paths into a small number of simpler economic objects. As equation (33) illustrates, apart from the expected return vector and covariance matrix, the CRRA investor's demand also responds to changes in the future investment opportunity sets in all periods $\tau > t$. In our structure, these changes can be decomposed to two parts. On one hand, the investor cares about how the future dD_τ shocks affect her investment opportunities, on the other hand, as long as the residual supply shock is not fully persistent, she cares how the eventual reversal affects future investment opportunities.

This also implies that our heuristic approach in defining the slope of the investors' demand curve for our special cases, that is, $\mathcal{C}^{\text{minimal}}$, $\mathcal{C}^{\text{static}}$, and/or $\mathcal{C}^{\log}$, will not be sufficient for our general case. When intertemporal hedging is present, the investor's response to an increase in expected returns from $\mathbb{E}_t(R)$ to $\mathbb{E}_t(R)+\delta$ will depend on how future price paths change beyond the next instant. Hence, our approach is to generalize and parametrize our slope concept in an intuitive and simple way to capture the idea that the investor's demand response is likely to depend on the persistence of the shock to expected returns, its size and relative magnitude across assets (i.e., its direction), and also the flow share of additional return versus the lump-sum when the shock reverses (i.e., its composition).

Dynamic slope. Our headline measure is *the dynamic slope* defined as follows. We then characterize it fully in our equilibrium model.

Definition 2. *[Dynamic Slope] Suppose that an investor is offered to buy any asset i with a δ_i additional instantaneous expected return where δ is an arbitrary vector. In particular, she is asked to form demand $\hat{y}_t^h$ facing the dividend process*

$$dD_t^h = \bar{D}dt + \sigma^\top dB_t + (1-\chi)\,\mathrm{diag}(P_t^n)\delta dt + \chi\frac{\mathrm{diag}(P_t^n)\delta}{\lambda}dN_t^h \tag{34}$$

instead of (3) *where* dN_t^h *is a Poisson process with intensity* λ*. After the first jump the dividend process switches to* (3) *forever. Then dynamic slope is measured as*

$$\mathcal{C}^{dyn}\left(w_t^n;\lambda,\delta,\chi\right) \equiv \frac{\left[\hat{y}_t^h - \hat{y}_t^n\right]_i}{\delta_i}. \tag{35}$$

where $\hat{y}_t^n$ *refers to the equilibrium demand in a normal state where* $\lambda^n = 0$*, and* λ, δ, χ *parameterizes the persistence, direction and composition of the shock to return path, respectively.*

The concept of dynamic slope is appealing for several reasons. First, it is simple to understand. It measures the optimal response of an investor's demand to a change in expected returns (implemented through a hypothetical dividend process) of given persistence, direction, and composition.

Second, as we argue below, it allows for a change in return path sufficiently close to its equilibrium counterpart in our dynamic setting. For parameter χ, in our model a shift in residual supply governed by a Poisson process also implies, apart from an increase in instantaneous expected returns, a higher flow return as long as residual supply is higher, and an added lump-sum at reversal. The parameter χ governs the share of the higher return which is received as a flow relative to the share received as lump sum at reversal. For parameter λ, its inverse $1/\lambda$ governs the persistence of the improvement of the investment opportunity set. In the $\lambda \to \infty$ limit, dynamic slope captures the investor's response to a hike (or drop) in expected return for an instant. In the other extreme, the $\lambda \to 0$ case captures a permanent change in expected return. Note that under our parameterization, irrespectively of λ and χ, the change in expected return is always δ.

Third, dynamic slope is closely tied to typical policy questions addressed in the growing literature of demand elasticity. For instance, the central question of quantitative easing (QE) is: to maintain the yield of certain bonds at x basis points below their non-intervention equilibrium level, how much must the Federal Reserve purchase as a fraction of outstanding bonds over a specific period? Dynamic slope provides the critical individual demand characteristic needed to answer this question: the change in the amount an investor is willing to hold in response to a change in return of given persistence.

Observe that just as our definition for the static case, dynamic slope captures the quantity response to a given change in the expected return vector, not to a given change in the price vector. In the dynamic context, the distinction is critical. Clearly, the same price shock implies a much lower expected return shock if it is persistent than if it is transitory.[15] Therefore, throughout our analysis we control for this by keeping the expected return shock constant in our comparative statics.

[15] See the pass-through effect in Davis et al. (2023) for an empirical counterpart of this observation.

Dynamic slope as an optimization problem. To calculate dynamic slope, we first solve for the equilibrium as described in Proposition 4 with a constant residual supply curve, $u = u^n, \lambda^n = 0$. This defines the portfolio holdings, $\hat{y}^n$, and value function $V^\xi(\hat{w}_t^\xi, w_t^\xi)$, when the economy is under the baseline normal state. Also, the equilibrium objects μ_t^w, σ_t^w represent the dynamics of the aggregate state variable which is not affected by the hypothetical dividend distribution (because the extra dividends are offered only to a given investor).

Then, we introduce $V^h(\cdot,\cdot)$, the value function of that investor who is receiving the hypothetical dividends (34). We conjecture that in the CRRA case, it has the form of

$$V^h(\hat{w}_t^h, w_t^n) = q^h(w_t^n; \lambda, \delta, \chi)\frac{(\hat{w}_t^h)^{1-\gamma}}{1-\gamma}. \tag{36}$$

Note that the marginal value of a dollar $q^h(w_t^n; \lambda, \delta, \chi)^{1-\gamma}$ depends on the parameters of the dividend process. Also, the HJB equation of this problem takes into account that when the offer ends so the dividend process reverts back to its equilibrium dynamics, this investor's value function jumps back to that under normal times.

Given these objects, we solve for the optimal portfolio choice $\hat{y}^h$ of the investor facing hypothetically higher dividends. Comparing to the investor's demand in normal times $\hat{y}^n$ and substituting both into the definition of dynamic slope, (35) gives a similar decomposition of dynamic slope as in the dynamic asset demand in (26).

The following proposition summarizes the results both for the log and the CRRA cases, with details provided in the Appendix.[16]

Proposition 5. *In Case 1 (log investors), demand under hypothetical dividend path* (34) *is*

$$\hat{y}_t^h = (\Omega_t^n)^{-1}(\mu_{Rt}^n - r\mathbf{1} + \delta) + \Big[\frac{\lambda}{\lambda + \hat{y}_t^{h\top}\chi\delta} - 1\Big](\Omega_t^n)^{-1}\chi\delta. \tag{37}$$

[16]Recall that in Section 4.1 investigating $\mathcal{C}^{\log}$ for the limiting cases $\lambda \to 0$ and ∞, we have derived (21) directly from the corresponding demand curve (20) in Proposition 1. Does (38) converge to (21) in these limits? The limit of $\lambda \to \infty$, so that the term of "hedging against reversal" in (38) goes to zero, is immediate. For the $\lambda \to 0$ case, to ensure consistency we must require that $\chi = \lambda\frac{k(\lambda)}{\delta}$ with an arbitrary bounded function $k(\lambda)$. That is, formally, we are taking the limit $\lim_{\lambda\to 0}\mathcal{C}^{\text{dyn}}(w_t; \lambda, \delta, \lambda\frac{k(\lambda)}{\delta})$ for convergence to (21). As we argue in the next section, the idea is that to match the concept of $\mathcal{C}^{\text{dyn}}$ and $\mathcal{O}$ in equilibrium, we require $\chi = \lambda\frac{J_{Rt}^s}{\delta}$ to be the jump component of $\Delta\mathbb{E}_t(R)$ in equilibrium. Since $\lim_{\lambda\to 0} J_{Rt}^s$ is bounded in equilibrium, $k(\lambda)$ has to be bounded too when taking the limit of $\mathcal{C}^{\log}$.

Hence, dynamic slope has only two components:

$$\begin{aligned}\mathcal{C}^{dyn}(w_t^n;\lambda,\delta,\chi) &= \frac{[\hat{y}_t^h-\hat{y}_t^n]_i}{\delta_i}\\ &= \underbrace{[(\Omega_t^n)^{-1}]_{ii}+\frac{\sum_{j\neq i}\left[(\Omega_t^n)^{-1}\right]_{ij}[\delta]_j}{[\delta]_i}}_{\textit{myopic slope}}+\underbrace{\Big[\frac{\lambda}{\lambda+\hat{y}_t^{h\top}\chi\delta}-1\Big]\frac{[(\Omega_t^n)^{-1}\chi\delta]_i}{[\delta]_i}}_{\textit{hedging against reversal}}.\end{aligned} \tag{38}$$

In Case 2 (CRRA investors) demand under hypothetical dividend path (34) *is*

$$\begin{aligned}\hat{y}_t^h =&\frac{1}{\gamma}(\Omega_t^n)^{-1}(\mu_{Rt}^n-r\mathbf{1}+\delta)+\frac{1}{\gamma}\frac{q^{h\prime}(w_t^n;\lambda,\delta,\chi)}{q^h(w_t^n;\lambda,\delta,\chi)}w_t^n y_t^n\\ &+\frac{\chi}{\gamma}(\Omega_t^n)^{-1}\delta\left(\frac{q^n(w_t^n)}{q^h(w_t^n;\lambda,\delta,\chi)}(1+\hat{y}_t^{h\top}\frac{\chi\delta}{\lambda})^{-\gamma}-1\right),\end{aligned} \tag{39}$$

where $q^h(w_t^n;\lambda,\delta,\chi)$ is the investor's marginal value of wealth given the hypothetical D^h (with ODE for $q^h(w_t^n;\lambda,\delta,\chi)$ given in the Appendix), while $q^n(w_t^n)$ is the marginal value of wealth in equilibrium with constant u^n. The implied demand slope is as follows:

$$\begin{aligned}\mathcal{C}^{dyn}(w_t^n;\lambda,\delta,\chi) =&\frac{[\hat{y}_t^h-\hat{y}_t^n]_i}{\delta_i} = \underbrace{\frac{1}{\gamma}[(\Omega_t^n)^{-1}]_{ii}+\frac{1}{\gamma}\frac{\sum_{j\neq i}\left[(\Omega_t^n)^{-1}\right]_{ij}[\delta]_j}{[\delta]_i}}_{\textit{myopic slope}}+\\ &+\underbrace{\frac{1}{\gamma}\Big[\frac{q^{h\prime}(w_t^n;\lambda,\delta,\chi)}{q^h(w_t^n;\lambda,\delta,\chi)}-\frac{q^{n\prime}(w_t^n)}{q^n(w_t^n)}\Big]\frac{[y_t^n w_t^n]_i}{[\delta]_i}}_{\textit{Brownian hedging}}\\ &+\underbrace{\frac{1}{\gamma}\Big[\frac{q^n(w_t^n)}{q^h(w_t^n;\lambda,\delta,\chi)}\Big(1+\hat{y}_t^{h\top}\frac{\chi\delta}{\lambda}\Big)^{-\gamma}-1\Big]\frac{[(\Omega_t^n)^{-1}\chi\delta]_i}{[\delta]_i}}_{\textit{hedging against reversal}}.\end{aligned} \tag{40}$$

Dynamic slope vs myopic slope. Observe that, in the decomposition (40), myopic demand has an identical structure to the static slope (13), but with a key difference that the exogenous covariance matrix Σ is replaced with the endogenous return covariance matrix Ω^n. Therefore, we refer to this as the *myopic slope*. Note that the myopic slope depends only on the relative risk-aversion parameter and the covariance matrix of returns; it is influenced by the size, persistence, and composition of changes in the return path (i.e., the parameters δ, λ, and χ) only insofar as these parameters affect the endogenous covariance matrix.

Figure 2 illustrates the differences between the myopic and dynamic slopes, under the calibrated parameters (see the details of calibration in Appendix E.1). The blue solid line plots how much the investor would change her demand as a function of the given change

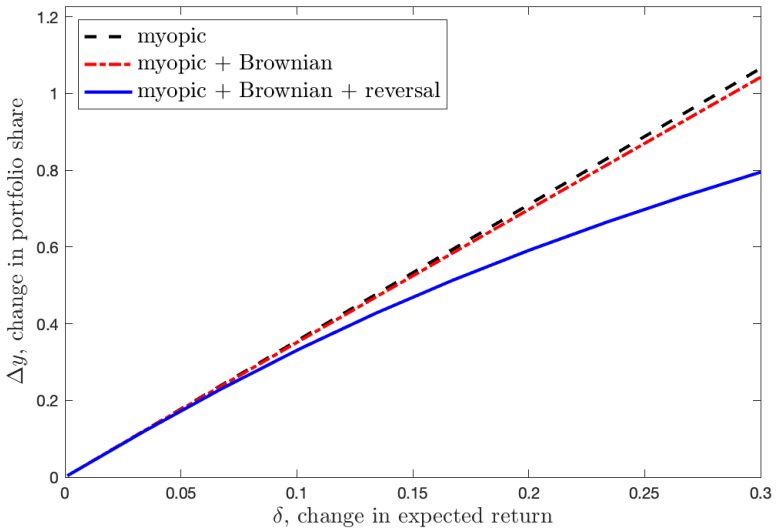


Figure 2: Change in demand as a function of change in expected return and its components. The blue solid line gives the investor's dynamic demand, which is the change in the investor's portfolio share Δy as a function of the change in the given instantaneous expected return δ. The black dashed line gives the myopic component, the red dash-dotted line gives the myopic demand plus the Brownian hedging component; so the gap between solid blue line and dash-dotted red line is the reversal hedging component. The parameter values are $\bar{D} = [0.27, 0.26]$, $\sigma = \text{diag}(0.485, 0.485)$, $\alpha = 1$, $\gamma = 3$, $r = 0.03$, $\rho = 0.08$, $s = [1, 1]$, $\lambda = 0.5$, and $\chi = 2/3$.

in instantaneous return δ in our dynamic setting. Hence, dynamic slope, $\mathcal{C}^{\text{dyn}}(w_t^n; \lambda, \delta, \chi)$ approximates the slope of this curve. We also plot the black dashed line as the corresponding myopic component given in (40); and the red dash-dotted line which gives the myopic demand plus the Brownian hedging component. Note that, just like the static demand curve $\mathcal{C}^{\text{static}}(\delta)$ in (13), the myopic component is also linear in δ given in (40).

The important take-aways from Figure 2 are that: i) the change of asset demand is non-linear so that the demand curve turns flatter for larger shocks, and ii) the curve tends to be flatter compared to the (linear) myopic demand. The difference is due to the classic Brownian and reversal hedging components à la Merton (1971), which, as shown, are both negative.[17]

The figure also illustrates that as long as the shock to expected returns, δ, is not too large, the intertemporal hedging effects tend to be small, hence the dynamic slope is close to its myopic component. This fact, and its simpler structure, might make the myopic slope an attractive alternative to dynamic slope as a concept to capture investors' demand sensitivity. Indeed, we shall demonstrate soon, that whichever of them the econometrician wish to recover based on shifts to residual supply, the problems remain the same.

Perhaps more importantly, this does not imply that the effect of intertemporal hedging

[17]For the interested reader, Appendix D discusses the economic intuition of why the Dynamic Slope tends to be flatter than its myopic counterpart.

can be ignored given the paper's main focus. In terms of our minimal example, the quantitatively small intertemporal hedging corresponds to a small $a(\cdot)$. However, as the last term in (2) illustrates, if the investor's demand curve shifts, the mismeasurment of its slope is not proportional to the absolute size of $a(\cdot)$; Instead, it depends on the ratio between the shift size and the corresponding change in expected return, $\frac{(a(y^s)-a(y^n))}{\mathbb{E}(R^s)-\mathbb{E}(R^n)}$. We will demonstrate below that this ratio can be large, even when the difference between dynamic slope and myopic slope is small.

5.3 Observed Slope versus Dynamic Slope

Per Definition 1, the observed slope of asset i is simply the ratio of the change in the representative investor's portfolio $[y^s - y^n]_i$ to the change in its expected return $[\Delta\mathbb{E}_t(R)]_i$ as the equilibrium responds to a one-off residual supply shock. With CRRA utility, the optimal portfolio policy (33) implies that the observed slope is

$$\begin{aligned}\mathcal{O}(w_t;\lambda,u^n,u^s) = \\ &= \underbrace{\frac{1}{\gamma}\frac{\left[(\Omega_t^s)^{-1}(\mathbb{E}_t(R^s)-r\mathbf{1}) - (\Omega_t^n)^{-1}(\mathbb{E}_t(R^n)-r\mathbf{1})\right]_i}{[\Delta\mathbb{E}_t(R)]_i}}_{\text{myopic observed slope}} \\ &\quad + \underbrace{\frac{1}{\gamma}\frac{\left[\frac{q^{s\prime}(w_t^s)}{q^s(w_t^s)}w_t^s y_t^s - \frac{q^{n\prime}(w_t^n)}{q^n(w_t^\xi)}w_t^n y_t^n\right]_i}{[\Delta\mathbb{E}_t(R)]_i}}_{\text{observed Brownian hedging}} + \underbrace{\frac{\lambda}{\gamma}(\Omega_t^s)^{-1}\left(\frac{q^n\left(w_t^s\left(1+y_t^{s\top}J_{Rt}^s\right)\right)}{q^s(w_t^s)[1+\hat{y}_t^{s\top}J_{Rt}^s]^\gamma}-1\right)J_{Rt}^s}_{\text{observed hedging against reversal}}.\end{aligned} \tag{41}$$

For easy comparison with Dynamic Slope, we add "observe" to the label of each components.

No we are ready to confront the main question of our paper in the general setting: to what extent an econometrician with an instrument capturing a shift in the residual supply curve can successfully measure the slope of an investor's demand curve? Under what conditions does the observed slope recover the dynamic slope? To give this exercise a fair chance, we pick i) δ to match the size and direction of the expected return shock due to the shift in supply, $\delta = \Delta\mathbb{E}_t(R)$; ii) χ to match the composition, $\chi = \frac{\lambda J_{Rt}^s}{\delta}$; and iii) λ to match the persistence of the shock.

We now show that even if we match the parameters of these objects as closely as possible, $\mathcal{C}^{\text{dyn}}$ and $\mathcal{O}$ in general differ, because $\mathcal{O}$ takes into account the adjustment of the dynamic equilibrium to the supply shock. We can trace back the difference between $\mathcal{C}^{\text{dyn}}$ and $\mathcal{O}$ to three distinct sources.

Endogenous risk effect. First, recall that the myopic component of the slope of demand is analogous to the static case: a product of risk-tolerance and the inverse of the covariance matrix of returns Ω_t. As the form (26) illustrates, both of these elements are equilibrium objects for general utility functions. In our CRRA setting, (31) implies the risk-tolerance to be an exogenous parameter $\frac{V_w(\cdot)}{wV_{ww}(\cdot)} \equiv \frac{1}{\gamma}$. Still, the covariance matrix of returns is endogenous because the shift in residual supply can change the riskiness of the asset's endogenous price and its covariance with other assets, even if cash-flows do not change.

As the slope of the dynamic demand curve is conceptually measuring how a single agent reacts to changes in returns, our measure $\mathcal{C}^{\text{dyn}}$ is not affected by these general equilibrium effects. Therefore, the myopic component of $\mathcal{C}^{\text{dyn}}$ will differ from that of $\mathcal{O}$. The following corollary states the result.

Corollary 1. *For a CRRA or log investor, the myopic slope differs from myopic observed slope because the return covariance matrix changes, i.e. there is an endogenous risk effect:*

$$\textit{myopic observed slope} - \textit{myopic slope} \;= \frac{1}{\gamma} \frac{\left[\left((\Omega_t^s)^{-1} - (\Omega_t^n)^{-1}\right)\left(\mathbb{E}_t\left(R^s\right) - r\mathbf{1}\right)\right]_i}{\left[\Delta \mathbb{E}_t(R)\right]_i}; \tag{42}$$

where $\gamma = 1$ for the log investor.

As we discussed in Section 4 right after Proposition 2, the economic intuition behind the endogenous risk effect is closely related to standard amplification effects in heterogeneous agent models. Ω_t^s differs from Ω_t^n because all investors increase their holdings of the asset when hedgers liquidate due to the shock. The investors' increased exposure to the fundamenal dividend shocks tend to imply an increased volatility (see (23)) and increased comovement with the other assets. Hence, given the inverse operator, this effects tend to be negative, making the observed slope flatter than the dynamic slope.

Note the similarity of (42) with the second term in (2) which we labeled as the endogenous risk effect due to the tilt of the demand curve. Just as in our minimal example, the size of this effect depends on the ratio of the tilt, i.e., the change in the (inverse) covariance matrix and the corresponding change in expected returns. Both of them are endogenously determined. Because of this, as we will see in our calibrated examples, the endogenous risk effect might be large, even if the change in covariance matrix (the numerator) is relatively small.

The effect of amplified intertemporal hedging. We now move on to the intertemporal hedging components, which are also different in dynamic slope and in observed slope. In particular, the difference of intertemporal hedging with respect to the Brownian shock in (40) and (41) gives the expressions in the next corollary.

Corollary 2. *The amplified intertemporal hedging effect is:*

$$\textit{observed Brownian hedging in } \mathcal{O} - \textit{Brownian hedging in } \mathcal{C}^{dyn}|_{\delta=\Delta\mathbb{E}_t(R),\chi=\frac{\lambda J^s_{Rt}}{\delta}} = \tag{43}$$

$$= \frac{\frac{q^{s\prime}(w^s_t)}{q^s(w^s_t)}\left[w^s_t y^s_t\right]_i - \frac{q^{h\prime}(w^n_t;\lambda,\delta,\chi)}{q^h(w^n_t;\lambda,\delta,\chi)}\left[w^n_t y^n_t\right]_i}{\left[\Delta\mathbb{E}_t\left(R\right)\right]_i} \tag{44}$$

for the Case 2 (CRRA investors) and 0 for Case 1 (log investors).

There are two sources of this difference. The first is that the shape of the value functions under the hypothetical and under the shocked state, that is, $q'(\cdot)/q(\cdot)$, are different. This effect tends to be small as in both cases the value functions respond to a similar change in expected returns. More importantly, unlike $\mathcal{C}^{\text{dyn}}$, the general equilibrium effect enters $\mathcal{O}$: in equilibrium, all investors react to the shift in residual demand, and therefore, aggregate positions change from $w^n_t y^n_t$ to $w^s_t y^s_t$.

Each investor understands that if others have a larger position of the asset that will influence the covariance of her wealth with future investment opportunities. In particular, when the portfolio of investors as a group performs well, the resulting increase in aggregate wealth implies larger risk-bearing capacity and lower risk premia. Investment opportunities are becoming worse. An investor with $\gamma > 1$ wants to hedge against that by increasing its exposure to these states, that is, a larger component of y implies a larger component of $\hat{y}$.

Linking the above observation back to our minimal example in Section 2, this is equivalent to having an endogenous intercept $a(y)$ that is increasing in investors' aggregate risk exposure y. Hence, $a(y^s) > a(y^n)$, so the investor's demand curve shifts upward.[18] This is the amplified intertemporal hedging effect and, for $\gamma > 1$, it makes the observed slope steeper than the dynamic slope, going against the endogenous risk effect.

Hedging against reversal effect. Finally, noticing that $\lambda J^s_{Rt} = \chi\Delta\mathbb{E}_t(R)$, the next corollary gives the last difference between $\mathcal{C}^{\text{dyn}}$ and $\mathcal{O}$ due to intertemporal hedging demand for the reversal. Again, the difference comes from that our observed slope is affected by the equilibrium change in all price paths due to the supply shock.

[18] Note that this general equilibrium effect is somewhat similar to the standard aggregate demand effect in macroeconomic models; for instance, in the Marginal Propensity to Consumption literature, one dollar of additional individual consumption will translate to some endogenous increase in individual wealth, which induces more consumption.

Corollary 3. *The hedging effect against the reversal effect for CRRA investors is:*

$$\begin{aligned}&\textit{hedging against reversal in } \mathcal{O} - \textit{hedging against reversal in } \mathcal{C}^{dyn}|_{\delta=\Delta\mathbb{E}_t(R),\chi=\frac{\lambda J^s_{Rt}}{\delta}} = \\ &= \frac{\frac{1}{\gamma}\left\{(\Omega^s_t)^{-1}\left[\frac{q^n(w^s_t(1+y^{s\top}_t J^s_{Rt}))}{q^s(w^s_t)(1+\hat{y}^{s\top}_t J^s_{Rt})^\gamma}-1\right]-(\Omega^n)^{-1}\left[\frac{q^n(w^n_t)}{q^h(w^n_t;\lambda,\delta,\chi)(1+\hat{y}^{h\top}_t\frac{\chi\delta}{\lambda})^\gamma}-1\right]\right\}[\chi\Delta\mathbb{E}_t(R)]_i}{[\Delta\mathbb{E}_t(R)]_i}. \quad (45)\end{aligned}$$

For the log investors $\gamma = 1$ and $q^h(\cdot) = q^n(\cdot) = q^s(\cdot) \equiv 1$.

Summary It is clear that shocks to residual supply exactly recover the slope of CRRA investors' demand only if the sum of the above three objects—endogenous risk effect, amplified intertemporal hedging effect, and hedging against reversal effect—is exactly zero. In the next subsection, through two applications, we demonstrate that there is no reason that any of these terms should be negligible under reasonable parameter values—to the opposite, they can be quantitatively substantial. Hence, generically, we should not expect that shifts in residual supply recover the slope of dynamic demand.

6 Applications and Discussions

In this section we perform a quantitative study on the difference between the dynamic and observed slope, with the help of two applications. We then discuss the connection between our model to the reduced-form factor-based asset demand system in HHHKL, and the way forward given our theoretical results regarding the wedge between these two objects.

6.1 Two Applications: Factor Elasticity and Micro Elasticity

We build up two applications to illustrate our results. These examples are intentionally placed into very different contexts to highlight the generality of our quantitative observations.

Factor elasticity Our first calibration focuses on the effect of a relative demand shock on a few large factors. To this end, we calibrate to the study of Ben-David et al. (2021) focusing on factor-level elasticities. Our baseline parameters are obtained by calibrating our model to asset return characteristics and fund flows patterns in Ben-David et al. (2021). They show that when investment styles are highly correlated with Morningstar ratings, changes in fund ratings result in style-level fund flows which lead to price effect. In our calibration, we set $N = 2$. We interpret the two assets as two portfolios of different investment styles. In particular, we consider the first asset as the small value portfolio and the second asset as the large growth portfolio. Our choice of parameters is summarized in Table 1 Panel A. For

(A) Factor elasticity			
Para.	Description	Value	Targeted Moments
$\bar{D}$	Expected cash flow of assets	[0.27, 0.26]	historical returns
σ	Cash flow volatility	diag(0.485, 0.485)	historical return volatilities
α	Hedgers' absolute risk aversion	1	normalization
γ	Investors' relative risk aversion	3	literature
r	Risk-free rate	0.03	historical average interest rate
ρ	Investors' time discount rate	0.08	investor holdings
s	Asset supply	[1,1]	normalization
u^s	Hedgers' endowment shock	[0.15,0]	fund flows Ben-David et al. (2021)
λ	Intensity of state transitions	0.5	persistence Ben-David et al. (2021)
(B) Micro elasticity			
Para.	Description	Value	Targeted Moments
$\bar{D}_i$	Expected cash flow of asset i	0.0369	historical returns
σ_i	Cash flow volatility of asset i	0.0692	historical return volatilities
κ	Cash flow correlation of i and j	0.3	idiosyncratic volatility
α	Hedgers' absolute risk aversion	1	normalization
γ	Investors' relative risk aversion	3	literature
r	Risk-free rate	0.03	historical average interest rate
ρ	Investors' time discount rate	0.08	investor holdings
S_i	Asset supply of asset i	0.0446	normalization ($1/\sqrt{I}$)
υ_1	Endowment to index assets	0	normalization
υ_2	Endowment to non-index assets	0.0107	fund flows GS (2024)
λ	Intensity of state transitions	0.2	persistence GS (2024)

Table 1: Parameters and Calibration Details.

instance, the persistence is 2 in our factor elasticity calibration, matching the empirical half life of fund flows as reported by Ben-David et al. (2021).

Micro elasticity In the second application, we are interested in an economy where the residual supply shock affects only very few assets out of many. Hence, we consider the case of index inclusion calibrating our model to the recent study of Greenwood and Sammon (2024), or GS (2024) for short. Similar to our multi-asset setting in Section 4.3, we picture a market with a large number of fundamentally identical assets with the only difference in hedgers' endowment. In particular, suppose that in the normal state, hedgers' endowment for the first I_1 asset differs from the rest:

$$[u^n]_{i=1}^{I_1} = \upsilon_1, \ [u^n]_{i=I_1+1}^{I} = \upsilon_2.$$

where $\upsilon_1 \neq \upsilon_2$ are scalars. Our interpretation is that the first I_1 assets are in an index. We are interested in a shock for which $[u^n]_{i=2}^{I_1} = [u^s]_{i=2}^{I_1}$ while $[u^s]_1 = \upsilon_2$, $[u^s]_I = \upsilon_1$. That is, the

(A) Factor elasticity					
	Persistence $(1/\lambda)$				
	0.20	0.50	1.00	**2.00**	5.00
Myopic Slope	3.19	3.19	3.20	**3.21**	3.23
Dynamic Slope	3.18	3.17	3.17	**3.17**	3.17
Observed Slope	2.14	1.95	1.67	**1.24**	0.46
Observed Slope/Dynamic Slope	67%	61%	53%	**39%**	14%
Dynamic minus Observed (DmO)	1.04	1.23	1.50	**1.93**	2.71
Shock to risk/DmO	134%	127%	120%	**112%**	104%
Diff Hedging Brownian/DmO	-34%	-27%	-20%	**-13%**	-6%
Diff Hedging Reversal/DmO	0%	0%	1%	**1%**	2%
(B) Micro elasticity					
	Persistence $(1/\lambda)$				
	0.20	0.50	1.00	2.00	**5.00**
Myopic Slope	8.87	8.58	8.32	7.96	**7.81**
Dynamic Slope	8.87	8.58	8.32	7.96	**7.81**
Observed Slope	6.72	6.14	5.49	4.60	**3.42**
Observed Slope/Dynamic Slope	76%	72%	66%	58%	**44%**
Dynamic minus Observed (DmO)	2.15	2.44	2.83	3.36	**4.39**
Shock to risk/DmO	166%	153%	141%	129%	**116%**
Diff Hedging Brownian/DmO	-66%	-53%	-41%	-29%	**-16%**
Diff Hedging Reversal/DmO	0%	0%	0%	0%	**0%**

Table 2: Comparative statics with respect to persistence $1/\lambda$

first asset is excluded from and the last asset is included in the index. Just as in the example of Section 4, we intentionally design this example, so that the shock itself has a minimal effect on the structure of the economy—effectively, we only swap the labels of the first and the last assets. However, the main difference here is that instead of focusing on diminishingly small, but infinitely persistent shock, we calibrate its size and persistence (and each of the other parameters) to a real-world example. Our choice of parameters is summarized in Table 1 Panel B; for instance, we set the persistence to be 5 ($\lambda = 0.2$), which targets the conditional probability of index inclusion or deletion. For detailed explanations of our calibration, see Appendix E.[19]

[19]In Appendix E, we rescale observed slope, $\mathcal{O}$, to make it comparable to standard estimates of price elasticity as in Gabaix and Koijen (2022). Under our baseline parameters, our factor elasticity calibration implies a price elasticity of 1.4. Ben-David et al. (2021) finds 0.2. The index-inclusion calibration yields an implied price elasticity of 399, which is substantially lower than the commonly cited value of around 6250 from Petajisto (2009). For comparison, the empirical micro elasticity estimated by Greenwood and Sammon (2024) is approximately 2.7. These results are in line with a broader observation that we do not dispute in this paper: that frictionless asset pricing models, in their simplest form, face challenges in rationalizing the empirically small price elasticity.

What Do We Learn from Two Calibrated Applications? Tables 2 reports our main calibration results. Panel A is for factor elasticity calibration while Panel B is for micro elasticity. In each panel the colored (with bold font) column contains our benchmark parametrization, while the other columns show how our effects vary with the persistence of the demand shift.[20] For each set of parameters, we calculate the myopic slope, the dynamic slope and the observed slope, and decompose the difference to our three effects.

There are a number of common takeaways across the two calibrations. First, while the dynamic slope is always smaller than myopic slope consistently with the illustration in Panel A and B in Figure 2, we find that the difference is quantitatively small. That is, whether one considers the myopic or the dynamic slope as the more intuitive measure for investor demand elasticity seems to have little practical significance.

Second, in all our examples, a shift in residual supply identifies a significantly flatter demand curve than investors' dynamic slope. In our baseline parameterizations, observed slope is approximately 40% of the dynamic (and myopic) slope. This is because, just as in our minimal demand system, a shock that shifts residual supply also changes the shape of investors' demand function as general equilibrium effects alter the covariance of returns and induce changes in hedging demand. The fraction decreases with the persistence of the identifying shock (shrinking to 14% in our first example), meaning that the empirical underestimation of our conceptual demand slope is more pronounced when the identifying shock is more persistent. This is intuitive as a more persistent shock should have larger effect on the dynamic equilibrium.

Third, the decomposition of the difference between demand slope and observed slope shows that the endogenous risk effect dominates the others. When hedgers demand less of the asset, investors have to hold a larger share. Therefore, shock to the aggregate wealth of this sector will have a larger effect on its return. From the point of view of individual investors, the assets are getting riskier and covary more with other stocks which decreases the slope of their demand. In fact, because the amplified intertemporal hedging component is going against the endogenous risk effect whenever $\gamma > 1$, in all the highlighted cases the endogenous risk effect is more than the 100% of the total under-measurement.[21] This share is shrinking with the persistence of the shock ranging from 104% to 166% across our examples.

Fourth, as intertemporal hedging is amplified in general equilibrium, this fact implies that this term plays a significant role in explaining the difference between dynamic slope and measured observed slope. We observe that while the direction of the effect is in line with the direction of Merton's intertemporal hedging, it is several magnitudes higher once the general

[20] For each column we recalibrate the parameters u^{ξ} to ensure that the implied expected return shocks remain constant at $\Delta\mathbb{E}_t(R) = 0.025$.

[21] In Appendix E, for both of our applications, we provide calibrations with $\gamma < 1$. In those cases, the endogenous risk effect is responsible for less than 100% of the total as our two main effects resulting in mismeasurements of the same sign.

equilibrium effect is taken into account. Recall from expression (44) that the size of this effect is mainly driven by the ratio of the shift in the equilibrium aggregate position to the shift in expected returns: $\frac{w^s y^s - w^n y^n}{[\Delta \mathbb{E}_t(R)]_i}$. In absolute terms, our results show that this effect can be responsible for up to 34% and 66% of the total mismeasurement in our two applications, respectively. Its absolute magnitude is especially large for less persistent shocks.

Finally, the component with respect to the intertemporal hedging for the reversal of the supply shock tends to be relatively small.[22]

6.2 Demand in Different Units and Connection to HHHKL

We follow the tradition of Merton (1971) and adopt CRRA preferences to link investors' portfolio share demand to asset returns, as in equation (33). Since the key identification challenge arises from the endogenous adjustment of investors' demand curve in response to supply shocks within a dynamic general equilibrium setting, it is evident that the issue highlighted by our analysis does not depend on the particular 'units' used when deriving individual investors' asset demand.

As an illustration, we switch here to the CARA-normal tradition to express asset demand in terms of units of assets as a function of their price. Let $\hat{Y}_t^\xi$ denote the individual investor's demand in units of assets—as opposed to $\hat{y}_t^\xi$ which is individual investor's portfolio share—when the economy is in state ξ, and let Y_t^ξ be the aggregate counterpart. Then, we can rewrite the optimal asset demand in (33) as

$$\begin{aligned}\underbrace{\hat{Y}_t^\xi}_{D_t} &= \underbrace{\frac{-r}{\gamma^A(w_t^\xi)}(\Pi_t^\xi)^{-1}}_{\mathcal{E}_t}\underbrace{P_t^\xi}_{P_t} \\ &+ \underbrace{\frac{1}{\gamma^A(w_t^\xi)}\frac{q^{\xi\prime}(w_t^\xi)}{q^\xi(w_t^\xi)}Y_t^\xi + \frac{1}{\gamma^A(w_t^\xi)}(\Pi_t^\xi)^{-1}\left[\bar{D} + \mu_{Pt}^\xi + \frac{\lambda^\xi q^{-\xi}(w_t^\xi + Y_t^{\xi\top}J_t^\xi)}{q^\xi(w_t^\xi)\left(1+\frac{1}{\hat{w}_t^\xi}\hat{Y}_t^{\xi\top}J_t^\xi\right)^\gamma}J_t^\xi\right]}_{e_t}.\end{aligned} \tag{46}$$

Here, we define $\gamma^A(w_t^\xi) \equiv \gamma/w_t^\xi$ as the investor's absolute risk aversion, $\Pi_t^n \equiv (\sigma_{Pt}^n + \sigma)^\top(\sigma_{Pt}^n + \sigma)$ as the dollar return variance-covariance matrix, and $J_t^\xi \equiv P_t^{-\xi} - P_t^\xi$ as the jump in price as the economy transitions from state ξ to state $-\xi$.

The labeling of each term in (46) shows how the dynamic demand curve would map to the standard formulation of demand elasticity estimation (with the notation of HHHKL):

$$D_t = \mathcal{E}_t P_t + e_t$$

[22] In Appendix E, we provide a sensitivity analysis with respect to the size of the shock to demonstrate that this effect is highly non-linear. In fact, it dominates the amplified hedging effect when the shock is large.

where D_t is the vector of investor's demand, P_t is the price vector, $\mathcal{E}_t$ is the matrix of own and cross elasticities while e_t is a, not necessarily zero-mean, vector of demand shifters.

This mapping illustrates the complementary insights in HHHKL and our work. HHHKL shows that under the assumption of *homogeneous substitution conditional on oberservables* there is an intuitive estimation procedure using both cross-sectional and time-series information to estimate the elasticity matrix $\mathcal{E}$, if the econometrician can find the instrumental variables satisfying the standard conditions for causual inference.

Namely, those instruments should shift the price vector without affecting the elasticity matrix $\mathcal{E}_t$ which the econometrician aims to estimate, and the demand shifter e_t that the econometrician cannot observe. The main insight of our paper is that standard instruments based on the shift in residual supply will not meet these conditions once we consider a dynamic environment. This is apparent in (46) as many elements of e_t and $\mathcal{E}_t$ are indexed by the state of residual supply ξ.[23]

To close the argument, in the rest of this section, we show that demand curve (46) does satisfy the assumption of homogeneous substitution conditional on observables. That is, the HHHKL estimation procedure, if one finds instruments satisfying the standard conditions, still remains valid.

For simplicity, we consider i.i.d. underlying cash flow shocks, so that the exogenous exogenous diffusion matrix in (3) for I assets is $\sigma = \bar{\sigma}\mathbf{I}$, with $\bar{\sigma}$ being a scalar and $\mathbf{I}$ an identity matrix $I \times I$. (We consider a factor structure for σ in Appendix B). Following the same logic as in (23) in Section 4.2, we can derive the endogenous dollar return diffusion matrix of the assets to be

$$\underbrace{\sigma^{\xi}_{Pt}}_{I\times I} = \frac{\bar{\sigma}}{1 - \underbrace{P^{\xi\prime}(w^{\xi}_t)^{\top} Y^{\xi}_t}_{\text{scalar}}} \cdot \underbrace{Y^{\xi}_t}_{I\times 1} \underbrace{P^{\xi\prime}(w^{\xi}_t)^{\top}}_{1\times I}. \tag{47}$$

As a result, the endogenous dollar variance-covariance matrix equals

$$\Pi^{\xi}_t \equiv (\sigma^{\xi}_{Pt} + \sigma)^{\top}(\sigma^{\xi}_{Pt} + \sigma) = \bar{\sigma}^2 \left[\mathbf{I} + \frac{P^{\xi\prime}(w^{\xi}_t) Y^{\xi\top}_t}{1 - P^{\xi\prime}(w^{\xi}_t)^{\top} Y^{\xi}_t}\right] \left[\mathbf{I} + \frac{Y^{\xi}_t P^{\xi\prime}(w^{\xi}_t)^{\top}}{1 - P^{\xi\prime}(w^{\xi}_t)^{\top} Y^{\xi}_t}\right]. \tag{48}$$

In Appendix B, we show that once we define the asset characteristic matrix to be

$$\boldsymbol{\beta} \equiv \left[Y^{\xi}_t \quad P^{\xi\prime}(w^{\xi}_t)\right]_{I\times 2} \tag{49}$$

[23]Note that this rewrite also illustrates that our insights apply to the estimation of demand elasticities with respect to prices just the same as to the estimation of elasticities with respect to expected return.

and the substitution matrix to be

$$\mathcal{E}_{sub} \equiv \begin{bmatrix} P^{\xi\prime}(w_t^\xi)^\top P^{\xi\prime}(w_t^\xi) & -1 \\ -1 & 0 \end{bmatrix}_{2\times 2}, \tag{50}$$

then we have the exact form of demand system studied in HHHKL:

$$\mathcal{E}_t \equiv \left[\frac{\partial \hat{Y}^\xi}{\partial P^\xi}\right]_{I\times I} = -\frac{r}{\gamma^A(w_t^\xi)\bar{\sigma}^2}\left(\mathbf{I} + \boldsymbol{\beta}\mathcal{E}_{sub}\boldsymbol{\beta}^\top\right). \tag{51}$$

In the language of HHHKL, in our model each asset has two (endogenous) characteristics, as suggested by (49): its aggregate holdings by investors Y_t^ξ, and its price loading on the aggregate investor wealth $P^{\xi\prime}(w_t^\xi)$. And, we have assumed away any exogenous factor structure for the dividend process; if we were to assume a K-factor structure for the dividend diffusion σ, then the total dimension of asset characteristics will be $K+2$, as shown in Appendix B.

6.3 The Way Forward

Our paper highlights a conceptual challenge for estimating simple asset demand systems in light of standard forces in dynamic asset pricing. A persistent shift in residual supply changes the equilibrium allocation of risk which affects the equilibrium return and covariance properties of assets. As investors' demand curves for the asset with the altered return and covariance properties are different, it implies that the identifying shock changes the estimand: a violation of the exclusion restriction.

So how do researchers move forward in light of these observations? One possibility is to focus on economies where the mismeasurement is smaller. We show that if the shift in residual supply is small and transitory, the difference between observed slope and dynamic (or myopic) slope is significantly muted. However, as shown in Section 4, this approach cannot eliminate the problem as we have also demonstrated in Proposition 2 that the bias remains even in the limit.

One question is whether carefully chosen controls can alleviate this problem. To investigate, we follow the insight of Merton (1971) and write the demand curve in a more general way. As we explained, the Merton approach represents the demand curve as a function of the instantaneous expected return, the instantaneous return covariance matrix, and the intertemporal hedging demand stemming from the covariance of marginal utility and future investment opportunities. Considering the case of a single risky asset, and with a slightly loose notation we write the demand curve for the representative investor as

$$\hat{y}_t^\xi = D(\mathbb{E}_t(R^\xi), (\sigma_{Rt}^\xi)^2, IHD_t^\xi).$$

Here, $D(\cdot,\cdot,\cdot)$ is a general function and the last argument, IHD_t^ξ, denotes the intertemporal hedging component. Then, in differences and after linearization, we can write

$$\Delta\hat{y}_t^\xi = D_1'\Delta\mathbb{E}_t(R^\xi) + D_2'\Delta(\sigma_{Rt}^\xi)^2 + D_3'\Delta IHD_t^\xi,$$

where D_1', D_2', D_3' are the three partial derivatives of D. We ask: can the standard approach of instrumenting $\Delta\mathbb{E}_t(R^\xi)$, augmented by the inclusion of controls for the volatility of the asset σ_{Rt}^ξ along with a proxy for variation in intertemporal hedging, yield a consistent approximation of the demand curve D, or, at least the slope with respect to $\mathbb{E}_t(R^\xi)$, D_1'?

Our observations in this paper imply that as long as the instrument is a shift in residual supply, for instance, $\xi = n$ changing to $\xi = s$ as in our model, then this approach leads to a bad control problem (e.g. Angrist and Pischke, 2009). Because of the tilt and shift, the potential control variables for $(\sigma_{Rt}^\xi)^2$ and IHD_t^ξ are likely to be affected by the instrument leading to an endogeneity bias. This is expected to distort the estimate of the slope coefficient corresponding to D_1'.

Therefore, a more promising avenue might be to impose more structure and separately model and instrument each argument of the demand function affected by the demand shift. Our fully solved model offers some guidance along this direction.

Another more reduced-form route is to step back from the view that asset markets can be seen as a demand system with well-defined supply and demand shocks. For instance, if the Fed wants to know the effect on yields of buying 1M worth of treasuries, they might not care whether the estimate corresponds to any investor's slope of demand. Perhaps the Fed only wants to know the resulting aggregate multiplier, taking into account of all complex equilibrium mechanisms triggering arbitrary tilts and shifts of some investors' demand curves. If this is the case, we can take the observed relationship between prices and quantities as the main object of interest. While, the estimated relationship would not correspond to the quantity response of a given investor group to a change in prices, it might provide sophisticated benchmarks for the literature on what "canonical" models imply on this aggregate multiplier and how sensitive it is to changes of parameters in the economy. We explore this route in the companion paper He et al. (2025).

7 Concluding Remarks

Measuring how investors were to react to exogenous shocks to returns, or its mirror image, the reaction of returns to changes in aggregate quantities is crucial for a wide range of policy relevant questions. However, according to standard finance theory the slope of the demand curve is determined in equilibrium. In particular, in standard models it depends on the dynamics of future asset prices in general, and in the covariance of asset returns in

particular. In this paper, we emphasize that exogenous shocks to the residual *supply* curve, a commonly used instrument in the empirical literature to identify the slope of investors' demand curve, tends to change the future dynamics of asset prices and therefore necessarily changes the investors' demand curve, violating the exclusion restriction. In this paper, we calibrate this effect and find that the dynamic slope tends to be two to seven times larger than its measured counterpart. Importantly, this general equilibrium effect does not vanish when the shock is small.

References

An, Yu and Zeyu Zheng, "Information from Uninformed Demand: How Predictable Demand Propagates Across Assets," Working Paper 2025.

Angrist, Joshua D and Jörn-Steffen Pischke, *Mostly harmless econometrics: An empiricist's companion*, Princeton University Press, 2009.

Ben-David, Itzhak, Jiacui Li, Andrea Rossi, and Yang Song, "Ratings-Driven Demand and Systematic Price Fluctuations," *The Review of Financial Studies*, 09 2021, *35* (6), 2790–2838.

Brunnermeier, Markus K and Yuliy Sannikov, "A macroeconomic model with a financial sector," *American Economic Review*, 2014, *104* (2), 379–421.

Campbell, John Y. and Luis Viceira, *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*, first ed., Oxford, UK: Oxford University Press, 2002.

_ **, Martin Lettau, Burton G. Malkiel, and Yexiao Xi**, "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk," *Journal of Finance*, 2001, *56* (1), 1–43.

Davis, Carter, Mahyar Kargar, and Jiacui Li, "Why is asset demand inelastic?," Working Paper 2023.

Fuchs, William, Satoshi Fukuda, and Daniel Neuhann, "Demand-System Asset Pricing: Theoretical Foundations," 2025. University of Austin, Texas.

Gabaix, Xavier and Ralph SJ Koijen, "In search of the origins of financial fluctuations: The inelastic markets hypothesis," Working Paper 2022.

_ **,** _ **, Federico Mainardi, Sangmin Simon Oh, and Motohiro Yogo**, "Limited Risk Transfer Between Investors: A New Benchmark for Macro-Finance Models," Technical Report, National Bureau of Economic Research 2025.

Greenwood, Robin and Marco Sammon, "The disappearing index effect," *The Journal of Finance*, 2024.

Haddad, Valentin, Alan Moreira, and Tyler Muir, "Asset purchase rules: How QE transformed the bond market," Technical Report, Working paper, UCLA, Rochester, and USC 2024.

_ **, Zhiguo He, Paul Huebner, Peter Kondor, and Erik Loualiche**, "Causal Inference for Asset Pricing," 2025. University of California, Los Angeles.

He, Zhiguo and Arvind Krishnamurthy, "Intermediary asset pricing," *American Economic Review*, 2013, *103* (2), 732–770.

_ , Péter Kondor, and Jessica Shi Li, "Price Multiplier in Dynamic Asset Pricing," 2025. Stanford University.

Koijen, Ralph S. J. and Motohiro Yogo, "A Demand System Approach to Asset Pricing," *Journal of Political Economy*, 2019, *127* (4), 1475–1515.

Koijen, Ralph S J, Robert J Richmond, and Motohiro Yogo, "Which Investors Matter for Equity Valuations and Expected Returns?," *Review of Economic Studies*, 2023, p. rdad083.

Kondor, Péter and Dimitri Vayanos, "Liquidity risk and the dynamics of arbitrage capital," *The Journal of Finance*, 2019, *74* (3), 1139–1173.

Longstaff, Francis A and Jiang Wang, "Asset pricing and the credit market," *The Review of Financial Studies*, 2012, *25* (11), 3169–3215.

Merton, Robert C, "Optimum consumption and portfolio rules in a continuous-time model," *Journal of Economic Theory*, 1971, *3* (4), 373–413.

Petajisto, Antti, "Why do demand curves for stocks slope down?," *Journal of Financial and Quantitative Analysis*, 2009, *44* (5), 1013–1044.

Shleifer, Andrei, "Do demand curves for stocks slope down?," *The Journal of Finance*, 1986, *41* (3), 579–590.

van Binsbergen, Jules H, Benjamin David, and Christian C Opp, "How (Not) to Identify Demand Elasticities in Dynamic Asset Markets," 2025. University of Pennsylvania.

Appendix

A Proofs

A.1 Proof of Proposition 1 and Proposition 4

A. Hedgers' Problem

We first consider a hedger's problem. Let $Q_t^\xi \equiv \text{diag}(P_t^\xi)^{-1} X_t^\xi$ denote the hedger's asset holdings in units when the economy is in state $\xi \in \{n, s\}$. From time t to $t+dt$, the hedger's change in wealth is given by

$$dv_t^\xi = rv_t^\xi dt + Q_t^{\xi\top}(dD_t + dP_t^\xi - rP_t^\xi dt) + u^{\xi\top} dD_t. \tag{A.1}$$

Using (3) and (6), we have

$$\begin{aligned} dv_t^\xi = &\Big[rv_t^\xi + Q_t^{\xi\top}(\bar{D} + \mu_{Pt}^\xi - rP_t^\xi) + u^{\xi\top}\bar{D}\Big]dt \\ &+ \Big[(\sigma_{Pt}^\xi + \sigma)Q_t^\xi + \sigma u_t^\xi\Big]^\top dB_t + Q_t^{\xi\top} J_t^\xi dN_t^\xi. \end{aligned} \tag{A.2}$$

Thus,

$$\frac{\mathbb{E}_t[dv_t^\xi]}{dt} = rv_t^\xi + Q_t^{\xi\top}(\bar{D} + \mu_{Pt}^\xi - rP_t^\xi + \lambda^\xi J_t^\xi) + u^{\xi\top}\bar{D}, \tag{A.3}$$

$$\frac{\text{Var}_t(dv_t^\xi)}{dt} = \big[Q_t^{\xi\top}(\sigma_{Pt}^\xi + \sigma)^\top + u_t^{\xi\top}\sigma^\top\big]\big[(\sigma_{Pt}^\xi + \sigma)Q_t^\xi + \sigma u_t^\xi\big] + \lambda^\xi (Q_t^{\xi\top} J_t^\xi)^2. \tag{A.4}$$

Substituting (A.3) and (A.4) into the hedger's mean-variance objective and taking first-order condition yields

$$\bar{D} + \mu_{Pt}^\xi - rP_t^\xi + \lambda^\xi J_t^\xi - \alpha\Big\{(\sigma_{Pt}^\xi + \sigma)^\top\big[(\sigma_{Pt}^\xi + \sigma)Q_t^\xi + \sigma u^\xi\big] + \lambda^\xi J_t^\xi J_t^{\xi\top} Q_t^\xi\Big\} = 0. \tag{A.5}$$

Rearranging yields

$$Q_t^\xi = \Big[(\sigma_{Pt}^\xi + \sigma)^\top(\sigma_{Pt}^\xi + \sigma) + \lambda^\xi J_t^\xi J_t^{\xi\top}\Big]^{-1}\Big[\frac{\bar{D} + \mu_{Pt}^\xi - rP_t^\xi + \lambda^\xi J_t^\xi}{\alpha} - (\sigma_{Pt}^\xi + \sigma)^\top \sigma u^\xi\Big]. \tag{A.6}$$

Note that $\mu_{Rt}^\xi \equiv \text{diag}(P_t^\xi)^{-1}(\bar{D} + \mu_{Pt}^\xi)$, $\sigma_{Rt}^\xi \equiv (\sigma_{Pt}^\xi + \sigma)\,\text{diag}(P_t^\xi)^{-1}$, and $J_{Rt}^\xi \equiv \text{diag}(P_t^\xi)^{-1} J_t^\xi$. Then from (A.6), $X_t^\xi = \text{diag}(P_t^\xi) Q_t^\xi$ can be written as

$$X_t^\xi = \Big[\Omega_t^\xi + \lambda^\xi J_{Rt}^\xi J_{Rt}^{\xi\top}\Big]^{-1}\Big[\frac{\mu_{Rt}^\xi - r\mathbf{1} + \lambda^\xi J_{Rt}^\xi}{\alpha} - (\sigma_{Rt}^\xi)^\top \sigma u^\xi\Big], \tag{A.7}$$

where $\Omega_{Rt}^\xi \equiv (\sigma_{Rt}^\xi)^\top \sigma_{Rt}^\xi$.

B. Investors' Problem

Let $\hat{w}_t^\xi$ denote an individual investor's time t wealth when the economy is in state ξ. Let $\hat{Y}_t^\xi \equiv \text{diag}(P_t^\xi)^{-1}\hat{w}_t^\xi \hat{y}_t^\xi$ denote the individual investor's corresponding asset holdings (in units). The investor's wealth $\hat{w}_t^\xi$ evolves according to

$$\begin{aligned} d\hat{w}_t^\xi &= r\hat{w}_t^\xi dt + \hat{Y}_t^{\xi\top}(dD_t + dP_t^\xi - rP_t^\xi dt) - \hat{c}_t^\xi dt \\ &= \left[r\hat{w}_t^\xi - \hat{c}_t^\xi + \hat{Y}_t^{\xi\top}(\bar{D} + \mu_{Pt}^\xi - rP_t^\xi)\right]dt + \hat{Y}_t^{\xi\top}(\sigma_{Pt}^\xi + \sigma)^\top dB_t + \hat{Y}_t^{\xi\top} J_t^\xi dN_t^\xi. \end{aligned} \tag{A.8}$$

Let w_t^ξ denote the aggregate wealth of investors at time t, and let Y_t^ξ denote the aggregate counterpart of $\hat{Y}_t^\xi$. Then w_t^ξ follows

$$dw_t^\xi = \left[rw_t^\xi - c_t^\xi + Y_t^{\xi\top}(\bar{D} + \mu_{Pt}^\xi - rP_t^\xi)\right]dt + Y_t^{\xi\top}(\sigma_{Pt}^\xi + \sigma)^\top dB_t + Y_t^{\xi\top} J_t^\xi dN_t^\xi. \tag{A.9}$$

The individual investor's problem is to choose the optimal consumption and asset holdings $\{\hat{c}_t^\xi, \hat{Y}_t^\xi\}$ to maximize their discounted utility over intertemporal consumption, subject to the wealth dynamic (A.8) while taking prices P_t^ξ as given. Let $V^\xi(\hat{w}_t^\xi, w_t^\xi)$ denote the investor's value function at time t in state ξ. The investor's HJB equation is

$$\begin{aligned} \rho V^\xi = \max_{\hat{c}_t^\xi, \hat{Y}_t^\xi} \Big\{ & u(\hat{c}_t^\xi) + V_{\hat{w}}^\xi \mu_t^{\hat{w},\xi} + \frac{1}{2} V_{\hat{w}\hat{w}}^\xi (\sigma_t^{\hat{w},\xi})^\top \sigma_t^{\hat{w},\xi} + V_w^\xi \mu_t^{w,\xi} + \frac{1}{2} V_{ww}^\xi (\sigma_t^{w,\xi})^\top \sigma_t^{w,\xi} \\ & + V_{\hat{w}w}^\xi (\sigma_t^{\hat{w},\xi})^\top \sigma_t^{w,\xi} + \lambda^\xi (V^{-\xi,+} - V^\xi) \Big\}, \end{aligned} \tag{A.10}$$

where V^ξ stands for the value function $V^\xi(\hat{w}_t^\xi, w_t^\xi)$, while the superscript $+$ on $V^{-\xi,+}$ denotes that the value function is evaluated after the jump, that is

$$V^{-\xi,+} \equiv V^{-\xi}\left(\hat{w}_t^\xi + \hat{Y}_t^{\xi\top} J_t^\xi, w_t^\xi + Y_t^{\xi\top} J_t^\xi\right). \tag{A.11}$$

Moreover, $\mu_t^{\hat{w},\xi}$ and $\sigma_t^{\hat{w},\xi}$ are the drift and the diffusion of the individual investor's wealth $\hat{w}_t^\xi$, while $\mu_t^{w,\xi}$ and $\sigma_t^{w,\xi}$ are the drift and the diffusion of investors' aggregate wealth w_t^ξ. From (A.8) and (A.9),

$$\mu_t^{\hat{w},\xi} \equiv r\hat{w}_t^\xi - \hat{c}_t^\xi + \hat{Y}_t^{\xi\top}(\bar{D} + \mu_{Pt}^\xi - rP_t^\xi), \tag{A.12}$$

$$\sigma_t^{\hat{w},\xi} \equiv (\sigma_{Pt}^\xi + \sigma)\hat{Y}_t^\xi, \tag{A.13}$$

$$\mu_t^{w,\xi} \equiv rw_t^\xi - c_t^\xi + Y_t^{\xi\top}(\bar{D} + \mu_{Pt}^\xi - rP_t^\xi), \tag{A.14}$$

$$\sigma_t^{w,\xi} \equiv (\sigma_{Pt}^\xi + \sigma)Y_t^\xi. \tag{A.15}$$

The first-order condition with respect to consumption implies that

$$u'(\hat{c}_t^\xi) = V_{\hat{w}}^\xi. \tag{A.16}$$

Similarly, the first-order condition with respect to asset holdings implies that $\hat{Y}_t^\xi$ satisfies

$$\begin{aligned} 0 =& V_{\hat{w}}^\xi(\bar{D} + \mu_{Pt}^\xi - rP_t^\xi) + V_{\hat{w}\hat{w}}^\xi(\sigma_{Pt}^\xi + \sigma)^\top(\sigma_{Pt}^\xi + \sigma)\hat{Y}_t^\xi \\ &+ V_{\hat{w}w}^\xi(\sigma_{Pt}^\xi + \sigma)^\top(\sigma_{Pt}^\xi + \sigma)Y_t^\xi + \lambda^\xi V_{\hat{w}}^{-\xi,+} J_t^\xi. \end{aligned} \tag{A.17}$$

Rearranging the first-order condition (A.17), we can write the individual investor's asset demand as

$$\hat{Y}_t^\xi = \frac{V_{\hat{w}}^\xi}{(-V_{\hat{w}\hat{w}}^\xi)}(\Pi_t^\xi)^{-1}(\bar{D} + \mu_{Pt}^\xi - rP_t^\xi) + \frac{V_{\hat{w}w}^\xi}{(-V_{\hat{w}\hat{w}}^\xi)}Y_t^\xi + \lambda^\xi \frac{V_{\hat{w}}^{-\xi,+}}{(-V_{\hat{w}\hat{w}}^\xi)}(\Pi_t^\xi)^{-1}J_t^\xi, \tag{A.18}$$

where we define $\Pi_t^\xi \equiv (\sigma_{Pt}^\xi + \sigma)^\top(\sigma_{Pt}^\xi + \sigma)$. By definition, $\hat{y}_t^\xi = \frac{1}{\hat{w}_t^\xi}\operatorname{diag}(P_t^\xi)\hat{Y}_t^\xi$ and $y_t^\xi = \frac{1}{w_t^\xi}\operatorname{diag}(P_t^\xi)Y_t^\xi$. Then (A.8) can be rewritten as

$$\begin{aligned} d\hat{w}_t^\xi &= r\hat{w}_t^\xi dt + \hat{w}_t^\xi \hat{y}_t^{\xi\top}\Big(dR_t^\xi - r\mathbf{1}dt\Big) - \hat{c}_t^\xi dt \\ &= \Big[(1 - \hat{y}_t^{\xi\top}\mathbf{1})r\hat{w}_t^\xi - \hat{c}_t^\xi + \hat{w}_t^\xi \hat{y}_t^{\xi\top}\mu_{Rt}^\xi\Big] + \hat{w}_t^\xi \hat{y}_t^{\xi\top}(\sigma_{Rt}^\xi)^\top dB_t + \hat{w}_t^\xi \hat{y}_t^{\xi\top}J_{Rt}^\xi dN_t^\xi. \end{aligned} \tag{A.19}$$

Similarly, (A.9) can be rewritten as

$$dw_t^\xi = \Big[(1 - y_t^{\xi\top}\mathbf{1})rw_t^\xi - c_t^\xi + w_t^\xi y_t^{\xi\top}\mu_{Rt}^\xi\Big] + w_t^\xi y_t^{\xi\top}(\sigma_{Rt}^\xi)^\top dB_t + w_t^\xi y_t^{\xi\top}J_{Rt}^\xi dN_t^\xi. \tag{A.20}$$

And from (A.17), we have

$$\begin{aligned} \hat{y}_t^\xi =& \frac{V_{\hat{w}}^\xi}{\hat{w}_t^\xi(-V_{\hat{w}\hat{w}}^\xi)}(\Omega_t^\xi)^{-1}(\mu_{Rt}^\xi - r\mathbf{1}) + \frac{V_{\hat{w}w}^\xi}{\hat{w}_t^\xi(-V_{\hat{w}\hat{w}}^\xi)}w_t^\xi y_t^\xi + \lambda^\xi \frac{V_{\hat{w}}^{-\xi,+}}{\hat{w}_t^\xi(-V_{\hat{w}\hat{w}}^\xi)}(\Omega_t^\xi)^{-1}J_{Rt}^\xi \\ =& \frac{V_{\hat{w}}^\xi}{\hat{w}_t^\xi(-V_{\hat{w}\hat{w}}^\xi)}(\Omega_t^\xi)^{-1}\Big(\mu_{Rt}^\xi - r\mathbf{1} + \lambda^\xi J_{Rt}^\xi\Big) + \frac{V_{\hat{w}w}^\xi}{\hat{w}_t^\xi(-V_{\hat{w}\hat{w}}^\xi)}w_t^\xi y_t^\xi \\ &+ \frac{1}{\hat{w}_t^\xi}\frac{V_{\hat{w}}^{-\xi,+} - V_{\hat{w}}^\xi}{(-V_{\hat{w}\hat{w}}^\xi)}(\Omega_t^\xi)^{-1}\lambda^\xi J_{Rt}^\xi. \end{aligned} \tag{A.21}$$

A.1.1 Log Investors

A log investor maximizes

$$\mathbb{E}_t\Big[\int_t^\infty e^{-\rho(\tau-t)}\log\hat{c}_\tau d\tau\Big], \tag{A.22}$$

subject to the wealth dynamic (A.8). To facilitate solution, we let $\hat{z}_t^\xi \equiv \hat{Y}_t^\xi/\hat{w}_t^\xi$ and $z_t^\xi \equiv Y_t^\xi/w_t^\xi$. Then the investor's wealth dynamic (A.8) becomes

$$d\hat{w}_t^\xi = r\hat{w}_t^\xi dt + \hat{w}_t^\xi \hat{z}_t^{\xi\top}(dD_t + dP_t^\xi - rP_t^\xi dt) - \hat{c}_t^\xi dt. \tag{A.23}$$

Substituting (3) and (6) into the individual investor's wealth dynamic (A.23) yields

$$\begin{aligned} d\hat{w}_t^\xi =& \Big[r\hat{w}_t^\xi - \hat{c}_t^\xi + \hat{w}_t^\xi \hat{z}_t^{\xi\top}(\bar{D} + \mu_{Pt}^\xi - rP_t^\xi) \Big] dt + \hat{w}_t^\xi \hat{z}_t^{\xi\top}(\sigma_{Pt}^\xi + \sigma)^\top dB_t \\ &+ \hat{w}_t^\xi \hat{z}_t^{\xi\top} J_t^\xi dN_t^\xi. \end{aligned} \tag{A.24}$$

Similarly, the investors' total wealth dynamic aggregates to

$$\begin{aligned} dw_t^\xi =& \Big[rw_t^\xi - c_t^\xi + w_t^\xi z_t^{\xi\top}(\bar{D} + \mu_{Pt}^\xi - rP_t^\xi) \Big] dt + w_t^\xi z_t^{\xi\top}(\sigma_{Pt}^\xi + \sigma)^\top dB_t \\ &+ w_t^\xi z_t^{\xi\top} J_t^\xi dN_t^\xi. \end{aligned} \tag{A.25}$$

The investor's HJB equation is

$$\begin{aligned} \rho V^\xi = \max_{\hat{c}_t^\xi, \hat{z}_t^\xi} \Big\{ & \log \hat{c}_t^\xi + V_{\hat{w}}^\xi \mu_t^{\hat{w},\xi} + \frac{1}{2} V_{\hat{w}\hat{w}}^\xi (\sigma_t^{\hat{w},\xi})^\top \sigma_t^{\hat{w},\xi} + V_w^\xi \mu_t^{w,\xi} + \frac{1}{2} V_{ww}^\xi (\sigma_t^{w,\xi})^\top \sigma_t^{w,\xi} \\ & + V_{\hat{w}w}^\xi (\sigma_t^{\hat{w},\xi})^\top \sigma_t^{w,\xi} + \lambda^\xi (V^{-\xi,+} - V^\xi) \Big\}, \end{aligned} \tag{A.26}$$

where

$$\mu_t^{\hat{w},\xi} = r\hat{w}_t^\xi - \hat{c}_t^\xi + \hat{w}_t^\xi \hat{z}_t^{\xi\top}(\bar{D} + \mu_{Pt}^\xi - rP_t^\xi), \tag{A.27}$$

$$\sigma_t^{\hat{w},\xi} = (\sigma_{Pt}^\xi + \sigma)\hat{z}_t^\xi \hat{w}_t^\xi, \tag{A.28}$$

$$\mu_t^{w,\xi} = rw_t^\xi - c_t^\xi + w_t^\xi z_t^{\xi\top}(\bar{D}_t + \mu_{Pt}^\xi - rP_t^\xi), \tag{A.29}$$

$$\sigma_t^{w,\xi} = (\sigma_{Pt}^\xi + \sigma) z_t^\xi w_t^\xi. \tag{A.30}$$

We conjecture (and later verify) that the investor's value function takes the form

$$V^\xi(\hat{w}_t^\xi, w_t^\xi) = \frac{1}{\rho} \log \hat{w}_t^\xi + \bar{q}^\xi(w_t^\xi). \tag{A.31}$$

Substituting (A.31) into (A.26) yields

$$\begin{aligned} &\log \hat{w}_t^\xi + \rho \bar{q}^\xi(w_t^\xi) = \\ &\max_{\hat{c}_t^\xi, \hat{z}_t^\xi} \Big\{ \log \hat{c}_t^\xi + \frac{1}{\rho \hat{w}_t^\xi} \mu_t^{\hat{w},\xi} - \frac{1}{2\rho(\hat{w}_t^\xi)^2} (\sigma_t^{\hat{w},\xi})^\top \sigma^{\hat{w},\xi} + \bar{q}^{\xi\prime}(w_t^\xi) \mu_t^{w,\xi} + \frac{1}{2} \bar{q}^{\xi\prime\prime}(w_t^\xi) (\sigma_t^{w,\xi})^\top \sigma_t^{w,\xi} \\ &+ \lambda^\xi \Big[\frac{1}{\rho} \log \Big(\hat{w}_t^\xi + \hat{w}_t^\xi \hat{z}_t^{\xi\top}(P_t^{-\xi} - P_t^\xi) \Big) + \bar{q}^{-\xi} \Big(w_t^\xi + w^\xi z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi) \Big) - \frac{1}{\rho} \log \hat{w}_t^\xi - \bar{q}^\xi(w_t^\xi) \Big] \Big\}. \end{aligned} \tag{A.32}$$

Taking the first-order condition with respect to consumption yields

$$\frac{1}{\hat{c}_t^\xi} - \frac{1}{\rho \hat{w}_t^\xi} = 0. \tag{A.33}$$

Thus,

$$\hat{c}_t^\xi = \rho \hat{w}_t^\xi. \tag{A.34}$$

Taking the first-order condition with respect to the investor's wealth-scaled asset holdings implies that $\hat{z}_t^\xi$ satisfies

$$0 =(\bar{D} + \mu_{Pt}^\xi - rP_t^\xi) - (\sigma_{Pt}^\xi + \sigma)^\top(\sigma_{Pt}^\xi + \sigma)\hat{z}_t^\xi + \frac{\lambda^\xi(P_t^{-\xi} - P_t^\xi)}{1 + \hat{z}_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)}. \tag{A.35}$$

Rearranging, we get

$$\bar{D} + \mu_{Pt}^\xi - rP_t^\xi = (\sigma_{Pt}^\xi + \sigma)^\top(\sigma_{Pt}^\xi + \sigma)\hat{z}_t^\xi - \frac{\lambda^\xi(P_t^{-\xi} - P_t^\xi)}{1 + \hat{z}_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)}. \tag{A.36}$$

Then

$$\hat{z}_t^\xi = (\Pi_t^\xi)^{-1}(\bar{D} + \mu_{Pt}^\xi - rP_t^\xi) + \lambda^\xi \frac{(\Pi_t^\xi)^{-1}(P_t^{-\xi} - P_t^\xi)}{1 + \hat{z}_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)}. \tag{A.37}$$

From (A.37), since $\hat{y}_t^\xi = \mathrm{diag}(P_t^\xi)\hat{z}_t^\xi$,

$$\begin{aligned}\hat{y}_t^\xi &= (\Omega_t^\xi)^{-1}(\mu_{Rt}^\xi - r\mathbf{1}) + \lambda^\xi \frac{(\Omega_t^\xi)^{-1}J_{Rt}^\xi}{1 + \hat{y}_t^{\xi\top}J_{Rt}^\xi} \\ &= (\Omega_t^\xi)^{-1}\left(\mu_{Rt}^\xi - r\mathbf{1} + \lambda^\xi J_{Rt}^\xi\right) + \lambda^\xi(\Omega_t^\xi)^{-1}\left(\frac{1}{1 + \hat{y}_t^{\xi\top}J_{Rt}^\xi} - 1\right)J_{Rt}^\xi.\end{aligned} \tag{A.38}$$

A. ODEs

Note that $\hat{z}_t^\xi = z_t^\xi$, $\hat{c}_t^\xi = c_t^\xi$ and $\hat{w}_t^\xi = w_t^\xi$ in equilibrium, then (A.36) implies

$$\bar{D} + \mu_{Pt}^\xi - rP_t^\xi = (\sigma_{Pt}^\xi + \sigma)^\top(\sigma_{Pt}^\xi + \sigma)z_t^\xi - \frac{\lambda^\xi(P_t^{-\xi} - P_t^\xi)}{1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)}, \tag{A.39}$$

and from (A.34), $c_t^\xi = \rho w_t^\xi$. Substituting into the HJB equation (A.32) yields

$$\begin{aligned}&\rho\bar{q}^\xi(w_t^\xi) = \log\rho + \frac{1}{\rho}\Big[r - \rho + \hat{z}_t^{\xi\top}(\bar{D} + \mu_{Pt}^\xi - rP_t^\xi)\Big] \\ &+ \frac{1}{2\rho}\hat{z}_t^{\xi\top}(\sigma_{Pt}^\xi + \sigma)^\top(\sigma_{Pt}^\xi + \sigma)\hat{z}_t^\xi \\ &+ \bar{q}^{\xi\prime}(w_t^\xi)w_t^\xi\Big[r - \rho + z_t^{\xi\top}(\bar{D} + \mu_{Pt}^\xi - rP_t^\xi)\Big] \\ &+ \frac{1}{2}\bar{q}^{\xi\prime\prime}(w_t^\xi)(w_t^\xi)^2 z_t^{\xi\top}(\sigma_{Pt}^\xi + \sigma)^\top(\sigma_{Pt}^\xi + \sigma)z_t^\xi \\ &+ \lambda^\xi\Big[\frac{1}{\rho}\log(1 + \hat{z}_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)) + \bar{q}^{-\xi}\Big(w_t^\xi + w_t^\xi z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)\Big) - \bar{q}^\xi(w_t^\xi)\Big].\end{aligned} \tag{A.40}$$

Note that $\hat{w}_t^\xi$ have all canceled out. Thus, we have verified the conjectured value function

(A.31). Using the fact that $\hat{z}_t^\xi = z_t^\xi$, we get

$$\begin{aligned}\rho\bar{q}^\xi(w_t^\xi) =& \log\rho + \Big[\bar{q}^{\xi\prime}(w_t^\xi)w_t^\xi + \frac{1}{\rho}\Big]\Big[r - \rho + z_t^{\xi\top}(\bar{D} + \mu_{Pt}^\xi - rP_t^\xi)\Big] \\ &+ \frac{1}{2}\Big[\bar{q}^{\xi\prime\prime}(w_t^\xi)(w_t^\xi)^2 - \frac{1}{\rho}\Big]z_t^{\xi\top}(\sigma_{Pt}^\xi + \sigma)^\top(\sigma_{Pt}^\xi + \sigma)z_t^\xi \\ &+ \lambda^\xi\Big[\frac{1}{\rho}\log(1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)) + \bar{q}^{-\xi}\Big(w_t^\xi + w_t^\xi z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)\Big) - \bar{q}^\xi(w_t^\xi)\Big].\end{aligned} \tag{A.41}$$

Using (A.39),

$$\begin{aligned}&\rho\bar{q}^\xi(w_t^\xi) \\ &= \log\rho + \Big[\bar{q}^{\xi\prime}(w_t^\xi)w_t^\xi + \frac{1}{\rho}\Big]\Big[r - \rho + z_t^{\xi\top}(\sigma_{Pt}^\xi + \sigma)^\top(\sigma_{Pt}^\xi + \sigma)z_t^\xi\Big] \\ &\quad - \Big[\bar{q}^{\xi\prime}(w_t^\xi)w_t^\xi + \frac{1}{\rho}\Big]\frac{\lambda^\xi z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)}{1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)} \\ &\quad + \frac{1}{2}\Big[\bar{q}^{\xi\prime\prime}(w_t^\xi)(w_t^\xi)^2 - \frac{1}{\rho}\Big]z_t^{\xi\top}(\sigma_{Pt}^\xi + \sigma)^\top(\sigma_{Pt}^\xi + \sigma)z_t^\xi \\ &\quad + \lambda^\xi\Big[\frac{1}{\rho}\log(1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)) + \bar{q}^{-\xi}\Big(w_t^\xi + w_t^\xi z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)\Big) - \bar{q}^\xi(w_t^\xi)\Big].\end{aligned} \tag{A.42}$$

Collecting terms, we get

$$\begin{aligned}\rho\bar{q}^\xi(w_t^\xi) =& \log\rho + \Big[\bar{q}^{\xi\prime}(w_t^\xi)w_t^\xi + \frac{1}{\rho}\Big](r - \rho) \\ &+ \Big[\bar{q}^{\xi\prime}(w_t^\xi)w_t^\xi + \frac{1}{2}\bar{q}^{\xi\prime\prime}(w_t^\xi)(w_t^\xi)^2 + \frac{1}{2\rho}\Big]z_t^{\xi\top}(\sigma_{Pt}^\xi + \sigma)^\top(\sigma_{Pt}^\xi + \sigma)z_t^\xi \\ &+ \lambda^\xi\Big[\frac{1}{\rho}\log(1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)) + \bar{q}^{-\xi}\Big(w_t^\xi + w_t^\xi z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)\Big) - \bar{q}^\xi(w_t^\xi) \\ &\qquad - \Big(\bar{q}^{\xi\prime}(w_t^\xi)w_t^\xi + \frac{1}{\rho}\Big)\frac{z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)}{1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)}\Big].\end{aligned} \tag{A.43}$$

Substituting the investors' aggregate consumption c_t^ξ into the total wealth dynamic for investors, we get

$$dw_t^\xi = w_t^\xi\Big[r - \rho + z_t^{\xi\top}(\bar{D} + \mu_{Pt}^\xi - rP_t^\xi)\Big]dt + w_t^\xi z_t^{\xi\top}(\sigma_{Pt}^\xi + \sigma)^\top dB_t + w_t^\xi z_t^{\xi\top}J_t^\xi dN_t^\xi. \tag{A.44}$$

The investors' total wealth w_t^ξ and the state of the economy ξ are the two state variables in the model. Thus, the risky assets' prices in state ξ must be functions of the investors' total wealth in the state, that is $P_t^\xi \equiv P^\xi(w_t^\xi)$. Let

$$\tilde{\mu}_t^{w,\xi} \equiv w_t^\xi\big[r - \rho + z_t^{\xi\top}(\bar{D} + \mu_{Pt}^\xi - rP_t^\xi)\big], \tag{A.45}$$

$$\tilde{\sigma}_t^{w,\xi} \equiv (\sigma_{Pt}^\xi + \sigma)z_t^\xi w_t^\xi. \tag{A.46}$$

By Itô's lemma, we have

$$\begin{aligned} dP_t^\xi =& \Big[\tilde{\mu}_t^{w,\xi} P^{\xi\prime}(w_t^\xi) + \frac{1}{2}(\tilde{\sigma}_t^{w,\xi})^\top \tilde{\sigma}_t^{w,\xi} P^{\xi\prime\prime}(w_t^\xi)\Big] dt + P^{\xi\prime}(w_t^\xi)(\tilde{\sigma}_t^{w,\xi})^\top dB_t \\ &+ (P_t^{-\xi} - P_t^\xi) dN_t^\xi. \end{aligned} \tag{A.47}$$

From (A.45) and (A.46), the drift of the risky assets' prices become

$$\begin{aligned} &\tilde{\mu}_t^{w,\xi} P^{\xi\prime}(w_t^\xi) + \frac{1}{2}(\tilde{\sigma}_t^{w,\xi})^\top \tilde{\sigma}_t^{w,\xi} P^{\xi\prime\prime}(w_t^\xi) \\ &= w_t^\xi \big[r - \rho + z_t^{\xi\top}(\bar{D} + \mu_{Pt}^\xi - r P_t^\xi)\big] P^{\xi\prime}(w_t^\xi) + \frac{1}{2} z_t^{\xi\top}(\sigma_{Pt}^\xi + \sigma)^\top (\sigma_{Pt}^\xi + \sigma) z_t^\xi P^{\xi\prime\prime}(w_t^\xi)(w_t^\xi)^2 \\ &= \Big[r - \rho + z_t^{\xi\top}(\sigma_{Pt}^\xi + \sigma)^\top (\sigma_{Pt}^\xi + \sigma) z_t^\xi - \frac{\lambda^\xi z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)}{1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)}\Big] P^{\xi\prime}(w_t^\xi) w_t^\xi \\ &\quad + \frac{1}{2} z_t^{\xi\top}(\sigma_{Pt}^\xi + \sigma)^\top (\sigma_{Pt}^\xi + \sigma) z_t^\xi P^{\xi\prime\prime}(w_t^\xi)(w_t^\xi)^2 \\ &= \Big[r - \rho - \frac{\lambda^\xi z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)}{1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)}\Big] P^{\xi\prime}(w_t^\xi) w_t^\xi \\ &\quad + z_t^{\xi\top}(\sigma_{Pt}^\xi + \sigma)^\top (\sigma_{Pt}^\xi + \sigma) z_t^\xi \Big[P^{\xi\prime}(w_t^\xi) w_t^\xi + \frac{1}{2} P^{\xi\prime\prime}(w_t^\xi)(w_t^\xi)^2\Big]. \end{aligned} \tag{A.48}$$

The diffusion of the risky assets' prices are

$$\tilde{\sigma}_t^{w,\xi} P^{\xi\prime}(w_t^\xi)^\top = (\sigma_{Pt}^\xi + \sigma) z_t^\xi w_t^\xi P^{\xi\prime}(w_t^\xi)^\top. \tag{A.49}$$

The risky assets' prices are conjectured to follow the Itô process (6). By matching coefficients, we get

$$J_t^\xi = P_t^{-\xi} - P_t^\xi, \tag{A.50}$$

while μ_{Pt}^ξ and σ_{Pt}^ξ are respectively given by

$$\begin{aligned} \mu_{Pt}^\xi =& \Big[r - \rho - \frac{\lambda^\xi z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)}{1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)}\Big] P^{\xi\prime}(w_t^\xi) w_t^\xi \\ &+ z_t^{\xi\top}(\sigma_{Pt}^\xi + \sigma)^\top (\sigma_{Pt}^\xi + \sigma) z_t^\xi \Big[P^{\xi\prime}(w_t^\xi) w_t^\xi + \frac{1}{2} P^{\xi\prime\prime}(w_t^\xi)(w_t^\xi)^2\Big], \end{aligned} \tag{A.51}$$

and

$$\sigma_{Pt}^\xi = (\sigma_{Pt}^\xi + \sigma) z_t^\xi w_t^\xi P^{\xi\prime}(w_t^\xi)^\top. \tag{A.52}$$

From (A.52),

$$(\sigma_{Pt}^\xi + \sigma) z_t^\xi = (\sigma_{Pt}^\xi + \sigma) z_t^\xi P^{\xi\prime}(w_t^\xi)^\top z_t^\xi w_t^\xi + \sigma z_t^\xi. \tag{A.53}$$

Thus,

$$\big[1 - P^{\xi\prime}(w_t^\xi)^\top z_t^\xi w_t^\xi\big](\sigma_{Pt}^\xi + \sigma) z_t^\xi = \sigma z_t^\xi. \tag{A.54}$$

Then

$$(\sigma_{Pt}^\xi + \sigma) z_t^\xi = \frac{\sigma z_t^\xi}{1 - P^{\xi\prime}(w_t^\xi)^\top z_t^\xi w_t^\xi}. \tag{A.55}$$

Substituting (A.55) into (A.52) yields

$$\sigma_{Pt}^\xi = \frac{\sigma z_t^\xi P^{\xi\prime}(w_t^\xi)^\top w_t^\xi}{1 - P^{\xi\prime}(w_t^\xi)^\top z_t^\xi w_t^\xi}. \tag{A.56}$$

Adding σ to both sides yields

$$\sigma_{Pt}^\xi + \sigma = \frac{\sigma z_t^\xi P^{\xi\prime}(w_t^\xi)^\top w_t^\xi + [1 - P^{\xi\prime}(w_t^\xi)^\top z_t^\xi w_t^\xi]\sigma}{1 - P^{\xi\prime}(w_t^\xi)^\top z_t^\xi w_t^\xi}. \tag{A.57}$$

Let

$$\hat{\sigma}_t^\xi \equiv \sigma z_t^\xi P^{\xi\prime}(w_t^\xi)^\top w_t^\xi + [1 - P^{\xi\prime}(w_t^\xi)^\top z_t^\xi w_t^\xi]\sigma. \tag{A.58}$$

Then

$$\sigma_{Pt}^\xi + \sigma = \frac{\hat{\sigma}_t^\xi}{1 - P^{\xi\prime}(w_t^\xi)^\top z_t^\xi w_t^\xi}. \tag{A.59}$$

Substituting (A.55) into (A.51) yields

$$\begin{aligned}\mu_{Pt}^\xi =& \Big[r - \rho - \frac{\lambda^\xi z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)}{1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)}\Big] P^{\xi\prime}(w_t^\xi) w_t^\xi \\ &+ \frac{z_t^{\xi\top}\Sigma z_t^\xi}{[1 - P^{\xi\prime}(w_t^\xi)^\top z_t^\xi w_t^\xi]^2}\Big[P^{\xi\prime}(w_t^\xi) w_t^\xi + \frac{1}{2} P^{\xi\prime\prime}(w_t^\xi)(w_t^\xi)^2\Big].\end{aligned} \tag{A.60}$$

Thus

$$\begin{aligned}&\bar{D} + \mu_{Pt}^\xi - r P_t^\xi \\ &= \bar{D} - r P^\xi(w_t^\xi) + \Big[r - \rho - \frac{\lambda^\xi z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)}{1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)}\Big] P^{\xi\prime}(w_t^\xi) w_t^\xi \\ &\quad + \frac{z_t^{\xi\top}\Sigma z_t^\xi}{[1 - P^{\xi\prime}(w_t^\xi)^\top z_t^\xi w_t^\xi]^2}\Big[P^{\xi\prime}(w_t^\xi) w_t^\xi + \frac{1}{2} P^{\xi\prime\prime}(w_t^\xi)(w_t^\xi)^2\Big].\end{aligned} \tag{A.61}$$

From (A.39), and using (A.55) and (A.59), we have

$$\bar{D} + \mu_{Pt}^{\xi} - rP_t^{\xi} = \frac{\hat{\sigma}_t^{\xi\top}\sigma z_t^{\xi}}{[1 - P^{\xi\prime}(w_t^{\xi})^{\top} z_t^{\xi} w_t^{\xi}]^2} - \frac{\lambda^{\xi}(P_t^{-\xi} - P_t^{\xi})}{1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})}. \tag{A.62}$$

Hence,

$$\begin{aligned} &\bar{D} - rP^{\xi}(w_t^{\xi}) + \Big[r - \rho - \frac{\lambda^{\xi} z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})}{1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})}\Big] P^{\xi\prime}(w_t^{\xi}) w_t^{\xi} \\ &+ \frac{z_t^{\xi\top}\Sigma z_t^{\xi}}{[1 - P^{\xi\prime}(w_t^{\xi})^{\top} z_t^{\xi} w_t^{\xi}]^2}\Big[P^{\xi\prime}(w_t^{\xi}) w_t^{\xi} + \frac{1}{2} P^{\xi\prime\prime}(w_t^{\xi})(w_t^{\xi})^2\Big] \\ &= \frac{\hat{\sigma}_t^{\xi\top}\sigma z_t^{\xi}}{[1 - P^{\xi\prime}(w_t^{\xi})^{\top} z_t^{\xi} w_t^{\xi}]^2} - \frac{\lambda^{\xi}(P_t^{-\xi} - P_t^{\xi})}{1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})}. \end{aligned} \tag{A.63}$$

The equation (A.63) follows from the investors' first-order condition and the Itô's lemma. Moreover, from (A.62) we have

$$z_t^{\xi\top}(\bar{D} + \mu_{Pt}^{\xi} - rP_t^{\xi}) = \frac{z_t^{\xi\top}\Sigma z_t^{\xi}}{[1 - P^{\xi\prime}(w_t^{\xi})^{\top} z_t^{\xi} w_t^{\xi}]^2} - \frac{\lambda^{\xi} z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})}{1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})}. \tag{A.64}$$

Since there is a measure one of symmetric hedgers, the aggregate risky asset holdings by hedgers at time t in state ξ are given by

$$\begin{aligned} Q_t^{\xi} =& \Big[(\sigma_{Pt}^{\xi} + \sigma)^{\top}(\sigma_{Pt} + \sigma) + \lambda^{\xi} J_t^{\xi} J_t^{\xi\top}\Big]^{-1}\Big[\frac{\bar{D} + \mu_{Pt}^{\xi} - rP_t^{\xi} + \lambda^{\xi} J_t^{\xi}}{\alpha} - (\sigma_{Pt}^{\xi} + \sigma)^{\top}\sigma u^{\xi}\Big] \\ =& \frac{1}{\alpha}\Big\{\frac{\hat{\sigma}_t^{\xi\top}\hat{\sigma}_t^{\xi}}{[1 - P^{\xi\prime}(w_t^{\xi})^{\top} z_t^{\xi} w_t^{\xi}]^2} + \lambda^{\xi}(P_t^{-\xi} - P_t^{\xi})(P_t^{-\xi} - P_t^{\xi})^{\top}\Big\}^{-1}\Big\{\lambda^{\xi}(P_t^{-\xi} - P_t^{\xi}) \\ &+ \bar{D} - rP^{\xi}(w_t^{\xi}) + \Big[r - \rho - \frac{\lambda^{\xi} z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})}{1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})}\Big] P^{\xi\prime}(w_t^{\xi}) w_t^{\xi} \\ &+ \frac{z_t^{\xi\top}\Sigma z_t^{\xi}}{[1 - P^{\xi\prime}(w_t^{\xi})^{\top} z_t^{\xi} w_t^{\xi}]^2}\Big[P^{\xi\prime}(w_t^{\xi}) w_t^{\xi} + \frac{1}{2} P^{\xi\prime\prime}(w_t^{\xi})(w_t^{\xi})^2\Big] - \frac{\alpha\hat{\sigma}_t^{\xi\top}\sigma u^{\xi}}{1 - P^{\xi\prime}(w_t^{\xi})^{\top} z_t^{\xi} w_t^{\xi}}\Big\}. \end{aligned} \tag{A.65}$$

Market clearing requires that

$$Q_t^{\xi} + Y_t^{\xi} = Q_t^{\xi} + z_t^{\xi} w_t^{\xi} = S. \tag{A.66}$$

That is,

$$\begin{aligned}
&\frac{1}{\alpha}\Big\{\frac{\hat{\sigma}_t^{\xi\top}\hat{\sigma}_t^{\xi}}{[1-P^{\xi\prime}(w_t^{\xi})^{\top}z_t^{\xi}w_t^{\xi}]^2}+\lambda^{\xi}(P_t^{-\xi}-P_t^{\xi})(P_t^{-\xi}-P_t^{\xi})^{\top}\Big\}^{-1}\Big\{\lambda^{\xi}(P_t^{-\xi}-P_t^{\xi})\\
&+\bar{D}-rP^{\xi}(w_t^{\xi})+\Big[r-\rho-\frac{\lambda^{\xi}z_t^{\xi\top}(P_t^{-\xi}-P_t^{\xi})}{1+z_t^{\xi\top}(P_t^{-\xi}-P_t^{\xi})}\Big]P^{\xi\prime}(w_t^{\xi})w_t^{\xi}\\
&+\frac{z_t^{\xi\top}\Sigma z_t^{\xi}}{[1-P^{\xi\prime}(w_t^{\xi})^{\top}z_t^{\xi}w_t^{\xi}]^2}\Big[P^{\xi\prime}(w_t^{\xi})w_t^{\xi}+\frac{1}{2}P^{\xi\prime\prime}(w_t^{\xi})(w_t^{\xi})^2\Big]-\frac{\alpha\hat{\sigma}_t^{\xi\top}\sigma u^{\xi}}{1-P^{\xi\prime}(w_t^{\xi})^{\top}z_t^{\xi}w_t^{\xi}}\Big\}\\
&=S-z_t^{\xi}w_t^{\xi}.
\end{aligned}\tag{A.67}$$

The equation (A.67) follows from the market clearing condition.

When the economy switches from state ξ to $-\xi$ at time t, the investor's wealth jumps from w_t^{ξ} to $w_t^{-\xi}$ according to the following fixed point condition

$$w_t^{-\xi}=w_t^{\xi}+w_t^{\xi}z_t^{\xi\top}(P_t^{-\xi}-P_t^{\xi}).\tag{A.68}$$

Substituting (A.68) into (A.63) yields

$$\begin{aligned}
&\bar{D}-rP^{\xi}(w_t^{\xi})+\left[r-\rho-\frac{\lambda^{\xi}\frac{w_t^{-\xi}-w_t^{\xi}}{w_t^{\xi}}}{\frac{w_t^{-\xi}}{w_t^{\xi}}}\right]P^{\xi\prime}(w_t^{\xi})w_t^{\xi}\\
&+\frac{z_t^{\xi\top}\Sigma z_t^{\xi}}{[1-P^{\xi\prime}(w_t^{\xi})^{\top}z_t^{\xi}w_t^{\xi}]^2}\Big[P^{\xi\prime}(w_t^{\xi})w_t^{\xi}+\frac{1}{2}P^{\xi\prime\prime}(w_t^{\xi})(w_t^{\xi})^2\Big]\\
&=\frac{\hat{\sigma}_t^{\xi\top}\sigma z_t^{\xi}}{[1-P^{\xi\prime}(w_t^{\xi})^{\top}z_t^{\xi}w_t^{\xi}]^2}-\frac{\lambda^{\xi}(P_t^{-\xi}-P_t^{\xi})}{\frac{w_t^{-\xi}}{w_t^{\xi}}}.
\end{aligned}\tag{A.69}$$

Furthermore, substituting (A.68) into (A.67) yields

$$\begin{aligned}
&\bar{D}-rP^{\xi}(w_t^{\xi})+\left[r-\rho-\frac{\lambda^{\xi}\frac{w_t^{-\xi}-w_t^{\xi}}{w_t^{\xi}}}{\frac{w_t^{-\xi}}{w_t^{\xi}}}\right]P^{\xi\prime}(w_t^{\xi})w_t^{\xi}\\
&+\frac{z_t^{\xi\top}\Sigma z_t^{\xi}}{[1-P^{\xi\prime}(w_t^{\xi})^{\top}z_t^{\xi}w_t^{\xi}]^2}\Big[P^{\xi\prime}(w_t^{\xi})w_t^{\xi}+\frac{1}{2}P^{\xi\prime\prime}(w_t^{\xi})(w_t^{\xi})^2\Big]\\
&=\alpha\Big\{\frac{\hat{\sigma}_t^{\xi\top}\hat{\sigma}_t^{\xi}}{[1-P^{\xi\prime}(w_t^{\xi})^{\top}z_t^{\xi}w_t^{\xi}]^2}+\lambda^{\xi}(P_t^{-\xi}-P_t^{\xi})(P_t^{-\xi}-P_t^{\xi})^{\top}\Big\}(S-z_t^{\xi}w_t^{\xi})\\
&\quad+\frac{\alpha\hat{\sigma}_t^{\xi\top}\sigma u^{\xi}}{1-P^{\xi\prime}(w_t^{\xi})^{\top}z_t^{\xi}w_t^{\xi}}-\lambda^{\xi}(P_t^{-\xi}-P_t^{\xi}).
\end{aligned}\tag{A.70}$$

Combining (A.69) and (A.70) yields

$$\begin{aligned}
&\frac{\hat{\sigma}_t^{\xi\top}\sigma z_t^\xi}{[1-P^{\xi\prime}(w_t^\xi)^\top z_t^\xi w_t^\xi]^2} - \frac{\lambda^\xi(P_t^{-\xi}-P_t^\xi)}{\frac{w_t^{-\xi}}{w_t^\xi}} \\
&= \frac{\alpha\hat{\sigma}_t^{\xi\top}\hat{\sigma}_t^\xi(S-z_t^\xi w_t^\xi)}{[1-P^{\xi\prime}(w_t^\xi)^\top z_t^\xi w_t^\xi]^2} + \alpha\lambda^\xi(P_t^{-\xi}-P_t^\xi)(P_t^{-\xi}-P_t^\xi)^\top(S-z_t^\xi w_t^\xi) \\
&\quad + \frac{\alpha\hat{\sigma}_t^{\xi\top}\sigma u^\xi}{1-P^{\xi\prime}(w_t^\xi)^\top z_t^\xi w_t^\xi} - \lambda^\xi(P_t^{-\xi}-P_t^\xi).
\end{aligned} \tag{A.71}$$

Rearranging yields

$$\begin{aligned}
&\frac{(1+\alpha w_t^\xi)\hat{\sigma}_t^{\xi\top}\sigma z_t^\xi - \alpha\hat{\sigma}_t^{\xi\top}\hat{\sigma}_t^\xi S}{[1-P^{\xi\prime}(w_t^\xi)^\top z_t^\xi w_t^\xi]^2} - \frac{\alpha\hat{\sigma}_t^{\xi\top}\sigma u^\xi}{1-P^{\xi\prime}(w_t^\xi)^\top z_t^\xi w_t^\xi} \\
&= \lambda^\xi\left[\left(\frac{1}{\frac{w_t^{-\xi}}{w_t^\xi}}-1\right)(P_t^{-\xi}-P_t^\xi) + \alpha(P_t^{-\xi}-P_t^\xi)(P_t^{-\xi}-P_t^\xi)^\top(S-z_t^\xi w_t^\xi)\right].
\end{aligned} \tag{A.72}$$

Note that

$$\hat{\sigma}_t^{\xi\top}\sigma z_t^\xi = w_t^\xi P^{\xi\prime}(w_t^\xi) z_t^{\xi\top}\Sigma z_t^\xi + [1-w_t^\xi P^{\xi\prime}(w_t^\xi)^\top z_t^\xi]\Sigma z_t^\xi. \tag{A.73}$$

Hence, (A.69) simplifies to

$$\begin{aligned}
&\bar{D} - rP^\xi(w_t^\xi) + \left[r-\rho-\frac{\lambda^\xi\frac{w_t^{-\xi}-w_t^\xi}{w_t^\xi}}{\frac{w_t^{-\xi}}{w_t^\xi}}\right]P^{\xi\prime}(w_t^\xi)w_t^\xi + \frac{z_t^{\xi\top}\Sigma z_t^\xi}{2[1-P^{\xi\prime}(w_t^\xi)^\top z_t^\xi w_t^\xi]^2}P^{\xi\prime\prime}(w_t^\xi)(w_t^\xi)^2 \\
&= \frac{\Sigma z_t^\xi}{1-P^{\xi\prime}(w_t^\xi)^\top z_t^\xi w_t^\xi} - \frac{\lambda^\xi(P_t^{-\xi}-P_t^\xi)}{\frac{w_t^{-\xi}}{w_t^\xi}}.
\end{aligned} \tag{A.74}$$

Moreover, $\bar{q}^\xi(w_t^\xi)$ is solution to the ODE (A.43), which can be rewritten as

$$\begin{aligned}
&\log\rho + \left[\bar{q}^{\xi\prime}(w_t^\xi)w_t^\xi + \frac{1}{\rho}\right](r-\rho) - \rho\bar{q}^\xi(w_t^\xi) \\
&+ \left[\bar{q}^{\xi\prime}(w_t^\xi)w_t^\xi + \frac{1}{2}\bar{q}^{\xi\prime\prime}(w_t^\xi)(w_t^\xi)^2 + \frac{1}{2\rho}\right]\frac{z_t^{\xi\top}\Sigma z_t^\xi}{[1-P^{\xi\prime}(w_t^\xi)^\top z_t^\xi w_t^\xi]^2} \\
&+ \lambda^\xi\left[\frac{1}{\rho}\log\left(\frac{w_t^{-\xi}}{w_t^\xi}\right) + \bar{q}^{-\xi}(w_t^{-\xi}) - \bar{q}^\xi(w_t^\xi) - \left(\bar{q}^{\xi\prime}(w_t^\xi)w_t^\xi + \frac{1}{\rho}\right)\frac{\frac{w_t^{-\xi}-w_t^\xi}{w_t^\xi}}{\frac{w_t^{-\xi}}{w_t^\xi}}\right] = 0.
\end{aligned} \tag{A.75}$$

The investors' wealth-scaled asset holdings z_t^ξ are determined by (A.72), while prices

$P^\xi(w_t^\xi)$ and hedging demand $\bar{q}^\xi(w_t^\xi)$ are solutions to the ODEs (A.74) and (A.75). Define

$$\pi^\xi(w_t^\xi) \equiv \frac{\bar{D}}{r} - P^\xi(w_t^\xi). \tag{A.76}$$

Then the investors' wealth-scaled asset holdings z_t^ξ satisfies

$$\begin{aligned}
&\frac{(1+\alpha w_t^\xi)\hat{\sigma}_t^{\xi\top}\sigma z_t^\xi - \alpha\hat{\sigma}_t^{\xi\top}\hat{\sigma}_t^\xi S}{[1+w_t^\xi\pi^{\xi\prime}(w_t^\xi)^\top z_t^\xi]^2} - \frac{\alpha\hat{\sigma}_t^{\xi\top}\sigma u^\xi}{1+w_t^\xi\pi^{\xi\prime}(w_t^\xi)^\top z_t^\xi} \\
&= \lambda^\xi\left[\Big(\frac{w_t^\xi}{w_t^{-\xi}} - 1\Big)(P_t^{-\xi} - P_t^\xi) + \alpha(P_t^{-\xi} - P_t^\xi)(P_t^{-\xi} - P_t^\xi)^\top(S - z_t^\xi w_t^\xi)\right],
\end{aligned} \tag{A.77}$$

where

$$\hat{\sigma}_t^\xi = -w_t^\xi\sigma z_t^\xi\pi^{\xi\prime}(w_t^\xi)^\top + \big[1+w_t^\xi\pi^{\xi\prime}(w_t^\xi)^\top z_t^\xi\big]\sigma, \tag{A.78}$$

$$P_t^{-\xi} - P_t^\xi = \pi^\xi(w_t^\xi) - \pi^{-\xi}(w_t^{-\xi}). \tag{A.79}$$

The functions $\pi^\xi(w_t^\xi)$ and $q^\xi(w_t^\xi)$ are solutions to the following ODEs

$$\begin{aligned}
&r\pi^\xi(w_t^\xi) - \Big[r - \rho - \frac{\lambda^\xi(w_t^{-\xi} - w_t^\xi)}{w_t^{-\xi}}\Big]\pi^{\xi\prime}(w_t^\xi)w_t^\xi - \frac{z_t^{\xi\top}\Sigma z_t^\xi}{2[1+w_t^\xi\pi^{\xi\prime}(w_t^\xi)^\top z_t^\xi]}\pi^{\xi\prime\prime}(w_t^\xi)(w_t^\xi)^2 \\
&= \frac{\Sigma z_t^\xi}{1+w_t^\xi\pi^{\xi\prime}(w_t^\xi)^\top z_t^\xi} - \frac{\lambda^\xi[\pi^\xi(w_t^\xi) - \pi^{-\xi}(w_t^{-\xi})]}{\frac{w_t^{-\xi}}{w_t^\xi}},
\end{aligned} \tag{A.80}$$

and

$$\begin{aligned}
&\log\rho + \Big[\bar{q}^{\xi\prime}(w_t^\xi)w_t^\xi + \frac{1}{\rho}\Big](r-\rho) - \rho\bar{q}^\xi(w_t^\xi) \\
&+ \Big[\bar{q}^{\xi\prime}(w_t^\xi)w_t^\xi + \frac{1}{2}\bar{q}^{\xi\prime\prime}(w_t^\xi)(w_t^\xi)^2 + \frac{1}{2\rho}\Big]\frac{z_t^{\xi\top}\Sigma z_t^\xi}{[1+w_t^\xi\pi^{\xi\prime}(w_t^\xi)^\top z_t^\xi]^2} \\
&+ \lambda^\xi\left[\frac{1}{\rho}\log\Big(\frac{w_t^{-\xi}}{w_t^\xi}\Big) + \bar{q}^{-\xi}(w_t^{-\xi}) - \bar{q}^\xi(w_t^\xi) - \Big(\bar{q}^{\xi\prime}(w_t^\xi)w_t^\xi + \frac{1}{\rho}\Big)\frac{w_t^{-\xi} - w_t^\xi}{w_t^{-\xi}}\right] = 0.
\end{aligned} \tag{A.81}$$

Moreover, the following fixed point condition must be satisfied

$$w_t^{-\xi} - w_t^\xi = w_t^\xi z_t^{\xi\top}\big[\pi^\xi(w_t^\xi) - \pi^{-\xi}(w_t^{-\xi})\big]. \tag{A.82}$$

B. Limit Cases with $\lambda \to 0$ and $\lambda \to \infty$

The log investor's optimal portfolio policy $\hat{y}_t^\xi$ is given by (A.38). Our setup focuses on $\lambda^n = 0$ and $\lambda^s = \lambda$. Since we define $\mathbb{E}_t(R^\xi) \equiv \mathbb{E}_t(dR_t^\xi)/dt = \mu_{Rt}^\xi + \lambda^\xi J_{Rt}^\xi$, we can then rewrite

$\hat{y}_t^\xi$ as

$$\hat{y}_t^\xi = (\Omega_t^\xi)^{-1}\left(\mathbb{E}_t(R^\xi) - r\mathbf{1}\right) + \lambda^\xi(\Omega_t^\xi)^{-1}\left(\frac{1}{1+\hat{y}_t^{\xi\top}J_{Rt}^\xi} - 1\right)J_{Rt}^\xi. \tag{A.83}$$

It is immediate that in state n,

$$\hat{y}_t^n = (\Omega_t^n)^{-1}\left(\mathbb{E}_t(R^n) - r\mathbf{1}\right). \tag{A.84}$$

Similarly, when $\lambda \to 0$, the investor's demand in state s is

$$\lim_{\lambda\to 0}\hat{y}_t^s = (\Omega_t^s)^{-1}\left(\mathbb{E}_t(R^s) - r\mathbf{1}\right). \tag{A.85}$$

On the other hand, when $\lambda \to \infty$, $J_{Rt}^s \to 0$. Thus, the second term on the right-hand side of (A.83) converges to zero, and

$$\lim_{\lambda\to\infty}\hat{y}_t^s = (\Omega_t^s)^{-1}\left(\mathbb{E}_t(R^s) - r\mathbf{1}\right). \tag{A.86}$$

Consequently, both in the limits $\lambda \to 0$ and $\lambda \to \infty$, we have

$$\hat{y}_t^\xi = (\Omega_t^\xi)^{-1}\big(\mathbb{E}_t(R^\xi) - r\mathbf{1}\big). \tag{A.87}$$

A.1.2 CRRA Investors

We now consider the case with CRRA investors. An CRRA investor maximizes

$$\mathbb{E}_t\Big[\int_t^\infty e^{-\rho(\tau-t)}\frac{\hat{c}_\tau^{1-\gamma}}{1-\gamma}d\tau\Big], \tag{A.88}$$

subject to the wealth dynamic (A.8). The investor's HJB equation can be written as

$$\begin{aligned}\rho V^\xi = \max_{\hat{c}_t^\xi,\hat{z}_t^\xi}\Big\{&\frac{\hat{c}_t^{1-\gamma}}{1-\gamma} + V_{\hat{w}}^\xi\mu_t^{\hat{w},\xi} + \frac{1}{2}V_{\hat{w}\hat{w}}^\xi(\sigma_t^{\hat{w},\xi})^\top\sigma_t^{\hat{w},\xi} + V_w^\xi\mu_t^{w,\xi} + \frac{1}{2}V_{ww}^\xi(\sigma_t^{w,\xi})^\top\sigma_t^{w,\xi}\\ &+ V_{\hat{w}w}^\xi(\sigma_t^{\hat{w},\xi})^\top\sigma_t^{w,\xi} + \lambda^\xi(V^{-\xi,+} - V^\xi)\Big\},\end{aligned} \tag{A.89}$$

where $\mu_t^{\hat{w},\xi}$, $\sigma_t^{\hat{w},\xi}$, $\mu_t^{w,\xi}$ and $\sigma_t^{w,\xi}$ are given by (A.27)-(A.30).

We conjecture (and later verify) that the investor's value function takes the form

$$V^\xi(\hat{w}_t^\xi, w_t^\xi) = q^\xi(w_t^\xi)\frac{(\hat{w}_t^\xi)^{1-\gamma}}{1-\gamma}. \tag{A.90}$$

Substituting the conjectured value function into the investor's HJB equation (A.89) yields

$$\begin{aligned}\rho q^{\xi}(w_t^{\xi})\frac{(\hat{w}_t^{\xi})^{1-\gamma}}{1-\gamma} = \max_{\hat{c}_t^{\xi},\hat{z}_t^{\xi}} \Big\{ &\frac{(\hat{c}_t^{\xi})^{1-\gamma}}{1-\gamma} + q^{\xi}(w_t^{\xi})(\hat{w}_t^{\xi})^{-\gamma}\mu_t^{\hat{w},\xi} - \frac{\gamma}{2}q^{\xi}(w_t^{\xi})(\hat{w}_t^{\xi})^{-\gamma-1}(\sigma_t^{\hat{w},\xi})^{\top}\sigma_t^{\hat{w},\xi} \\ &+ q^{\xi\prime}(w_t^{\xi})\frac{(\hat{w}_t^{\xi})^{1-\gamma}}{1-\gamma}\mu_t^{w,\xi} + \frac{1}{2}q^{\xi\prime\prime}(w_t^{\xi})\frac{(\hat{w}_t^{\xi})^{1-\gamma}}{1-\gamma}(\sigma_t^{w,\xi})^{\top}\sigma_t^{w,\xi} + q^{\xi\prime}(w_t^{\xi})(\hat{w}_t^{\xi})^{-\gamma}(\sigma_t^{\hat{w},t})^{\top}\sigma_t^{w,\xi} \\ &+ \lambda^{\xi}\Big[q^{-\xi}\big(w_t^{\xi} + w_t^{\xi}z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})\big)\frac{[\hat{w}_t^{\xi} + \hat{w}_t^{\xi}\hat{z}_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})]^{1-\gamma}}{1-\gamma} - q^{\xi}(w_t^{\xi})\frac{(\hat{w}_t^{\xi})^{1-\gamma}}{1-\gamma}\Big]\Big\}.\end{aligned} \tag{A.91}$$

Taking the first-order condition with respect to consumption yields

$$(\hat{c}_t^{\xi})^{-\gamma} - q^{\xi}(w_t^{\xi})(\hat{w}_t^{\xi})^{-\gamma} = 0. \tag{A.92}$$

Thus,

$$\hat{c}_t^{\xi} = q^{\xi}(w_t^{\xi})^{-\frac{1}{\gamma}}\hat{w}_t^{\xi}. \tag{A.93}$$

Let $\hat{z}_t^{\xi} \equiv \hat{Y}_t^{\xi}/\hat{w}_t^{\xi}$ denote the investor's optimal wealth-scaled asset holdings at time t if the economy is in state ξ. Let $z_t^{\xi} \equiv Y_t^{\xi}/w_t^{\xi}$ denote the investors' aggregate wealth-scaled asset holdings. Taking the first-order condition, $\hat{z}_t^{\xi}$ satisfies

$$\begin{aligned}0 =&q^{\xi}(w_t^{\xi})(\hat{w}_t^{\xi})^{1-\gamma}(\bar{D} + \mu_{Pt}^{\xi} - rP_t^{\xi}) - \gamma q^{\xi}(w_t^{\xi})(\hat{w}_t^{\xi})^{1-\gamma}(\sigma_{Pt}^{\xi} + \sigma)^{\top}(\sigma_{Pt}^{\xi} + \sigma)\hat{z}_t^{\xi} \\ &+ q^{\xi\prime}(w_t^{\xi})(\hat{w}_t^{\xi})^{1-\gamma}(\sigma_{Pt}^{\xi} + \sigma)^{\top}(\sigma_{Pt}^{\xi} + \sigma)z_t^{\xi}w_t^{\xi} \\ &+ \lambda^{\xi}q^{-\xi}\big(w_t^{\xi} + w_t^{\xi}z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})\big)[1 + \hat{z}_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})]^{-\gamma}(\hat{w}_t^{\xi})^{1-\gamma}(P_t^{-\xi} - P_t^{\xi}).\end{aligned} \tag{A.94}$$

Rearranging, we get

$$\begin{aligned}\bar{D} + \mu_{Pt}^{\xi} - rP_t^{\xi} =&\gamma(\sigma_{Pt}^{\xi} + \sigma)^{\top}(\sigma_{Pt}^{\xi} + \sigma)\hat{z}_t^{\xi} - \frac{q^{\xi\prime}(w_t^{\xi})}{q^{\xi}(w_t^{\xi})}w_t^{\xi}(\sigma_{Pt}^{\xi} + \sigma)^{\top}(\sigma_{Pt}^{\xi} + \sigma)z_t^{\xi} \\ &- \frac{\lambda^{\xi}q^{-\xi}\big(w_t^{\xi}[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})]\big)(P_t^{-\xi} - P_t^{\xi})}{q^{\xi}(w_t^{\xi})[1 + \hat{z}_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})]^{\gamma}},\end{aligned} \tag{A.95}$$

or equivalently

$$\begin{aligned}\hat{z}_t^{\xi} =&\frac{1}{\gamma}(\Pi_t^{\xi})^{-1}(\bar{D} + \mu_{Pt}^{\xi} - rP_t^{\xi}) + \frac{1}{\gamma}\frac{q^{\xi\prime}(w_t^{\xi})}{q^{\xi}(w_t^{\xi})}w_t^{\xi}z_t^{\xi} \\ &+ \frac{\lambda^{\xi}}{\gamma}\frac{q^{-\xi}(w_t^{\xi} + w_t^{\xi}z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi}))}{q^{\xi}(w_t^{\xi})[1 + \hat{z}_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})]^{\gamma}}(\Pi_t^{\xi})^{-1}(P_t^{-\xi} - P_t^{\xi}),\end{aligned} \tag{A.96}$$

where $\Pi_t^\xi \equiv (\sigma_{Pt}^\xi + \sigma)^\top (\sigma_{Pt}^\xi + \sigma)$. Since $\hat{y}_t^\xi = \text{diag}(P_t^\xi)\hat{z}_t^\xi$, then

$$\begin{aligned}\hat{y}_t^\xi =& \frac{1}{\gamma}(\Omega_t^\xi)^{-1}(\mu_{Rt}^\xi - r\mathbf{1} + \lambda^\xi J_{Rt}^\xi) + \frac{1}{\gamma}\frac{q^{\xi\prime}(w_t^\xi)}{q^\xi(w_t^\xi)} w_t^\xi y_t^\xi \\ &+ \frac{\lambda^\xi}{\gamma}(\Omega_t^\xi)^{-1}\left(\frac{q^{-\xi}(w_t^\xi(1 + y_t^{\xi\top} J_{Rt}^\xi))}{q^\xi(w_t^\xi)[1 + \hat{y}_t^{\xi\top} J_{Rt}^\xi]^\gamma} - 1\right) J_{Rt}^\xi.\end{aligned} \tag{A.97}$$

Note that $\hat{z}_t^\xi = z_t^\xi$, $\hat{c}_t^\xi = c_t^\xi$ and $\hat{w}_t^\xi = w_t^\xi$ in equilibrium, then (A.95) implies

$$\begin{aligned}\bar{D} + \mu_{Pt}^\xi - rP_t^\xi =& \Big[\gamma - \frac{q^{\xi\prime}(w_t^\xi)}{q^\xi(w_t^\xi)} w_t^\xi\Big](\sigma_{Pt}^\xi + \sigma)^\top (\sigma_{Pt}^\xi + \sigma) z_t^\xi \\ &- \frac{\lambda^\xi q^{-\xi}\big(w_t^\xi[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)]\big)(P_t^{-\xi} - P_t^\xi)}{q^\xi(w_t^\xi)[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)]^\gamma}.\end{aligned} \tag{A.98}$$

Let

$$\Gamma^\xi(w_t^\xi) \equiv \gamma - \frac{q^{\xi\prime}(w_t^\xi)}{q^\xi(w_t^\xi)} w_t^\xi, \tag{A.99}$$

Then, the above equation can be rewritten as

$$\begin{aligned}\bar{D} + \mu_{Pt}^\xi - rP_t^\xi =& \Gamma^\xi(w_t^\xi)(\sigma_{Pt}^\xi + \sigma)^\top (\sigma_{Pt}^\xi + \sigma) z_t^\xi \\ &- \frac{\lambda^\xi q^{-\xi}\big(w_t^\xi[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)]\big)(P_t^{-\xi} - P_t^\xi)}{q^\xi(w_t^\xi)[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)]^\gamma}.\end{aligned} \tag{A.100}$$

Moreover, from (A.93), the investors' total consumption aggregates to $c_t^\xi = q^\xi(w_t^\xi)^{-\frac{1}{\gamma}} w_t^\xi$. Substituting it into the HJB equation (A.91) yields

$$\begin{aligned}&\rho q^\xi(w_t^\xi)\frac{(\hat{w}_t^\xi)^{1-\gamma}}{1-\gamma} = \frac{q^\xi(w_t^\xi)^{-\frac{1-\gamma}{\gamma}}(\hat{w}_t^\xi)^{1-\gamma}}{1-\gamma} \\ &+ q^\xi(w_t^\xi)(\hat{w}_t^\xi)^{1-\gamma}\Big[r - q^\xi(w_t^\xi)^{-\frac{1}{\gamma}} + \hat{z}_t^{\xi\top}(\bar{D} + \mu_{Pt}^\xi - rP_t^\xi)\Big] \\ &- \frac{\gamma}{2} q^\xi(w_t^\xi)(\hat{w}_t^\xi)^{1-\gamma}\hat{z}_t^{\xi\top}(\sigma_{Pt}^\xi + \sigma)^\top(\sigma_{Pt}^\xi + \sigma)\hat{z}_t^\xi \\ &+ q^{\xi\prime}(w_t^\xi)\frac{(\hat{w}_t^\xi)^{1-\gamma}}{1-\gamma} w_t^\xi\Big[r - q^\xi(w_t^\xi)^{-\frac{1}{\gamma}} + z_t^{\xi\top}(\bar{D} + \mu_{Pt}^\xi - rP_t^\xi)\Big] \\ &+ \frac{1}{2} q^{\xi\prime\prime}(w_t^\xi)\frac{(\hat{w}_t^\xi)^{1-\gamma}}{1-\gamma}(w_t^\xi)^2 z_t^{\xi\top}(\sigma_{Pt}^\xi + \sigma)^\top(\sigma_{Pt}^\xi + \sigma) z_t^\xi \\ &+ q^{\xi\prime}(w_t^\xi)(\hat{w}_t^\xi)^{1-\gamma} w_t^\xi \hat{z}_t^{\xi\top}(\sigma_{Pt}^\xi + \sigma)^\top(\sigma_{Pt}^\xi + \sigma) z_t^\xi \\ &+ \lambda^\xi\Big[q^{-\xi}\big(w_t^\xi[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)]\big)\frac{[1 + \hat{z}_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)]^{1-\gamma}(\hat{w}_t^\xi)^{1-\gamma}}{1-\gamma} - q^\xi(w_t^\xi)\frac{(\hat{w}_t^\xi)^{1-\gamma}}{1-\gamma}\Big].\end{aligned} \tag{A.101}$$

Note that $\hat{w}_t^{1-\gamma}$ cancel out on both sides. Thus, we have verified the conjectured value function (A.90). Using $\hat{z}_t^\xi = z_t^\xi$, we get

$$\begin{aligned}
&\rho q^\xi(w_t^\xi) \\
&= q^\xi(w_t^\xi)^{-\frac{1-\gamma}{\gamma}} + \Big[q^{\xi\prime}(w_t^\xi)w_t^\xi + (1-\gamma)q^\xi(w_t^\xi)\Big]\Big[r - q^\xi(w_t^\xi)^{-\frac{1}{\gamma}} + z_t^{\xi\top}(\bar{D} + \mu_{Pt}^\xi - rP_t^\xi)\Big] \\
&+ \frac{1}{2}\Big[q^{\xi\prime\prime}(w_t^\xi)(w_t^\xi)^2 - \gamma(1-\gamma)q^\xi(w_t^\xi) + 2(1-\gamma)q^{\xi\prime}(w_t^\xi)w_t^\xi\Big] z_t^{\xi\top}(\sigma_{Pt}^\xi + \sigma)^\top(\sigma_{Pt}^\xi + \sigma)z_t^\xi \\
&+ \lambda^\xi\Big[q^{-\xi}\big(w_t^\xi[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)]\big)[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)]^{1-\gamma} - q^\xi(w_t^\xi)\Big].
\end{aligned} \tag{A.102}$$

Using (A.100) and dividing both sides by $q^\xi(w_t^\xi)$ and rearranging, we find

$$\begin{aligned}
&q^\xi(w_t^\xi)^{-\frac{1}{\gamma}} + \Big[\frac{q^{\xi\prime}(w_t^\xi)}{q^\xi(w_t^\xi)}w_t^\xi + 1 - \gamma\Big]\Big[r - q^\xi(w_t^\xi)^{-\frac{1}{\gamma}} + \Gamma^\xi(w_t^\xi)z_t^{\xi\top}(\sigma_{Pt}^\xi + \sigma)^\top(\sigma_{Pt}^\xi + \sigma)z_t^\xi\Big] \\
&- \Big[\frac{q^{\xi\prime}(w_t^\xi)}{q^\xi(w_t^\xi)}w_t^\xi + 1 - \gamma\Big]\frac{\lambda^\xi q^{-\xi}(w_t^\xi[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)])z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)}{q^\xi(w_t^\xi)[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)]^\gamma} \\
&+ \frac{1}{2}\Big[\frac{q^{\xi\prime\prime}(w_t^\xi)}{q^\xi(w_t^\xi)}(w_t^\xi)^2 - \gamma(1-\gamma) + 2(1-\gamma)\frac{q^{\xi\prime}(w_t^\xi)}{q^\xi(w_t^\xi)}w_t^\xi\Big] z_t^{\xi\top}(\sigma_{Pt}^\xi + \sigma)^\top(\sigma_{Pt}^\xi + \sigma)z_t^\xi \\
&+ \lambda^\xi\Big[\frac{q^{-\xi}(w_t^\xi[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)])}{q^\xi(w_t^\xi)}[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)]^{1-\gamma} - 1\Big] = \rho.
\end{aligned} \tag{A.103}$$

From (A.99),

$$\begin{aligned}
\Gamma^{\xi\prime}(w_t^\xi) &= -\frac{q^{\xi\prime\prime}(w_t^\xi)q^\xi(w_t^\xi) - q^{\xi\prime}(w_t^\xi)^2}{q^\xi(w_t^\xi)^2}w_t^\xi - \frac{q^{\xi\prime}(w_t^\xi)}{q^\xi(w_t^\xi)} \\
&= -\frac{q^{\xi\prime\prime}(w_t^\xi)}{q^\xi(w_t^\xi)}w_t^\xi + \frac{q^{\xi\prime}(w_t^\xi)^2}{q^\xi(w_t^\xi)^2}w_t^\xi - \frac{q^{\xi\prime}(w_t^\xi)}{q^\xi(w_t^\xi)}.
\end{aligned} \tag{A.104}$$

Thus,

$$\begin{aligned}
&\frac{q^{\xi\prime\prime}(w_t^\xi)}{q^\xi(w_t^\xi)}(w_t^\xi)^2 - \gamma(1-\gamma) + 2(1-\gamma)\frac{q^{\xi\prime}(w_t^\xi)}{q^\xi(w_t^\xi)}w_t^\xi \\
&= \Big[\frac{q^{\xi\prime\prime}(w_t^\xi)}{q^\xi(w_t^\xi)}w_t^\xi - \frac{q^{\xi\prime}(w_t^\xi)^2}{q^\xi(w_t^\xi)^2}w_t^\xi + \frac{q^{\xi\prime}(w_t^\xi)}{q^\xi(w_t^\xi)}\Big]w_t^\xi + \Big[\frac{q^{\xi\prime}(w_t^\xi)}{q^\xi(w_t^\xi)}w_t^\xi - \gamma\Big]^2 + \frac{q^{\xi\prime}(w_t^\xi)}{q^\xi(w_t^\xi)}w_t^\xi - \gamma \\
&= -\Gamma^{\xi\prime}(w_t^\xi)w_t^\xi + \Gamma^\xi(w_t^\xi)^2 - \Gamma^\xi(w_t^\xi).
\end{aligned} \tag{A.105}$$

Substituting into (A.103) and subtracting r from both sides, we get

$$\begin{aligned}
\rho - r =& q^{\xi}(w_t^{\xi})^{-\frac{1}{\gamma}} - r + \Big[\frac{q^{\xi\prime}(w_t^{\xi})}{q^{\xi}(w_t^{\xi})} w_t^{\xi} + 1 - \gamma\Big]\Big[r - q^{\xi}(w_t^{\xi})^{-\frac{1}{\gamma}}\Big] \\
&- \lambda^{\xi}\Big[\frac{q^{\xi\prime}(w_t^{\xi})}{q^{\xi}(w_t^{\xi})} w_t^{\xi} + 1 - \gamma\Big]\frac{q^{-\xi}(w_t^{\xi}[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})])z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})}{q^{\xi}(w_t^{\xi})[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})]^{\gamma}} \\
&+ \frac{1}{2}\Big[-\Gamma^{\xi\prime}(w_t^{\xi})w_t^{\xi} + \Gamma^{\xi}(w_t^{\xi})^2 - \Gamma^{\xi}(w_t^{\xi}) + 2[1 - \Gamma^{\xi}(w_t^{\xi})]\Gamma^{\xi}(w_t^{\xi})\Big] z_t^{\xi\top}(\sigma_{Pt}^{\xi} + \sigma)^{\top}(\sigma_{Pt} + \sigma)z_t^{\xi} \\
&+ \lambda^{\xi}\Big[\frac{q^{-\xi}(w_t^{\xi}[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})])}{q^{\xi}(w_t^{\xi})}[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})]^{1-\gamma} - 1\Big] \\
=& \Big[q^{\xi}(w_t^{\xi})^{-\frac{1}{\gamma}} - r\Big]\Gamma^{\xi}(w_t^{\xi}) - \frac{1}{2}\Big[\Gamma^{\xi\prime}(w_t^{\xi})w_t^{\xi} - \Gamma^{\xi}(w_t^{\xi}) + \Gamma^{\xi}(w_t^{\xi})^2\Big] z_t^{\xi\top}(\sigma_{Pt}^{\xi} + \sigma)^{\top}(\sigma_{Pt}^{\xi} + \sigma)z_t^{\xi} \\
&+ \lambda^{\xi}\Big\{\frac{[\Gamma^{\xi}(w_t^{\xi})z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi}) + 1]q^{-\xi}(w_t^{\xi}[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})])}{q^{\xi}(w_t^{\xi})[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})]^{\gamma}} - 1\Big\}.
\end{aligned} \tag{A.106}$$

Substituting the investors' aggregate consumption c_t^{ξ} into the total wealth dynamic for investors, we get

$$\begin{aligned}
dw_t^{\xi} =& w_t^{\xi}\big[r - q^{\xi}(w_t^{\xi})^{-\frac{1}{\gamma}} + z_t^{\xi\top}(\bar{D} + \mu_{Pt}^{\xi} - rP_t^{\xi})\big]dt + w_t^{\xi}z_t^{\xi\top}(\sigma_{Pt}^{\xi} + \sigma)^{\top}dB_t \\
&+ w_t^{\xi}z_t^{\xi\top}J_t^{\xi}dN_t^{\xi}.
\end{aligned} \tag{A.107}$$

The investors' total wealth w_t^{ξ} and the state of the economy ξ are the two state variables in the model. Thus, the risky assets' prices in state ξ must be functions of the investors' total wealth in the state, that is, $P_t^{\xi} = P^{\xi}(w_t^{\xi})$. Let

$$\tilde{\mu}_t^{w,\xi} \equiv w_t^{\xi}\big[r - q^{\xi}(w_t^{\xi})^{-\frac{1}{\gamma}} + z_t^{\xi\top}(\bar{D} + \mu_{Pt}^{\xi} - rP_t^{\xi})\big], \tag{A.108}$$

$$\tilde{\sigma}_t^{w,\xi} \equiv (\sigma_{Pt}^{\xi} + \sigma)z_t^{\xi}w_t^{\xi}, \tag{A.109}$$

which are the drift and the diffusion of the investors' total wealth w_t^{ξ} after substituting in the investors' optimal consumption and wealth-scaled asset holdings. By Itô's lemma, we have

$$\begin{aligned}
dP_t^{\xi} =& \Big[\tilde{\mu}_t^{w,\xi}P^{\xi\prime}(w_t^{\xi}) + \frac{1}{2}(\tilde{\sigma}_t^{w,\xi})^{\top}\tilde{\sigma}_t^{w,\xi}P^{\xi\prime\prime}(w_t^{\xi})\Big]dt + P^{\xi\prime}(w_t^{\xi})(\tilde{\sigma}_t^{w,\xi})^{\top}dB_t \\
&+ (P_t^{-\xi} - P_t^{\xi})dN_t^{\xi}.
\end{aligned} \tag{A.110}$$

From (A.108) and (A.109), the drift of the risky assets' prices become

$$
\begin{aligned}
&\tilde{\mu}_t^{w,\xi} P^{\xi\prime}(w_t^\xi) + \frac{1}{2}(\tilde{\sigma}_t^{w,\xi})^\top \tilde{\sigma}_t^{w,\xi} P^{\xi\prime\prime}(w_t^\xi) \\
&= w_t^\xi \big[r - q^\xi (w_t^\xi)^{-\frac{1}{\gamma}} + z_t^{\xi\top}(\bar{D} + \mu_{Pt}^\xi - r P_t^\xi) \big] P^{\xi\prime}(w_t^\xi) \\
&\quad + \frac{1}{2} z_t^{\xi\top} (\sigma_{Pt}^\xi + \sigma)^\top (\sigma_{Pt}^\xi + \sigma) z_t^\xi P^{\xi\prime\prime}(w_t^\xi)(w_t^\xi)^2 \\
&= w_t^\xi \Big[r - q^\xi (w_t^\xi)^{-\frac{1}{\gamma}} + \Gamma^\xi(w_t^\xi) z_t^{\xi\top} (\sigma_{Pt}^\xi + \sigma)^\top (\sigma_{Pt}^\xi + \sigma) z_t^\xi \\
&\qquad - \frac{\lambda^\xi q^{-\xi} \big(w_t^\xi [1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)] \big) z_t^{\xi\top} (P_t^{-\xi} - P_t^\xi)}{q^\xi(w_t^\xi)[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)]^\gamma} \Big] P^{\xi\prime}(w_t^\xi) \\
&\quad + \frac{1}{2} z_t^{\xi\top} (\sigma_{Pt}^\xi + \sigma)^\top (\sigma_{Pt}^\xi + \sigma) z_t^\xi P^{\xi\prime\prime}(w_t^\xi)(w_t^\xi)^2 \\
&= \Big[r - q^\xi (w_t^\xi)^{-\frac{1}{\gamma}} - \frac{\lambda^\xi q^{-\xi} \big(w_t^\xi [1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)] \big) z_t^{\xi\top} (P_t^{-\xi} - P_t^\xi)}{q^\xi(w_t^\xi)[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)]^\gamma} \Big] P^{\xi\prime}(w_t^\xi) w_t^\xi \\
&\quad + z_t^{\xi\top} (\sigma_{Pt}^\xi + \sigma)^\top (\sigma_{Pt}^\xi + \sigma) z_t^\xi \Big[\Gamma^\xi(w_t^\xi) P^{\xi\prime}(w_t^\xi) w_t^\xi + \frac{1}{2} P^{\xi\prime\prime}(w_t^\xi)(w_t^\xi)^2 \Big],
\end{aligned}
\tag{A.111}
$$

where the second equality follows from (A.100). The diffusion of the risky assets' prices are

$$
\tilde{\sigma}_t^{w,\xi} P^{\xi\prime}(w_t^\xi)^\top = (\sigma_{Pt}^\xi + \sigma) z_t^\xi w_t^\xi P^{\xi\prime}(w_t^\xi)^\top. \tag{A.112}
$$

The risky assets' prices are conjectured to follow the Itô process (6). Thus, by matching coefficients, we get

$$
J_t^\xi = P_t^{-\xi} - P_t^\xi, \tag{A.113}
$$

while μ_{Pt}^ξ and σ_{Pt}^ξ are given by

$$
\begin{aligned}
\mu_{Pt}^\xi = &\Big[r - q^\xi (w_t^\xi)^{-\frac{1}{\gamma}} - \frac{\lambda^\xi q^{-\xi} \big(w_t^\xi [1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)] \big) z_t^{\xi\top} (P_t^{-\xi} - P_t^\xi)}{q^\xi(w_t^\xi)[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)]^\gamma} \Big] P^{\xi\prime}(w_t^\xi) w_t^\xi \\
&+ z_t^{\xi\top} (\sigma_{Pt}^\xi + \sigma)^\top (\sigma_{Pt}^\xi + \sigma) z_t^\xi \Big[\Gamma^\xi(w_t^\xi) P^{\xi\prime}(w_t^\xi) w_t^\xi + \frac{1}{2} P^{\xi\prime\prime}(w_t^\xi)(w_t^\xi)^2 \Big],
\end{aligned}
\tag{A.114}
$$

and

$$
\sigma_{Pt}^\xi = (\sigma_{Pt}^\xi + \sigma) z_t^\xi w_t^\xi P^{\xi\prime}(w_t^\xi)^\top. \tag{A.115}
$$

From (A.115),

$$
(\sigma_{Pt}^\xi + \sigma) z_t^\xi = (\sigma_{Pt}^\xi + \sigma) z_t^\xi P^{\xi\prime}(w_t^\xi)^\top z_t^\xi w_t^\xi + \sigma z_t^\xi. \tag{A.116}
$$

Note that $P^{\xi\prime}(w_t^\xi)^\top z_t^\xi w_t^\xi$ is a scalar. Thus,

$$
\big[1 - P^{\xi\prime}(w_t^\xi)^\top z_t^\xi w_t^\xi \big] (\sigma_{Pt}^\xi + \sigma) z_t^\xi = \sigma z_t^\xi. \tag{A.117}
$$

Then

$$(\sigma_{Pt}^{\xi}+\sigma)z_t^{\xi}=\frac{\sigma z_t^{\xi}}{1-P^{\xi\prime}(w_t^{\xi})^{\top}z_t^{\xi}w_t^{\xi}}. \tag{A.118}$$

Substituting (A.118) into (A.115) yields

$$\sigma_{Pt}^{\xi}=\frac{\sigma z_t^{\xi}P^{\xi\prime}(w_t^{\xi})^{\top}w_t^{\xi}}{1-P^{\xi\prime}(w_t^{\xi})^{\top}z_t^{\xi}w_t^{\xi}}. \tag{A.119}$$

Adding σ to both sides yields

$$\sigma_{Pt}^{\xi}+\sigma=\frac{\sigma z_t^{\xi}P^{\xi\prime}(w_t^{\xi})^{\top}w_t^{\xi}+[1-P^{\xi\prime}(w_t^{\xi})^{\top}z_t^{\xi}w_t^{\xi}]\sigma}{1-P^{\xi\prime}(w_t^{\xi})^{\top}z_t^{\xi}w_t^{\xi}}. \tag{A.120}$$

Let

$$\hat{\sigma}_t^{\xi}\equiv\sigma z_t^{\xi}P^{\xi\prime}(w_t^{\xi})^{\top}w_t^{\xi}+[1-P^{\xi\prime}(w_t^{\xi})^{\top}z_t^{\xi}w_t^{\xi}]\sigma. \tag{A.121}$$

Then

$$\sigma_{Pt}^{\xi}+\sigma=\frac{\hat{\sigma}_t^{\xi}}{1-P^{\xi\prime}(w_t^{\xi})^{\top}z_t^{\xi}w_t^{\xi}}. \tag{A.122}$$

Substituting (A.118) into (A.114) yields

$$\begin{aligned}\mu_{Pt}^{\xi}=&\Big[r-q^{\xi}(w_t^{\xi})^{-\frac{1}{\gamma}}-\frac{\lambda^{\xi}q^{-\xi}\big(w_t^{\xi}[1+z_t^{\xi\top}(P_t^{-\xi}-P_t^{\xi})]\big)z_t^{\xi\top}(P_t^{-\xi}-P_t^{\xi})}{q^{\xi}(w_t^{\xi})[1+z_t^{\xi\top}(P_t^{-\xi}-P_t^{\xi})]^{\gamma}}\Big]P^{\xi\prime}(w_t^{\xi})w_t^{\xi}\\&+\frac{z_t^{\xi\top}\Sigma z_t^{\xi}}{[1-P^{\xi\prime}(w_t^{\xi})^{\top}z_t^{\xi}w_t^{\xi}]^2}\Big[\Gamma^{\xi}(w_t^{\xi})P^{\xi\prime}(w_t^{\xi})w_t^{\xi}+\frac{1}{2}P^{\xi\prime\prime}(w_t^{\xi})(w_t^{\xi})^2\Big].\end{aligned} \tag{A.123}$$

It is thus immediate that

$$\begin{aligned}&\bar{D}+\mu_{Pt}^{\xi}-rP_t^{\xi}\\&=\Big[r-q^{\xi}(w_t^{\xi})^{-\frac{1}{\gamma}}-\frac{\lambda^{\xi}q^{-\xi}\big(w_t^{\xi}[1+z_t^{\xi\top}(P_t^{-\xi}-P_t^{\xi})]\big)z_t^{\xi\top}(P_t^{-\xi}-P_t^{\xi})}{q^{\xi}(w_t^{\xi})[1+z_t^{\xi\top}(P_t^{-\xi}-P_t^{\xi})]^{\gamma}}\Big]P^{\xi\prime}(w_t^{\xi})w_t^{\xi}\\&\quad+\frac{z_t^{\xi\top}\Sigma z_t^{\xi}}{[1-P^{\xi\prime}(w_t^{\xi})^{\top}z_t^{\xi}w_t^{\xi}]^2}\Big[\Gamma^{\xi}(w_t^{\xi})P^{\xi\prime}(w_t^{\xi})w_t^{\xi}+\frac{1}{2}P^{\xi\prime\prime}(w_t^{\xi})(w_t^{\xi})^2\Big]+\bar{D}-rP^{\xi}(w_t^{\xi}).\end{aligned} \tag{A.124}$$

From (A.100) and using (A.118) and (A.122), we have

$$\bar{D}+\mu_{Pt}^{\xi}-rP_t^{\xi}=\frac{\Gamma^{\xi}(w_t^{\xi})\hat{\sigma}_t^{\xi\top}\sigma z_t^{\xi}}{[1-P^{\xi\prime}(w_t^{\xi})^{\top}z_t^{\xi}w_t^{\xi}]^2}-\frac{\lambda^{\xi}q^{-\xi}\big(w_t^{\xi}[1+z_t^{\xi\top}(P_t^{-\xi}-P_t^{\xi})]\big)(P_t^{-\xi}-P_t^{\xi})}{q^{\xi}(w_t^{\xi})[1+z_t^{\xi\top}(P_t^{-\xi}-P_t^{\xi})]^{\gamma}}. \tag{A.125}$$

From (A.124) and (A.125),

$$\begin{aligned}
&\Big[r - q^{\xi}(w_t^{\xi})^{-\frac{1}{\gamma}} - \frac{\lambda^{\xi} q^{-\xi}\big(w_t^{\xi}[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})]\big) z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})}{q^{\xi}(w_t^{\xi})[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})]^{\gamma}}\Big] P^{\xi\prime}(w_t^{\xi}) w_t^{\xi} \\
&\quad + \frac{z_t^{\xi\top}\Sigma z_t^{\xi}}{[1 - P^{\xi\prime}(w_t^{\xi})^{\top} z_t^{\xi} w_t^{\xi}]^2}\Big[\Gamma^{\xi}(w_t^{\xi}) P^{\xi\prime}(w_t^{\xi}) w_t^{\xi} + \frac{1}{2} P^{\xi\prime\prime}(w_t^{\xi})(w_t^{\xi})^2\Big] + \bar{D} - r P^{\xi}(w_t^{\xi}) \qquad \text{(A.126)} \\
&= \frac{\Gamma^{\xi}(w_t^{\xi}) \hat{\sigma}^{\xi\top} \sigma z_t^{\xi}}{[1 - P^{\xi\prime}(w_t^{\xi})^{\top} z_t^{\xi} w_t^{\xi}]^2} - \frac{\lambda^{\xi} q^{-\xi}\big(w_t^{\xi}[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})]\big)(P_t^{-\xi} - P_t^{\xi})}{q^{\xi}(w_t^{\xi})[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})]^{\gamma}}.
\end{aligned}$$

The equation (A.126) follows from the investors' first-order condition and the Itô's lemma. Moreover, (A.125) implies that

$$\begin{aligned}
z_t^{\xi\top}(\bar{D} + \mu_{Pt}^{\xi} - r P_t^{\xi}) = \Gamma^{\xi}(w_t^{\xi}) &\frac{z_t^{\xi\top}\Sigma z_t^{\xi}}{[1 - z_t^{\xi\top} P^{\xi\prime}(w_t^{\xi}) w_t^{\xi}]^2} \\
&- \frac{\lambda^{\xi} q^{-\xi}\big(w_t^{\xi}[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})]\big) z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})}{q^{\xi}(w_t^{\xi})[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})]^{\gamma}}.
\end{aligned} \qquad \text{(A.127)}$$

Since there is a measure one of symmetric hedgers, the aggregate risky asset holdings by hedgers at time t in state ξ are given by

$$Q_t^{\xi} = \Big[(\sigma_{Pt}^{\xi} + \sigma)^{\top}(\sigma_{Pt}^{\xi} + \sigma) + \lambda^{\xi} J_t^{\xi} J_t^{\xi\top}\Big]^{-1}\Big[\frac{\bar{D} + \mu_{Pt}^{\xi} - r P_t^{\xi} + \lambda^{\xi} J_t^{\xi}}{\alpha} - (\sigma_{Pt}^{\xi} + \sigma)^{\top}\sigma u^{\xi}\Big]. \qquad \text{(A.128)}$$

From (A.118), (A.122) and (A.124) and using the fact that $J_t^{\xi} = P_t^{-\xi} - P_t^{\xi}$,

$$\begin{aligned}
Q_t^{\xi} = \frac{1}{\alpha}\Big\{ &\frac{\hat{\sigma}_t^{\xi\top}\hat{\sigma}_t^{\xi}}{[1 - P^{\xi\prime}(w_t^{\xi})^{\top} z_t^{\xi} w_t^{\xi}]^2} + \lambda^{\xi}(P_t^{-\xi} - P_t^{\xi})(P_t^{-\xi} - P_t^{\xi})^{\top}\Big\}^{-1}\Big\{\lambda^{\xi}(P_t^{-\xi} - P_t^{\xi}) \\
&+ \Big[r - q^{\xi}(w_t^{\xi})^{-\frac{1}{\gamma}} - \frac{\lambda^{\xi} q^{-\xi}\big(w_t^{\xi}[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})]\big) z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})}{q^{\xi}(w_t^{\xi})[1 + z_t^{\xi\top}(P_t^{-\xi} - P_t^{\xi})]^{\gamma}}\Big] P^{\xi\prime}(w_t^{\xi}) w_t^{\xi} \\
&+ \frac{z_t^{\xi\top}\Sigma z_t^{\xi}}{[1 - P^{\xi\prime}(w_t^{\xi})^{\top} z_t^{\xi} w_t^{\xi}]^2}\Big[\Gamma^{\xi}(w_t^{\xi}) P^{\xi\prime}(w_t^{\xi}) w_t^{\xi} + \frac{1}{2} P^{\xi\prime\prime}(w_t^{\xi})(w_t^{\xi})^2\Big] + \bar{D} - r P^{\xi}(w_t^{\xi}) \\
&- \frac{\alpha \hat{\sigma}_t^{\xi\top}\sigma u^{\xi}}{1 - P^{\xi\prime}(w_t^{\xi})^{\top} z_t^{\xi} w_t^{\xi}}\Big\}.
\end{aligned} \qquad \text{(A.129)}$$

Market clearing requires that

$$Q_t^{\xi} + Y_t^{\xi} = Q_t^{\xi} + z_t^{\xi} w_t^{\xi} = S. \qquad \text{(A.130)}$$

Combining (A.129) and (A.130) yields

$$\begin{aligned}
&\frac{1}{\alpha}\Big\{\frac{\hat{\sigma}_t^{\xi\top}\hat{\sigma}_t^{\xi}}{[1-P^{\xi\prime}(w_t^{\xi})^{\top}z_t^{\xi}w_t^{\xi}]^2}+\lambda^{\xi}(P_t^{-\xi}-P_t^{\xi})(P_t^{-\xi}-P_t^{\xi})^{\top}\Big\}^{-1}\Big\{\lambda^{\xi}(P_t^{-\xi}-P_t^{\xi})\\
&+\Big[r-q^{\xi}(w_t^{\xi})^{-\frac{1}{\gamma}}-\frac{\lambda^{\xi}q^{-\xi}\big(w_t^{\xi}[1+z_t^{\xi\top}(P_t^{-\xi}-P_t^{\xi})]\big)z_t^{\xi\top}(P_t^{-\xi}-P_t^{\xi})}{q^{\xi}(w_t^{\xi})[1+z_t^{\xi\top}(P_t^{-\xi}-P_t^{\xi})]^{\gamma}}\Big]P^{\xi\prime}(w_t^{\xi})w_t^{\xi}\\
&+\frac{z_t^{\xi\top}\Sigma z_t^{\xi}}{[1-P^{\xi\prime}(w_t^{\xi})^{\top}z_t^{\xi}w_t^{\xi}]^2}\Big[\Gamma^{\xi}(w_t^{\xi})P^{\xi\prime}(w_t^{\xi})w_t^{\xi}+\frac{1}{2}P^{\xi\prime\prime}(w_t^{\xi})(w_t^{\xi})^2\Big]+\bar{D}-rP^{\xi}(w_t^{\xi})\\
&-\frac{\alpha\hat{\sigma}_t^{\xi\top}\sigma u^{\xi}}{1-P^{\xi\prime}(w_t^{\xi})^{\top}z_t^{\xi}w_t^{\xi}}\Big\}=S-z_t^{\xi}w_t^{\xi}.
\end{aligned}\tag{A.131}$$

The equation (A.131) follows from the market clearing condition.

When the economy switches from state ξ to state $-\xi$ at time t, the investor's wealth jumps from w_t^{ξ} to $w_t^{-\xi}$ according to the following fixed point condition

$$w_t^{-\xi}=w_t^{\xi}+w_t^{\xi}z_t^{\xi\top}(P_t^{-\xi}-P_t^{\xi}).\tag{A.132}$$

Substituting (A.132) into (A.126) yields

$$\begin{aligned}
&\bar{D}-rP^{\xi}(w_t^{\xi})+\Big[r-q^{\xi}(w_t^{\xi})^{-\frac{1}{\gamma}}-\frac{\lambda^{\xi}q^{-\xi}(w_t^{-\xi})\frac{w_t^{-\xi}-w_t^{\xi}}{w_t^{\xi}}}{q^{\xi}(w_t^{\xi})(\frac{w_t^{-\xi}}{w_t^{\xi}})^{\gamma}}\Big]P^{\xi\prime}(w_t^{\xi})w_t^{\xi}\\
&+\frac{z_t^{\xi\top}\Sigma z_t^{\xi}}{[1-P^{\xi\prime}(w_t^{\xi})^{\top}z_t^{\xi}w_t^{\xi}]^2}\Big[\Gamma^{\xi}(w_t^{\xi})P^{\xi\prime}(w_t^{\xi})w_t^{\xi}+\frac{1}{2}P^{\xi\prime\prime}(w_t^{\xi})(w_t^{\xi})^2\Big]\\
&=\frac{\Gamma^{\xi}(w_t^{\xi})\hat{\sigma}_t^{\xi\top}\sigma z_t^{\xi}}{[1-P^{\xi\prime}(w_t^{\xi})^{\top}z_t^{\xi}w_t^{\xi}]^2}-\frac{\lambda^{\xi}q^{-\xi}(w_t^{-\xi})(P_t^{-\xi}-P_t^{\xi})}{q^{\xi}(w_t^{\xi})(\frac{w_t^{-\xi}}{w_t^{\xi}})^{\gamma}}.
\end{aligned}\tag{A.133}$$

Furthermore, substituting (A.132) into (A.131) yields

$$\begin{aligned}
&\bar{D}-rP_t^{\xi}+\Big[r-q^{\xi}(w_t^{\xi})^{-\frac{1}{\gamma}}-\frac{\lambda^{\xi}q^{-\xi}(w_t^{-\xi})\frac{w_t^{-\xi}-w_t^{\xi}}{w_t^{\xi}}}{q^{\xi}(w_t^{\xi})(\frac{w_t^{-\xi}}{w_t^{\xi}})^{\gamma}}\Big]P^{\xi\prime}(w_t^{\xi})w_t^{\xi}\\
&+\frac{z_t^{\xi\top}\Sigma z_t^{\xi}}{[1-P^{\xi\prime}(w_t^{\xi})^{\top}z_t^{\xi}w_t^{\xi}]^2}\Big[\Gamma^{\xi}(w_t^{\xi})P^{\xi\prime}(w_t^{\xi})w_t^{\xi}+\frac{1}{2}P^{\xi\prime\prime}(w_t^{\xi})(w_t^{\xi})^2\Big]\\
&=\alpha\Big\{\frac{\hat{\sigma}_t^{\xi\top}\hat{\sigma}_t^{\xi}}{[1-P^{\xi\prime}(w_t^{\xi})^{\top}z_t^{\xi}w_t^{\xi}]^2}+\lambda^{\xi}(P^{-\xi}-P^{\xi})(P^{-\xi}-P^{\xi})^{\top}\Big\}(S-z_t^{\xi}w_t^{\xi})\\
&+\frac{\alpha\hat{\sigma}_t^{\xi\top}\sigma u^{\xi}}{1-P^{\xi\prime}(w_t^{\xi})^{\top}z_t^{\xi}w_t^{\xi}}-\lambda^{\xi}(P_t^{-\xi}-P_t^{\xi}).
\end{aligned}\tag{A.134}$$

Combining (A.133) and (A.134) yields

$$\begin{aligned}
&\frac{\Gamma^\xi(w_t^\xi)\hat{\sigma}_t^{\xi\top}\sigma z_t^\xi}{[1-w_t^\xi P^{\xi\prime}(w_t^\xi)^\top z_t^\xi]^2} - \frac{\lambda^\xi q^{-\xi}(w_t^{-\xi})(P_t^{-\xi}-P_t^\xi)}{q^\xi(w_t^\xi)(\frac{w_t^{-\xi}}{w_t^\xi})^\gamma} \\
&= \frac{\alpha\hat{\sigma}_t^{\xi\top}\hat{\sigma}_t^\xi(S-z_t^\xi w_t^\xi)}{[1-w_t^\xi P^{\xi\prime}(w_t^\xi)^\top z_t^\xi]^2} + \alpha\lambda^\xi(P^{-\xi}-P^\xi)(P^{-\xi}-P^\xi)^\top(S-z_t^\xi w_t^\xi) \\
&\quad + \frac{\alpha\hat{\sigma}^\top\sigma u^\xi}{1-w_t^\xi P^{\xi\prime}(w_t^\xi)^\top z_t^\xi} - \lambda^\xi(P^{-\xi}-P_t^\xi).
\end{aligned} \tag{A.135}$$

Rearranging yields

$$\begin{aligned}
&\frac{[\Gamma^\xi(w_t^\xi)+\alpha w_t^\xi]\hat{\sigma}_t^{\xi\top}\sigma z_t^\xi - \alpha\hat{\sigma}_t^{\xi\top}\sigma(\sigma^{-1}\hat{\sigma}_t^\xi)S}{[1-w_t^\xi P^{\xi\prime}(w_t^\xi)^\top z_t^\xi]^2} - \frac{\alpha\hat{\sigma}_t^{\xi\top}\sigma u^\xi}{1-w_t^\xi P^{\xi\prime}(w_t^\xi)^\top z_t^\xi} \\
&= \lambda^\xi\Big\{\Big[\frac{q^{-\xi}(w_t^{-\xi})}{q^\xi(w_t^\xi)(\frac{w_t^{-\xi}}{w_t^\xi})^\gamma}-1\Big](P_t^{-\xi}-P_t^\xi) + \alpha(P_t^{-\xi}-P_t^\xi)(P_t^{-\xi}-P_t^\xi)^\top(S-z_t^\xi w_t^\xi)\Big\}.
\end{aligned} \tag{A.136}$$

Note that

$$\hat{\sigma}_t^\xi = w_t^\xi\sigma z_t^\xi P^{\xi\prime}(w_t^\xi)^\top + [1-w_t^\xi P^{\xi\prime}(w_t^\xi)^\top z_t^\xi]\sigma. \tag{A.137}$$

Thus,

$$\hat{\sigma}_t^{\xi\top}\sigma z_t^\xi = w_t^\xi P^{\xi\prime}(w_t^\xi) z_t^{\xi\top}\Sigma z_t^\xi + [1-w_t^\xi P^{\xi\prime}(w_t^\xi)^\top z_t^\xi]\Sigma z_t^\xi. \tag{A.138}$$

Then, (A.133) simplifies to

$$\begin{aligned}
&\bar{D} - rP^\xi(w_t^\xi) + \Big[r - q^\xi(w_t^\xi)^{-\frac{1}{\gamma}} - \frac{\lambda^\xi q^{-\xi}(w_t^{-\xi})\frac{w_t^{-\xi}-w_t^\xi}{w_t^\xi}}{q^\xi(w_t^\xi)(\frac{w_t^{-\xi}}{w_t^\xi})^\gamma}\Big]P^{\xi\prime}(w_t^\xi)w_t^\xi \\
&+ \frac{z_t^{\xi\top}\Sigma z_t^\xi}{2[1-w_t^\xi P^{\xi\prime}(w_t^\xi)^\top z_t^\xi]^2}P^{\xi\prime\prime}(w_t^\xi)(w_t^\xi)^2 = \frac{\Gamma^\xi(w_t^\xi)\Sigma z_t^\xi}{1-w_t^\xi P^{\xi\prime}(w_t^\xi)^\top z_t^\xi} - \frac{\lambda^\xi q^{-\xi}(w_t^{-\xi})(P_t^{-\xi}-P_t^\xi)}{q^\xi(w_t^\xi)(\frac{w_t^{-\xi}}{w_t^\xi})^\gamma}.
\end{aligned} \tag{A.139}$$

Moreover, $q^\xi(w_t^\xi)$ is solution to the ODE (A.106). Using (A.118), we can rewrite (A.106) as

$$\begin{aligned}
&\frac{q^\xi(w_t^\xi)^{-\frac{1}{\gamma}}-r}{\rho-r}\Gamma^\xi(w_t^\xi) - \frac{\Gamma^{\xi\prime}(w_t^\xi)w_t^\xi - \Gamma^\xi(w_t^\xi)+\Gamma^\xi(w_t^\xi)^2}{2(\rho-r)}\frac{z_t^{\xi\top}\Sigma z_t^\xi}{[1-w_t^\xi P^{\xi\prime}(w_t^\xi)^\top z_t^\xi]^2} \\
&+ \frac{\lambda^\xi}{\rho-r}\Big\{\frac{[\Gamma^\xi(w_t^\xi)\frac{w_t^{-\xi}-w_t^\xi}{w_t^\xi}+1]q^{-\xi}(w_t^{-\xi})}{q^\xi(w_t^\xi)(\frac{w_t^{-\xi}}{w_t^\xi})^\gamma}-1\Big\} = 1.
\end{aligned} \tag{A.140}$$

The investors' wealth-scaled asset holdings z_t^ξ are determined by (A.136), while prices $P^\xi(w_t^\xi)$ and hedging demand $q^\xi(w_t^\xi)$ are solutions to the ODEs (A.139) and (A.140). We can define

$$\pi^\xi(w_t^\xi) \equiv \frac{\bar{D}}{r} - P^\xi(w_t^\xi). \tag{A.141}$$

Then, $P^{\xi\prime}(w_t^\xi) = -\pi^{\xi\prime}(w_t^\xi)$ and $P^{\xi\prime\prime}(w_t^\xi) = -\pi^{\xi\prime\prime}(w_t^\xi)$. The ODE (A.139) becomes

$$\begin{aligned} &r\pi^\xi(w_t^\xi) - \Big[r - q^\xi(w_t^\xi)^{-\frac{1}{\gamma}} - \frac{\lambda^\xi q^{-\xi}(w_t^{-\xi})\frac{w_t^{-\xi}-w_t^\xi}{w_t^\xi}}{q^\xi(w_t^\xi)(\frac{w_t^{-\xi}}{w_t^\xi})^\gamma}\Big]\pi^{\xi\prime}(w_t^\xi)w_t^\xi - \frac{z_t^{\xi\top}\Sigma z_t^\xi \pi^{\xi\prime\prime}(w_t^\xi)(w_t^\xi)^2}{2[1+w_t^\xi\pi^{\xi\prime}(w_t^\xi)^\top z_t^\xi]^2} \\ &= \frac{\Gamma^\xi(w_t^\xi)\Sigma z_t^\xi}{1+w_t^\xi\pi^{\xi\prime}(w_t^\xi)^\top z_t^\xi} - \frac{\lambda^\xi q^{-\xi}(w_t^{-\xi})[\pi^\xi(w_t^\xi)-\pi^{-\xi}(w_t^{-\xi})]}{q^\xi(w_t^\xi)(\frac{w_t^{-\xi}}{w_t^\xi})^\gamma}. \end{aligned} \tag{A.142}$$

The ODE (A.140) becomes

$$\begin{aligned} &\frac{q^\xi(w_t^\xi)^{-\frac{1}{\gamma}} - r}{\rho - r}\Gamma^\xi(w_t^\xi) - \frac{\Gamma^{\xi\prime}(w_t^\xi)w_t^\xi - \Gamma^\xi(w_t^\xi) + \Gamma^\xi(w_t^\xi)^2}{2(\rho-r)}\frac{z_t^{\xi\top}\Sigma z_t^\xi}{[1+w_t^\xi\pi^{\xi\prime}(w_t^\xi)^\top z_t^\xi]^2} \\ &+ \frac{\lambda^\xi}{\rho-r}\Big\{\frac{[\Gamma^\xi(w_t^\xi)\frac{w_t^{-\xi}-w_t^\xi}{w_t^\xi}+1]q^{-\xi}(w_t^{-\xi})}{q^\xi(w_t^\xi)(\frac{w_t^{-\xi}}{w_t^\xi})^\gamma} - 1\Big\} = 1. \end{aligned} \tag{A.143}$$

Moreover, (A.136) can be written as

$$\begin{aligned} &\frac{[\Gamma^\xi(w_t^\xi)+\alpha w_t^\xi]\hat{\sigma}_t^{\xi\top}\sigma z_t^\xi - \alpha\hat{\sigma}_t^{\xi\top}\sigma(\sigma^{-1}\hat{\sigma}_t^\xi)S}{[1+w_t^\xi\pi^{\xi\prime}(w_t^\xi)^\top z_t^\xi]^2} - \frac{\alpha\hat{\sigma}_t^{\xi\top}\sigma u^\xi}{1+w_t^\xi\pi^{\xi\prime}(w_t^\xi)^\top z_t^\xi} \\ &= \lambda^\xi\Big\{\Big[\frac{q^{-\xi}(w_t^{-\xi})}{q^\xi(w_t^\xi)(\frac{w_t^{-\xi}}{w_t^\xi})^\gamma} - 1\Big](P_t^{-\xi}-P_t^\xi) + \alpha(P_t^{-\xi}-P_t^\xi)(P_t^{-\xi}-P_t^\xi)^\top(S - z_t^\xi w_t^\xi)\Big\}. \end{aligned} \tag{A.144}$$

Here,

$$\hat{\sigma}_t^\xi = -w_t^\xi\sigma z_t^\xi\pi^{\xi\prime}(w_t^\xi)^\top + [1+w_t^\xi\pi^{\xi\prime}(w_t^\xi)^\top z_t^\xi]\sigma, \tag{A.145}$$

$$\Gamma^\xi(w_t^\xi) = \gamma - \frac{q^{\xi\prime}(w_t^\xi)}{q^\xi(w_t^\xi)}w_t^\xi. \tag{A.146}$$

Moreover, the following fixed point condition must be satisfied

$$w_t^{-\xi} - w_t^\xi = w_t^\xi y_t^{\xi\top}\big[\pi^\xi(w_t^\xi) - \pi^{-\xi}(w_t^{-\xi})\big]. \tag{A.147}$$

A.2 Proof of Proposition 2

Based on the derivations for Proposition 1 and Proposition 4 and specializing the ODEs to the one-asset case, we find that investors' aggregate wealth-scaled asset holding z_t^ξ satisfies

$$\begin{aligned}&\frac{(1+\alpha w_t^\xi)\sigma^2 z_t^\xi - \alpha\sigma^2 S}{[1+\pi^{\xi\prime}(w_t^\xi)w_t^\xi z_t^\xi]^2} - \frac{\alpha\sigma^2 u^\xi}{1+\pi^{\xi\prime}(w_t^\xi)w_t^\xi z_t^\xi}\\ &= \lambda^\xi\Big[\Big(\frac{w_t^\xi}{w_t^{-\xi}}-1\Big)\frac{w_t^{-\xi}-w_t^\xi}{w_t^\xi z_t^\xi} + \alpha(S-w_t^\xi z_t^\xi)\Big(\frac{w_t^{-\xi}-w_t^\xi}{w_t^\xi z_t^\xi}\Big)^2\Big].\end{aligned} \tag{A.148}$$

The functions $\pi^\xi(w_t^\xi)$ and $\bar{q}^\xi(w_t^\xi)$ are solutions to the following ODEs

$$\begin{aligned}&r\pi^\xi(w_t^\xi) - \Big[r-\rho-\lambda^\xi\frac{w_t^{-\xi}-w_t^\xi}{w_t^{-\xi}}\Big]\pi^{\xi\prime}(w_t^\xi)w_t^\xi - \frac{1}{2}\Big[\frac{\sigma z_t^\xi}{1+\pi^{\xi\prime}(w_t^\xi)w_t^\xi z_t^\xi}\Big]^2\pi^{\xi\prime\prime}(w_t^\xi)(w_t^\xi)^2\\ &= \frac{\sigma^2 z_t^\xi}{1+\pi^{\xi\prime}(w_t^\xi)w_t^\xi z_t^\xi} - \lambda^\xi\frac{\pi^\xi(w_t^\xi)-\pi^{-\xi}(w_t^{-\xi})}{\frac{w_t^{-\xi}}{w_t^\xi}},\end{aligned} \tag{A.149}$$

and

$$\begin{aligned}&\log\rho + \Big[\bar{q}^{\xi\prime}(w_t^\xi)w_t^\xi + \frac{1}{\rho}\Big](r-\rho) - \rho\bar{q}^\xi(w_t^\xi)\\ &+ \Big[\bar{q}^{\xi\prime}(w_t^\xi)w_t^\xi + \frac{1}{2}\bar{q}^{\xi\prime\prime}(w_t^\xi)(w_t^\xi)^2 + \frac{1}{2\rho}\Big]\Big[\frac{\sigma z_t^\xi}{1+\pi^{\xi\prime}(w_t^\xi)w_t^\xi z_t^\xi}\Big]^2\\ &+ \lambda^\xi\Big[\frac{1}{\rho}\log\Big(\frac{w_t^{-\xi}}{w_t^\xi}\Big) + \bar{q}^{-\xi}(w_t^{-\xi}) - \bar{q}^\xi(w_t^\xi) - \Big(\bar{q}^{\xi\prime}(w_t^\xi)w_t^\xi + \frac{1}{\rho}\Big)\frac{w_t^{-\xi}-w_t^\xi}{w_t^{-\xi}}\Big] = 0.\end{aligned} \tag{A.150}$$

Moreover, the following fixed point condition must be satisfied

$$w_t^{-\xi} - w_t^\xi = w_t^\xi z_t^\xi[\pi^\xi(w_t^\xi) - \pi^{-\xi}(w_t^{-\xi})]. \tag{A.151}$$

Suppose that we are standing at $\xi = s$ and the state prior to shock is denoted by $n-$. We want to analyze the asymptotic behavior at $\lambda^s = \lambda \to \infty$. Let $\epsilon \equiv 1/\lambda$. Expand the function $\pi^s(w)$ according to

$$\pi^s(w) = \pi^n(w) + \epsilon\pi_1(w) + \epsilon^2\pi_2(w) + \dots \tag{A.152}$$

where higher-order terms have been omitted. If $w_t^s = w$, then from the fixed point condition (A.151), we have

$$\begin{aligned}w_t^{n-} &= w - (S-Q_t^{n-})\big[\pi^n(w_t^{n-}) - \pi^s(w)\big]\\ &= w - (S-Q_t^{n-})\big[\pi^n(w_t^{n-}) - \pi^n(w) - \epsilon\pi_1(w) - \epsilon^2\pi_2(w)\big].\end{aligned} \tag{A.153}$$

For small ϵ, we have

$$w_t^{n-} = w - (S-Q_t^{n-})\big[\pi^{n\prime}(w)(w_t^{n-}-w) - \epsilon\pi_1(w) - \epsilon^2\pi_2(w)\big], \tag{A.154}$$

which implies

$$w_t^{n-} = w + \frac{S - Q_t^{n-}}{1 + (S - Q_t^{n-})\pi^{n\prime}(w)}\big[\pi_1(w)\epsilon + \pi_2(w)\epsilon^2\big]. \tag{A.155}$$

When $\lambda \to \infty$, $w_t^{n-} \to w$. Moreover, $\pi^n(w_t^{n-}) \to \pi^s(w)$.

From (A.58) and (A.59) and specializing to the one-asset case, we get

$$\sigma_{Pt}^{\xi} + \sigma = \frac{\sigma}{1 + \pi^{\xi\prime}(w_t^{\xi})w_t^{\xi} z_t^{\xi}} = \frac{\sigma}{1 + \pi^{\xi\prime}(w_t^{\xi})(S - Q_t^{\xi})}. \tag{A.156}$$

The second equality follows from the market clearing condition $w_t z_t^{\xi} = S - Q_t^{\xi}$. Then

$$\sigma_{Rt}^{\xi} \equiv \frac{\sigma_{Pt}^{\xi} + \sigma}{P_t^{\xi}} = \frac{\sigma}{[1 + \pi^{\xi\prime}(w_t^{\xi})(S - Q_t^{\xi})][\frac{\bar{D}}{r} - \pi^{\xi}(w_t^{\xi})]}. \tag{A.157}$$

Note that in the one-asset case,

$$Q_t^{\xi} = \Big[(\sigma_{Pt}^{\xi} + \sigma)^2 + \lambda^{\xi}(J_t^{\xi})^2\Big]^{-1}\Big[\frac{\bar{D} + \mu_{Pt}^{\xi} - rP_t^{\xi} + \lambda^{\xi} J_t^{\xi}}{\alpha} - (\sigma_{Pt}^{\xi} + \sigma)\sigma u^{\xi}\Big]. \tag{A.158}$$

In the limit with $\lambda \to \infty$, (A.37) implies

$$(\sigma_{Pt}^{s} + \sigma)^2 \hat{z}_t^{s} = \bar{D} + \mu_{Pt}^{s} - rP_t^{s} + \lambda J_t^{s}, \tag{A.159}$$

where λJ_t^s is bounded away from zero, and (A.158) becomes

$$Q_t^s = \frac{1}{\alpha \hat{w}_t^s}(S - Q_t^s) - \frac{\sigma}{\sigma_{Pt}^s + \sigma}u^{\xi} = \frac{1}{\alpha \hat{w}_t^s}(S - Q_t^s) - [1 + \pi^{s\prime}(w_t^s)(S - Q_t^s)]u^s. \tag{A.160}$$

Rearranging yields

$$Q_t^s = \frac{\frac{S}{\alpha \hat{w}_t^s} - [1 + \pi^{s\prime}(w_t^s)S]u^s}{1 + \frac{1}{\alpha \hat{w}_t^s} - \pi^{s\prime}(w_t^s)u^s}. \tag{A.161}$$

Similarly, in the normal state,

$$Q_t^n = \frac{\frac{S}{\alpha \hat{w}_t^n} - [1 + \pi^{n\prime}(w_t^n)S]u^n}{1 + \frac{1}{\alpha \hat{w}_t^n} - \pi^{n\prime}(w_t^n)u^n}. \tag{A.162}$$

Since $\pi^{n\prime}(w) < 0$, we can show that $Q_t^s < Q_t^n$ for $u^s > u^n$. Substituting into (A.157), we see that $\sigma_{Rt}^s > \sigma_{Rt}^n$. Note that $\Delta\mathbb{E}_t(R) = \mu_{Rt}^s + \lambda J_{Rt}^s - \mu_{Rt}^n$, that is

$$\Delta\mathbb{E}_t(R) = \frac{(\sigma_{Pt}^s + \sigma)^2(S - Q_t^s)}{\hat{w}_t^s(\frac{\bar{D}}{r} - \pi^s(w_t^s))} - \frac{(\sigma_{Pt}^n + \sigma)^2(S - Q_t^n)}{\hat{w}_t^n(\frac{\bar{D}}{r} - \pi^n(w_t^n))}. \tag{A.163}$$

In the limit, $\Delta\mathbb{E}_t(R) > 0$. From (22), $\mathcal{O}^{\log}(w_t^n; u^n, u^s) < \mathcal{C}^{\log}(w_t^n; \delta)|_{\delta = \Delta\mathbb{E}_t(R)}$.

(The proof for

$$\lim_{u^n \to u^s} \left(\mathcal{O}^{log}(w_t^n; u^n, u^s) - \mathcal{C}^{log}(w_t^n; \delta)|_{\delta = \Delta \mathbb{E}_t(R)} \right) < 0.$$

is TBC)

A.3 Proof of Proposition 3

In this proof we show that

$$\lim_{\Delta \upsilon \to 0} \frac{1}{\gamma} \frac{\left[\left((\Omega_t^s)^{-1} - (\Omega_t^n)^{-1} \right) \left(\mathbb{E}_t \left(R^s \right) - r\mathbf{1} \right) \right]}{\Delta E \left(R \right)} =$$

$$= -\frac{P_t^2}{\gamma} \left(I \left(s + \upsilon \right) F \left(w_t \right) + \frac{r}{\bar{D}} 2\pi \left(w_t \right) \left(1 - \sigma^2 I \left(s + \upsilon \right)^2 F \left(w_t \right) \right) \right) \left(s + \upsilon \right) < 0$$

where $\frac{1}{\sigma^2 I (s+\upsilon)^2} > F \left(w_t \right) > 0$ and $\bar{F} \left(w_t \right) > 0$ are scalar functions defined by ODEs and P_t, w_t and $\pi \left(\cdot \right)$ are the limit of P_t^ξ, w_t^ξ and $\pi^\xi \left(\cdot \right)$ for both states $\xi = s, n$.

We start with the following Lemma which is based on results in Kondor-Vayanos.

Lemma A.1. *When $\lambda^s, \lambda^n \to 0$, price is*

$$P_t^\xi = \frac{\bar{D}}{r} - \pi^\xi \left(\omega_t \right) \Sigma \left(s + u^\xi \right)$$

expected return has the form of

$$\mathbb{E}_t(R^\xi) - r\mathbf{1} = diag \left(P_t^\xi \right)^{-1} \left(\bar{D} + \mu_{Pt}^\xi - r P_t^\xi \right) = diag \left(P_t^\xi \right)^{-1} B^\xi \left(w_t^\xi \right) \Sigma \left(s + u^\xi \right)$$

investors equilibrium position has the form of

$$Y_t = G^\xi \left(w_t^\xi \right) \left(s + u^\xi \right)$$

and, the inverse of the return covariance matrix has the form of

$$\left(\Omega_t^\xi \right)^{-1} = diag \left(P_t^\xi \right) \Sigma^{-1} diag \left(P_t^\xi \right) - F^\xi \left(w_t^\xi \right) diag \left(P_t^\xi \right) \left(s + u^\xi \right) \left(s + u^\xi \right)^\intercal diag \left(P_t^\xi \right)$$

where $\pi^\xi \left(\cdot \right), F^\xi \left(\cdot \right)$ and $B^\xi \left(\cdot \right)$ and $G^\xi \left(\cdot \right)$ are scalar functions depending on the vectors of u^ξ and s only through the terms $\left(s + u^\xi \right)^\intercal \Sigma \left(s + u^\xi \right)$ and $\left(s + u^\xi \right)^\intercal \Sigma s$.

Proof. With $\lambda^s, \lambda^n \to 0$, our model is equivalent with the positive supply, short-term hedgers' version of KV discussed in Section 6 . Expected return is given by (6.1) in Proposition 6.2 in KV which is, with our notation,

$$\bar{D} + \mu_{Pt}^\xi - r P_t^\xi = B^\xi \left(w_t^\xi \right) \Sigma \left(s + u^\xi \right)$$

where

$$B^{\xi}\left(w_t^{\xi}\right)=\frac{\alpha\left(\frac{1}{w_t^{\xi}}\right)}{\alpha+\frac{1}{w_t^{\xi}}+\alpha\pi^{\xi\prime}\left(w_t^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma s}\left[1-\left(s+u^{\xi}\right)^{\intercal}\Sigma\left(s+u^{\xi}\right)\frac{\alpha\pi^{\xi\prime}\left(w_t\right)}{\alpha+\frac{1}{w_t^{\xi}}+\alpha\pi^{\xi\prime}\left(w_t^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma s}\right]$$

where $\pi^{\xi}\left(w_t^{\xi}\right)$ solves

$$\begin{aligned}&\frac{\alpha^2\left(s+u^{\xi}\right)^{\intercal}\Sigma\left(s+u^{\xi}\right)}{2\left[\alpha+\frac{1}{w_t^{\xi}}+\alpha\pi^{\xi\prime}(w_t)\left(s+u^{\xi}\right)^{\intercal}\Sigma s\right]^2}\left(-\pi^{\prime\prime\xi}\left(w_t^{\xi}\right)\right)-\left(r-q(w_t)^{-\frac{1}{\gamma}}\right)\pi^{\xi\prime}(w_t)w_t+r\pi^{\xi}\left(w_t^{\xi}\right)\\&=\frac{\alpha\left(\frac{\gamma}{w_t^{\xi}}-\frac{q^{\xi\prime}(w_t)}{q^{\xi}(w_t)}\right)}{\alpha+\frac{1}{w_t^{\xi}}+\alpha\pi^{\xi\prime}(w_t)\left(s+u^{\xi}\right)^{\intercal}\Sigma s},\end{aligned}\tag{A.164}$$

with boundary conditions

$$\pi^{\xi}\left(0\right)=\frac{\alpha}{r},\ \lim_{w_t\to\infty}\pi^{\xi}\left(w_t\right)=0.$$

Theorem 3 in Kondor and Vayanos (2019) implies that $\pi^{\xi}\left(w_t\right)$ is monotonically decreasing. For Π_t^{ξ}, as the proof of Proposition 6.1 shows, we have the following expressions (using our notation).

$$\sigma_{P_t}^{\xi}=\frac{\alpha}{\alpha+\frac{1}{w_t^{\xi}}}\pi^{\xi\prime}\left(w_t^{\xi}\right)b_t\left(s+u^{\xi}\right)^{\intercal}\Sigma$$

where

$$b_t=\frac{\sigma\left(s+u^{\xi}\right)}{1-\frac{\alpha}{\alpha+\frac{1}{w_t^{\xi}}}\pi^{\xi\prime}\left(w_t^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma s}$$

Hence, we can write

$$\sigma_{P_t}^{\xi}=\frac{\alpha}{\alpha+\frac{1}{w_t^{\xi}}}\pi^{\xi\prime}\left(w_t^{\xi}\right)\frac{1}{1-\frac{\alpha}{\alpha+\frac{1}{w_t^{\xi}}}\pi^{\xi\prime}\left(w_t^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma s}\sigma\left(s+u^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma$$

or

$$\sigma_{P_t}^{\xi}=\hat{F}^{\xi}\left(w_t^{\xi}\right)\sigma\left(s+u^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma$$

with

$$\hat{F}^{\xi}\left(w_t^{\xi}\right)\equiv\frac{\alpha\pi^{\xi\prime}\left(w_t^{\xi}\right)}{\alpha+\frac{1}{w_t^{\xi}}-\alpha\pi^{\xi\prime}\left(w_t^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma s}.$$

Then

$$\sigma_{P_t}^{\xi\intercal} = \hat{F}^{\xi}\left(w_t^{\xi}\right)\left(\sigma\left(s+u^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma\right)^{\intercal} = \hat{F}^{\xi}\left(w_t^{\xi}\right)\left(\left(s+u^{\xi}\right)^{\intercal}\Sigma\right)^{\intercal}\left(\sigma\left(s+u^{\xi}\right)\right)^{\intercal} =$$
$$= \hat{F}^{\xi}\left(w_t^{\xi}\right)\Sigma\left(s+u^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\sigma^{\intercal}$$

implying

$$\sigma_{P_t}^{\xi\intercal}\sigma_{P_t}^{\xi} = \left(s+u^{\xi}\right)^{\intercal}\Sigma\left(s+u^{\xi}\right)\left(\hat{F}^{\xi}\left(w_t^{\xi}\right)\right)^2\Sigma\left(s+u^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma.$$

Then

$$\Pi_t^{\xi} = \left(\sigma_{P_t}^{\xi}+\sigma\right)^{\intercal}\left(\sigma_{P_t}^{\xi}+\sigma\right) = \sigma_{P_t}^{\xi\intercal}\sigma_{P_t}^{\xi} + \sigma^{\intercal}\sigma_{P_t}^{\xi} + \Sigma + \sigma_{P_t}^{\xi\intercal}\sigma =$$
$$= \left(\hat{F}^{\xi}\left(w_t^{\xi}\right)\right)^2\Sigma\left(s+u^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma\left(s+u^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma + 2\hat{F}^{\xi}\left(w_t^{\xi}\right)\Sigma\left(s+u^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma + \Sigma$$
$$= \left(s+u^{\xi}\right)^{\intercal}\Sigma\left(s+u^{\xi}\right)\left(\hat{F}^{\xi}\left(w_t^{\xi}\right)\right)^2\Sigma\left(s+u^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma + 2\hat{F}^{\xi}\left(w_t^{\xi}\right)\Sigma\left(s+u^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma + \Sigma$$
$$= \hat{F}^{\xi}\left(w_t^{\xi}\right)\left(\hat{F}^{\xi}\left(w_t^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma\left(s+u^{\xi}\right)+2\right)\Sigma\left(s+u^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma + \Sigma$$

Or, defining

$$\bar{F}^{\xi}\left(w_t^{\xi}\right) = \hat{F}^{\xi}\left(w_t^{\xi}\right)\left(\hat{F}^{\xi}\left(w_t^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma\left(s+u^{\xi}\right)+2\right) =$$
$$= \frac{-\alpha\pi^{\xi\prime}\left(w_t^{\xi}\right)}{\frac{1}{w_t^{\xi}}+\alpha\left(1+\pi^{\xi\prime}\left(w_t^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma s\right)}\left(\frac{-\alpha\pi^{\xi\prime}\left(w_t^{\xi}\right)}{\frac{1}{w_t^{\xi}}+\alpha\left(1+\pi^{\xi\prime}\left(w_t^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma s\right)}\left(s+u^{\xi}\right)^{\intercal}\Sigma\left(s+u^{\xi}\right)+2\right)$$

we have simply

$$\Pi_t^{\xi} = \bar{F}^{\xi}\left(w_t^{\xi}\right)\Sigma\left(s+u^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma + \Sigma$$

where $\bar{F}^{\xi}\left(w_t^{\xi}\right)$ is a scalar. For the inverse, note that by the Sherman Morrison formula

$$\left(\Sigma + vv^{\intercal}\right)^{-1} = \Sigma^{-1} - \frac{\Sigma^{-1}vv^{\intercal}\Sigma^{-1}}{1+v^{\intercal}\Sigma^{-1}v}.$$

Picking

$$v = \sqrt{\bar{F}^{\xi}\left(w_t\right)}\Sigma\left(s+u^{\xi}\right)$$

gives $\left(\Pi_t^{\xi}\right)^{-1} = \left(\Sigma + vv^{\intercal}\right)^{-1}$ and

$$\left(\Pi_t^{\xi}\right)^{-1} = \Sigma^{-1} - \frac{\Sigma^{-1}vv^{\intercal}\Sigma^{-1}}{1+v^{\intercal}\Sigma^{-1}v}$$
$$= \Sigma^{-1} - \frac{\bar{F}^{\xi}\left(w_t^{\xi}\right)}{1+\bar{F}^{\xi}\left(w_t^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma\left(s+u^{\xi}\right)}\left(s+u^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}.$$

Introducing the short-cut

$$F^{\xi}\left(w_t^{\xi}\right) \equiv \frac{\bar{F}^{\xi}\left(w_t^{\xi}\right)}{1+\bar{F}^{\xi}\left(w_t^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma\left(s+u^{\xi}\right)}$$

and the definition of Ω_t^{ξ} givs that the inverse of the return covariance matrix has the form of

$$\left(\Omega_t^{\xi}\right)^{-1} = diag\left(P_t^{\xi}\right)\Sigma^{-1}diag\left(P_t^{\xi}\right) - F^{\xi}\left(w_t^{\xi}\right)diag\left(P_t^{\xi}\right)\left(s+u^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}diag\left(P_t^{\xi}\right)$$

Finally, Equation (IA148) gives equilibirum positions with the choice of $F'\left(w_t\right)=0$ as

$$Y_t = \frac{\alpha}{\alpha+\frac{1}{w_t^{\xi}}+\alpha\pi'\left(w_t^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma s}\left(\sigma_{P_t}+\sigma\right)^{-1}\sigma\left(s+u\right)$$

then

$$\left(\hat{F}^{\xi}\left(w_t^{\xi}\right)\Sigma\left(s+u^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma+\Sigma\right)Y_t = \frac{\alpha}{\alpha+\frac{1}{w_t^{\xi}}+\alpha\pi'\left(w_t^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma s}\Sigma\left(s+u\right)$$

or

$$Y_t = \frac{\alpha}{\alpha+\frac{1}{w_t^{\xi}}+\alpha\pi'\left(w_t^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma s}\left(\hat{F}^{\xi}\left(w_t^{\xi}\right)\Sigma\left(s+u^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma+\Sigma\right)^{-1}\Sigma\left(s+u\right)$$

which gives the statement with the choice of

$$G^{\xi}\left(w_t^{\xi}\right) = \frac{\alpha}{\alpha+\frac{1}{w_t^{\xi}}+\alpha\pi'\left(w_t^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma s}\frac{1}{1+\hat{F}^{\xi}\left(w_t^{\xi}\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma\left(s+u^{\xi}\right)}.$$

□

First note, that in our example, $\left(s+u^s\right)^{\intercal}\Sigma\left(s+u^s\right) = \left(s+u^n\right)^{\intercal}\Sigma\left(s+u^n\right)$ and $\left(s+u^s\right)^{\intercal}\Sigma s = \left(s+u^n\right)^{\intercal}\Sigma s$ hence the Lemma implies that $\pi^{\xi}\left(\cdot\right), F^{\xi}\left(\cdot\right)$ and $B^{\xi}\left(\cdot\right)$ and $G^{\xi}\left(\cdot\right)$ are the same functions in the states $\xi = n, s$. Therefore, in the rest of the proof we omit the subscript ξ.

We are interested in

$$\left[\left(\left(\Omega_t^s\right)^{-1}-\left(\Omega_t^n\right)^{-1}\right)\mathbb{E}\left(R_t^s\right)\right]_1$$

which we can write as a weighted combination of the terms

$$\left[\left(\left(\Omega_t^s\right)^{-1}-\left(\Omega_t^n\right)^{-1}\right)\right]_{1j}.$$

So consider first

$$\left[\left(\left(\Omega_t^s\right)^{-1}-\left(\Omega_t^n\right)^{-1}\right)\right]_{11}=$$
$$=\left(\frac{\bar{D}}{r}-\pi\left(w_t^s\right)\sigma^2\left(s+\upsilon+\Delta\upsilon\right)\right)^2\left(\frac{1}{\sigma^2}-F\left(w_t^s\right)\left(s+\upsilon+\Delta\upsilon\right)^2\right)$$
$$-\left(\frac{\bar{D}}{r}-\pi\left(w_t^s\right)\sigma^2\left(s+\upsilon\right)\right)^2\left(\frac{1}{\sigma^2}-F\left(w_t^s\right)\left(s+\upsilon\right)^2\right)$$

Hence, to a first order we have

$$\left[\left(\left(\Omega_t^s\right)^{-1}-\left(\Omega_t^n\right)^{-1}\right)\right]_{11}\approx\frac{\partial\frac{\left(\frac{\bar{D}}{r}-\pi\left(w_t\right)\sigma^2\left(s+\upsilon\right)\right)^2}{\sigma^2}-F\left(w_t\right)\left(\frac{\bar{D}}{r}-\pi\left(w_t\right)\sigma^2\left(s+\upsilon\right)\right)^2\left(s+\upsilon\right)^2}{\partial w_t}\left(w_t^s-w_t^n\right)+$$
$$+\frac{\partial\frac{\left(\frac{\bar{D}}{r}-\pi\left(w_t\right)\sigma^2\left(s+\upsilon\right)\right)^2}{\sigma^2}-F\left(w_t\right)\left(\frac{\bar{D}}{r}-\pi\left(w_t\right)\sigma^2\left(s+\upsilon\right)\right)^2\left(s+\upsilon\right)^2}{\partial\upsilon}\Delta\upsilon$$

First, we show that the first element is of order $\left(\Delta\upsilon\right)^2$. Indeed

$$w_t^s-w_t^n=\hat{Y}_t^{n\intercal}\left(\left(-\pi\left(w_t^s\right)\right)\left(s+u^s\right)-\left(-\pi\left(w_t^n\right)\right)\left(s+u^n\right)\right)=$$
$$=G\left(w_t^n\right)\left(-\pi\left(w_t^s\right)\right)\left(s+u^n\right)^\intercal\Sigma\left(s+u^s\right)-G\left(w_t^n\right)\left(-\pi\left(w_t^n\right)\right)\left(s+u^n\right)^\intercal\Sigma\left(s+u^n\right)=$$
$$=G\left(w_t^n\right)\left(-\pi\left(w_t^s\right)\right)\sigma^2\left(\Delta\upsilon\right)^2+G\left(w_t^n\right)\left(\left(-\pi\left(w_t^s\right)\right)-\left(-\pi\left(w_t^n\right)\right)\right)\left(s+u^n\right)^\intercal\Sigma\left(s+u^n\right)$$

using

$$\left(\left(-\pi\left(w_t^s\right)\right)-\left(-\pi\left(w_t^n\right)\right)\right)\approx-\pi'\left(w_t^n\right)\left(w_t^s-w_t^n\right).$$

this implies

$$w_t^s-w_t^n\approx\frac{G\left(w_t^n\right)\left(-\pi\left(w_t^s\right)\right)}{1+G\left(w_t^n\right)\left(s+u^n\right)^\intercal\Sigma\left(s+u^n\right)\pi'\left(w_t^n\right)}\left(\left(s+u^n\right)^\intercal\Sigma\left(s+u^s\right)-\left(s+u^n\right)^\intercal\Sigma\left(s+u^n\right)\right)=$$
$$=-\frac{G\left(w_t^n\right)\left(-\pi\left(w_t^s\right)\right)}{1+G\left(w_t^n\right)\left(s+u^n\right)^\intercal\Sigma\left(s+u^n\right)\pi'\left(w_t^n\right)}\sigma^2\left(\Delta\upsilon\right)^2.$$

Therefore, we can replace w^n, w^s with its limit w_t. Hence,

$$\left[\left(\left(\Omega_t^s\right)^{-1}-\left(\Omega_t^n\right)^{-1}\right)\right]_{11}\approx\frac{\partial\frac{\left(\frac{\bar{D}}{r}-\pi\left(w_t\right)\sigma^2\left(s+\upsilon\right)\right)^2}{\sigma^2}-F\left(w_t\right)\left(\frac{\bar{D}}{r}-\pi\left(w_t\right)\sigma^2\left(s+\upsilon\right)\right)^2\left(s+\upsilon\right)^2}{\partial\upsilon}\Delta\upsilon=$$
$$=-2P_t\left(F\left(w_t\right)\left(s+\upsilon\right)P_t-\pi\left(w_t\right)\left(F\left(w_t\right)\left(s+\upsilon\right)\sigma^2\left(s+\upsilon\right)-1\right)\right)\Delta\upsilon$$

Note that

$$F\left(w_t\right)\left(s+\upsilon\right)\sigma^2\left(s+\upsilon\right)=\frac{\bar{F}\left(w_t\right)\left(s+\upsilon\right)\sigma^2\left(s+\upsilon\right)}{1+\bar{F}\left(w_t\right)\sigma^2\left(s+u^\xi\right)^\intercal\left(s+u^\xi\right)}<1$$

and that $\pi\left(w_t\right), P_t>0$ hence $\left[\left(\left(\Omega_t^s\right)^{-1}-\left(\Omega_t^n\right)^{-1}\right)\right]_{11}$ is always negative.

Now consider

$$\left[\left(\left(\Omega_t^s\right)^{-1}-\left(\Omega_t^n\right)^{-1}\right)\right]_{1j}$$

First we consider the case when $j \leq I_1$, then

$$\begin{aligned}
&\left[\left(\left(\Omega_t^s\right)^{-1}-\left(\Omega_t^n\right)^{-1}\right)\right]_{1j} = \\
&= \left(-F\left(w_t\right)\left(\frac{\bar{D}}{r}-\pi\left(w_t\right)\sigma^2\left(s+\upsilon\right)\right)\left(\frac{\bar{D}}{r}-\pi\left(w_t\right)\sigma^2\left(s+\upsilon+\Delta\upsilon\right)\right)\left(s+\upsilon\right)\left(s+\upsilon+\Delta\upsilon\right)\right) \\
&- \left(-F\left(w_t\right)\left(\frac{\bar{D}}{r}-\pi\left(w_t\right)\sigma^2\left(s+\upsilon\right)\right)\left(\frac{\bar{D}}{r}-\pi\left(w_t\right)\sigma^2\left(s+\upsilon\right)\right)\left(s+\upsilon\right)\left(s+\upsilon\right)\right)
\end{aligned}$$

Just as before, the effect of w_t is lower order, so we can write

$$\begin{aligned}
\left[\left(\left(\Omega_t^s\right)^{-1}-\left(\Omega_t^n\right)^{-1}\right)\right]_{1j} &\approx \frac{\partial\left(-F(w_t)\left(\frac{\bar{D}}{r}-\pi(w_t)\sigma^2(s+\upsilon)\right)\left(\frac{\bar{D}}{r}-\pi(w_t)\sigma^2(s+\upsilon+\Delta\upsilon)\right)(s+\upsilon)(s+\upsilon+\Delta\upsilon)\right)}{\partial\Delta\upsilon}|_{\Delta\upsilon=0}\Delta\upsilon = \\
&= -F\left(w_t\right)\left[P_t^s\right]_1\left(s+\upsilon\right)\left(\left[P_t^s\right]_1-\pi\left(w_t\right)\sigma^2\left(s+\upsilon\right)\right)\Delta\upsilon
\end{aligned}$$

Finally, we have the $I > j > I_1$assets, for those we can write

$$\begin{aligned}
\left[\left(\left(\Omega_t^s\right)^{-1}-\left(\Omega_t^n\right)^{-1}\right)\right]_{1k} &\approx -\frac{\partial\left(-F(w_t)\left(\frac{\bar{D}}{r}-\pi(w_t)\sigma^2(s+\upsilon)\right)\left(\frac{\bar{D}}{r}-\pi(w_t)\sigma^2(s+\upsilon+\Delta\upsilon)\right)(s+\upsilon)(s+\upsilon+\Delta\upsilon)\right)}{\partial\Delta\upsilon}|_{\Delta\upsilon=0}\left(-\Delta\upsilon\right) \\
&= -F\left(w_t\right)\left[P_t^s\right]_1\left(s+\upsilon\right)\left(\left[P_t^s\right]_1-\pi\left(w_t\right)\sigma^2\left(s+\upsilon\right)\right)\Delta\upsilon.
\end{aligned}$$

(It is easy to see that $\left[\left(\left(\Omega_t^s\right)^{-1}-\left(\Omega_t^n\right)^{-1}\right)\right]_{1I}=0$, as the role of the first and the last asset before and after the shock are symmetric.)

These terms are the same at the limit with uncertain sign. Then, using $\mathbb{E}_t\left(R_t^s\right)-r\mathbf{1} = \operatorname{diag}\left(P_t^s\right)^{-1}B\left(w_t^s\right)\Sigma\left(s+u^s\right)$ gives

$$\begin{aligned}
&\left[\left(\left(\Omega_t^s\right)^{-1}-\left(\Omega_t^n\right)^{-1}\right)\right]_1\left(\mathbb{E}_t\left(R_t^s\right)-r\mathbf{1}\right) = \\
&= -2\left(F\left(w_t\right)\left(s+\upsilon\right)P_t-\pi\left(w_t\right)\left(F\left(w_t\right)\left(s+\upsilon\right)\sigma^2\left(s+\upsilon\right)-1\right)\right)B\left(w_t\right)\sigma^2\left(s+\upsilon\right)\Delta\upsilon- \\
&\qquad (I-2)F\left(w_t\right)\left(s+\upsilon\right)\left(\left[P_t^s\right]_1-\pi\left(w_t\right)\sigma^2\left(s+\upsilon\right)\right)B\left(w_t\right)\sigma^2\left(s+\upsilon\right)\Delta\upsilon = \\
&\qquad = \left(-I\left(s+\upsilon\right)F\left(w_t\right)\frac{\bar{D}}{r}-2\pi\left(w_t\right)\left(1-\sigma^2 I\left(s+\upsilon\right)^2F\left(w_t\right)\right)\right)B\left(w_t^s\right)\sigma^2\left(s+\upsilon\right)\Delta\upsilon
\end{aligned}$$

which we scale by $\left[\Delta\mathbb{E}_t\left(R\right)\right]_1 = B\left(w_t\right)\frac{\bar{D}}{r}\frac{\sigma^2}{P_t^2}\Delta\upsilon$ giving

$$-\frac{\left(I\left(s+\upsilon\right)\frac{\bar{D}}{r}F\left(w_t\right)+2\pi\left(w_t\right)\left(1-\sigma^2 I\left(s+\upsilon\right)^2F\left(w_t\right)\right)\right)}{\frac{\bar{D}}{r}\frac{1}{P_t^2}}\left(s+\upsilon\right)$$

which is negative as

$$1-\sigma^2 I\left(s+\upsilon\right)^2 F\left(w_t\right)=1-\sigma^2 I\left(s+\upsilon\right)^2 \frac{\bar{F}\left(w_t\right)}{1+\bar{F}\left(w_t\right)\left(s+u^{\xi}\right)^{\intercal}\Sigma\left(s+u^{\xi}\right)}=$$
$$=1-\frac{\sigma^2 I\left(s+\upsilon\right)^2 \bar{F}\left(w_t\right)}{1+\bar{F}\left(w_t\right)\sigma^2 I\left(s+\upsilon\right)^2}>0.$$

Finally, we turn to the results on the covariances. Note from Lemma A.1 that

$$[\Omega_t^s]_{1i}=\bar{F}^s\left(w_t^s\right)\sigma^2\left\{\begin{array}{ll}\frac{(s+\upsilon+\Delta\upsilon)^2}{\left(\frac{\bar{D}}{r}+g(\omega_t)\sigma^2(s+\upsilon+\Delta\upsilon)\right)^2} & for \quad i=1\\ \frac{(s+\upsilon+\Delta\upsilon)(s+\upsilon)}{\left(\frac{\bar{D}}{r}+g(\omega_t)\sigma^2(s+\upsilon+\Delta\upsilon)\right)\left(\frac{\bar{D}}{r}+g(\omega_t)\sigma^2(s+\upsilon)\right)} & for \quad I_1>i>1\\ \frac{(s+\upsilon+\Delta\upsilon)^2}{\left(\frac{\bar{D}}{r}+g(\omega_t)\sigma^2(s+\upsilon+\Delta\upsilon)\right)^2} & for \quad I>i>I_1\\ \frac{(s+\upsilon+\Delta\upsilon)(s+\upsilon)}{\left(\frac{\bar{D}}{r}+g(\omega_t)\sigma^2(s+\upsilon+\Delta\upsilon)\right)\left(\frac{\bar{D}}{r}+g(\omega_t)\sigma^2(s+\upsilon)\right)} & for \quad i=I\end{array}\right\}$$

while

$$[\Omega_t^n]_{1i}=\bar{F}^n\left(w_t^n\right)\sigma^2\left\{\begin{array}{ll}\frac{(s+\upsilon)^2}{\left(\frac{\bar{D}}{r}+g(\omega_t)\sigma^2(s+\upsilon)\right)^2} & for \quad i=1\\ \frac{(s+\upsilon)(s+\upsilon)}{\left(\frac{\bar{D}}{r}+g(\omega_t)\sigma^2(s+\upsilon)\right)\left(\frac{\bar{D}}{r}+g(\omega_t)\sigma^2(s+\upsilon)\right)} & for \quad I_1>i>1\\ \frac{(s+\upsilon+\Delta\upsilon)(s+\upsilon)}{\left(\frac{\bar{D}}{r}+g(\omega_t)\sigma^2(s+\upsilon+\Delta\upsilon)\right)\left(\frac{\bar{D}}{r}+g(\omega_t)\sigma^2(s+\upsilon)\right)} & for \quad I>i>I_1\\ \frac{(s+\upsilon+\Delta\upsilon)(s+\upsilon)}{\left(\frac{\bar{D}}{r}+g(\omega_t)\sigma^2(s+\upsilon+\Delta\upsilon)\right)\left(\frac{\bar{D}}{r}+g(\omega_t)\sigma^2(s+\upsilon)\right)} & for \quad i=I\end{array}\right\}$$

Therefore, following the argument above, using the first order Taylor expension around $\Delta\upsilon=0$ and recognizing that $w_t^n-w_t^s$ is of order $\left(\Delta\upsilon\right)^2$, we have

$$\lim_{\Delta\upsilon\to 0}\frac{[\Omega_t^s-\Omega_t^n]_{1i}}{\Delta\upsilon}=\bar{F}\left(w_t\right)\sigma^2\left\{\begin{array}{lll}2\frac{\bar{D}}{r}\frac{(s+\upsilon)^2}{(P_t^n)^3} & for & i=1\\ \frac{\bar{D}}{r}\frac{s+\upsilon}{(P_t^n)^3} & for & I>i>1\\ 0 & for & i=I\end{array}\right\}$$

A.4 Proof of Proposition 5

Suppose that an investor is offered to buy any asset with a δ additional instantaneous expected return. The hypothetical dividend process is

$$dD_t^h=\bar{D}dt+\sigma^{\top}dB_t+(1-\chi)\,\mathrm{diag}(P_t^n)\delta dt+\chi\,\mathrm{diag}(P_t^n)\frac{\delta}{\lambda}dN_t^h. \tag{A.165}$$

Let $\epsilon\equiv\mathrm{diag}(P_t^n)\delta$. Then

$$dD_t^h=\bar{D}dt+\sigma^{\top}dB_t+(1-\chi)\epsilon dt+\chi\frac{\epsilon}{\lambda}dN_t^h. \tag{A.166}$$

Given the investor's hypothetical dividend process, let $\hat{z}_t^h \equiv \hat{Y}_t^h/\hat{w}_t^h$ denote the individual investor's wealth-scaled asset holdings and let $\hat{c}_t^h$ denote the investor's consumption. The individual investor's wealth dynamic is

$$\begin{aligned} d\hat{w}_t^h &= r\hat{w}_t^h dt + \hat{w}_t^h \hat{z}_t^{h\top}(dD_t^h + dP_t^n - rP_t^n dt) - \hat{c}_t^h dt \\ &= \Big[r\hat{w}_t^h - \hat{c}_t^h + \hat{w}_t^h \hat{z}_t^{h\top}\big(\bar{D} + \mu_{Pt}^n - rP_t^n + (1-\chi)\epsilon\big)\Big]dt + \hat{w}_t^h \hat{z}_t^{h\top}(\sigma_{Pt}^n + \sigma)^\top dB_t \\ &\quad + \frac{\chi}{\lambda}\hat{w}_t^h \hat{z}_t^{h\top}\epsilon dN_t^h. \end{aligned} \tag{A.167}$$

The corresponding aggregate wealth of investors

$$dw_t^n = \Big[rw_t^n - c_t^n + w_t^n z_t^{n\top}(\bar{D} + \mu_{Pt}^n - rP_t^n)\Big]dt + w_t^n z_t^{n\top}(\sigma_{Pt}^n + \sigma)^\top dB_t. \tag{A.168}$$

A.4.1 Log Investors

The individual investor's HJB equation is

$$\begin{aligned} \rho V^h = \max_{\hat{c}_t^h, \hat{z}_t^h} \Big\{ &\frac{(\hat{c}_t^h)^{1-\gamma}}{1-\gamma} + V_{\hat{w}^h}^h \mu_t^{\hat{w}^h} + \frac{1}{2}V_{\hat{w}^h\hat{w}^h}^h (\sigma_t^{\hat{w}^h})^\top \sigma_t^{\hat{w}^h} + V_w^h \mu_t^{w,n} \\ &+ \frac{1}{2}V_{ww}^h (\sigma_t^{w,n})^\top \sigma_t^{w,n} + V_{\hat{w}^h w}^h (\sigma_t^{\hat{w}^h})^\top \sigma_t^{w,n} + \lambda(V^{n,+} - V^h)\Big\}, \end{aligned} \tag{A.169}$$

where

$$\mu_t^{\hat{w}^h} = r\hat{w}_t^h - \hat{c}_t^h + \hat{w}_t^h \hat{z}_t^{h\top}\big[\bar{D} + \mu_{Pt}^n - rP_t^n + (1-\chi)\epsilon\big], \tag{A.170}$$

$$\sigma_t^{\hat{w}^h} = (\sigma_{Pt}^n + \sigma)\hat{z}_t^h \hat{w}_t^h, \tag{A.171}$$

$$\mu_t^{w,n} = rw_t^n - c_t^n + w_t^n z_t^{n\top}(\bar{D} + \mu_{Pt}^n - rP_t^n), \tag{A.172}$$

$$\sigma_t^{w,n} = (\sigma_{Pt}^n + \sigma)z_t^n w_t^n, \tag{A.173}$$

and

$$V^{n,+} = V^n\left(\hat{w}_t^h + \frac{\chi}{\lambda}\hat{w}_t^h \hat{z}_t^{h\top}\epsilon, w_t^n\right) = \frac{1}{\rho}\log\left(\hat{w}_t^h + \frac{\chi}{\lambda}\hat{w}_t^h \hat{z}_t^{h\top}\epsilon\right) + \bar{q}^n(w_t^n). \tag{A.174}$$

Here, $\bar{q}^n(w_t^n)$ is solution to the ODEs given by Proposition 4.

We conjecture (and verify) that the investor's value function under the hypothetical scenario takes the form

$$V^h(\hat{w}_t^h, w_t^n) = \frac{1}{\rho}\log\hat{w}_t^h + \bar{q}^h(w_t^n). \tag{A.175}$$

Substituting into the HJB equation yields

$$\begin{aligned}
&\log \hat{w}_t^h + \rho \bar{q}^h(w_t^n) = \\
&\max_{\hat{c}_t^h, \hat{z}_t^h} \Big\{ \log \hat{c}_t^h + \frac{1}{\rho \hat{w}_t^h} \mu_t^{\hat{w}^h} - \frac{1}{2\rho(\hat{w}_t^h)^2} (\sigma_t^{\hat{w}^h})^\top \sigma_t^{\hat{w}^h} + \bar{q}^{h\prime}(w_t^n) \mu_t^{w,n} + \frac{1}{2} \bar{q}^{h\prime\prime}(w_t^n) (\sigma_t^{w,n})^\top \sigma_t^{w,n} \\
&\qquad + \lambda \Big[\frac{1}{\rho} \log \Big(\hat{w}_t^h + \frac{\chi}{\lambda} \hat{w}_t^h \hat{z}_t^{h\top} \epsilon \Big) + \bar{q}^n(w_t^n) - \frac{1}{\rho} \log \hat{w}_t^h - \bar{q}^h(w_t^n) \Big] \Big\}.
\end{aligned} \tag{A.176}$$

Taking the first-order condition with respect to consumption yields the optimal consumption by the investor under the hypothetical scenario

$$\hat{c}_t^h = \rho \hat{w}_t^h. \tag{A.177}$$

Taking the first-order condition with respect to wealth-scaled asset holdings,

$$0 = \bar{D} + \mu_{Pt}^n - r P_t^n + (1-\chi)\epsilon - (\sigma_{Pt}^n + \sigma)^\top (\sigma_{Pt}^n + \sigma) \hat{z}_t^h + \frac{\chi \epsilon}{1 + \frac{\chi}{\lambda} \hat{z}_t^{h\top} \epsilon}. \tag{A.178}$$

Hence,

$$\hat{z}_t^h = (\Pi_t^n)^{-1} \big[\bar{D} + \mu_{Pt}^n - r P_t^n + (1-\chi)\epsilon \big] + \frac{\chi}{1 + \frac{\chi}{\lambda} \hat{z}^{h\top} \epsilon} (\Pi_t^n)^{-1} \epsilon. \tag{A.179}$$

From (A.177), the investors' aggregate consumption is given by $c_t^n = \rho w_t^n$. Then we can rewrite the HJB equation as

$$\begin{aligned}
\rho \bar{q}^h(w_t^n) &= \log \rho + \frac{1}{\rho} \Big[r - \rho + \hat{z}_t^{h\top} \left(\bar{D} + \mu_{Pt}^n - r P_t^n + (1-\chi)\epsilon \right) \Big] \\
&- \frac{1}{2\rho} z_t^{h\top} (\sigma_{Pt}^n + \sigma)^\top (\sigma_{Pt}^n + \sigma) \hat{z}_t^h + \bar{q}^{h\prime}(w_t^n) w_t^n \Big[r - \rho + z_t^{n\top} (\bar{D} + \mu_{Pt}^n - r P_t^n) \Big] \\
&+ \frac{1}{2} \bar{q}^{h\prime\prime}(w_t^n) (w_t^n)^2 z_t^{n\top} (\sigma_{Pt}^n + \sigma)^\top (\sigma_{Pt}^n + \sigma) z_t^n \\
&+ \lambda \Big[\frac{1}{\rho} \log \Big(1 + \frac{\chi}{\lambda} z_t^{h\top} \epsilon \Big) + \bar{q}^n(w_t^n) - \bar{q}^h(w_t^n) \Big].
\end{aligned} \tag{A.180}$$

Note that $\hat{w}_t^h$ have all canceled out.

Since $\hat{y}_t^h = \mathrm{diag}(P_t^n) \hat{z}_t^h$, then

$$\begin{aligned}
\hat{y}_t^h &= (\Omega_t^n)^{-1} \left[\mu_{Rt}^n - r\mathbf{1} + (1-\chi)\delta \right] + \frac{\lambda}{\lambda + \hat{y}^{h\top} \chi \delta} (\Omega_t^n)^{-1} \chi \delta \\
&= (\Omega_t^n)^{-1} (\mu_{Rt}^n - r\mathbf{1} + \delta) + \Big[\frac{\lambda}{\lambda + \hat{y}_t^{h\top} \chi \delta} - 1 \Big] (\Omega_t^n)^{-1} \chi \delta.
\end{aligned} \tag{A.181}$$

From (A.38), the investor's demand in the normal state is given by

$$\hat{y}_t^n = (\Omega_t^n)^{-1} (\mu_{Rt}^n - r\mathbf{1}). \tag{A.182}$$

The dynamic slope is thus

$$\begin{aligned}\mathcal{C}(w_t^n;\lambda,\delta,\chi) &= \frac{[\hat{y}_t^h - \hat{y}_t^n]_i}{[\delta]_i} \\ &= \frac{[(\Omega_t^n)^{-1}\delta]_i}{[\delta]_i} + \Big[\frac{\lambda}{\lambda + \hat{y}_t^{h\top}\chi\delta} - 1\Big]\frac{[(\Omega_t^n)^{-1}\chi\delta]_i}{[\delta]_i} \\ &= [(\Omega_t^n)^{-1}]_{ii} + \frac{\sum_{j\neq i}[(\Omega_t^n)^{-1}]_{ij}[\delta]_j}{[\delta]_i} + \Big[\frac{\lambda}{\lambda + \hat{y}_t^{h\top}\chi\delta} - 1\Big]\frac{[(\Omega_t^n)^{-1}\chi\delta]_i}{[\delta]_i}.\end{aligned} \tag{A.183}$$

When $\lambda \to \infty$, the last component goes to zero. When $\lambda \to 0$ and $\chi = 0$, the last component is also zero in the limit.

A.4.2 CRRA Investors

The individual investor's HJB equation is

$$\begin{aligned}\rho V^h = \max_{\hat{c}_t^h, \hat{z}_t^h} \Big\{ &\frac{(\hat{c}_t^h)^{1-\gamma}}{1-\gamma} + V^h_{\hat{w}^h}\mu_t^{\hat{w}^h} + \frac{1}{2}V^h_{\hat{w}^h\hat{w}^h}(\sigma_t^{\hat{w}^h})^\top\sigma_t^{\hat{w}^h} + V^h_w\mu_t^{w,n} \\ &+ \frac{1}{2}V^h_{ww}(\sigma_t^{w,n})^\top\sigma_t^{w,n} + V^h_{\hat{w}^h w}(\sigma_t^{\hat{w}^h})^\top\sigma_t^{w,n} + \lambda(V^{n,+} - V^h)\Big\},\end{aligned} \tag{A.184}$$

where

$$\mu_t^{\hat{w}^h} = r\hat{w}_t^h - \hat{c}_t^h + \hat{w}_t^h\hat{z}_t^{h\top}\big[\bar{D} + \mu_{Pt}^n - rP_t^n + (1-\chi)\epsilon\big], \tag{A.185}$$

$$\sigma_t^{\hat{w}^h} = (\sigma_{Pt}^n + \sigma)\hat{z}_t^h\hat{w}_t^h, \tag{A.186}$$

$$\mu_t^{w,n} = rw_t^n - c_t^n + w_t^n z_t^{n\top}(\bar{D} + \mu_{Pt}^n - rP_t^n), \tag{A.187}$$

$$\sigma_t^{w,n} = (\sigma_{Pt}^n + \sigma)z_t^n w_t^n, \tag{A.188}$$

and

$$V^{n,+} = V^n\left(\hat{w}_t^h + \frac{\chi}{\lambda}\hat{w}_t^h\hat{z}_t^{h\top}\epsilon, w_t^n\right) = q^n(w_t^n)\frac{(\hat{w}_t^h + \frac{\chi}{\lambda}\hat{w}_t^h\hat{z}_t^{h\top}\epsilon)^{1-\gamma}}{1-\gamma}. \tag{A.189}$$

Here, $q^n(w_t^n)$ is solution to the ODEs given by Proposition 4.

We conjecture (and verify) that the investor's value function takes the form

$$V^h(\hat{w}_t^h, w_t^n) = q^h(w_t^n;\lambda,\delta,\chi)\frac{(\hat{w}_t^h)^{1-\gamma}}{1-\gamma}. \tag{A.190}$$

Substituting into the HJB equation

$$\begin{aligned}
&\rho q^h(w_t^n;\lambda,\delta,\chi)\frac{(\hat{w}_t^h)^{1-\gamma}}{1-\gamma}\\
&=\max_{\hat{c}_t^h,\hat{z}_t^h}\Big\{\frac{(\hat{c}_t^h)^{1-\gamma}}{1-\gamma}+q^h(w_t^n;\lambda,\delta,\chi,)(\hat{w}_t^h)^{-\gamma}\mu_t^{\hat{w}^h}-\frac{\gamma}{2}q^h(w_t^n;\lambda,\delta,\chi)(\hat{w}_t^h)^{-\gamma-1}(\sigma_t^{\hat{w}^h})^\top\sigma_t^{\hat{w}^h}\\
&\qquad+q^{h\prime}(w_t^n;\lambda,\delta,\chi,)\frac{(\hat{w}_t^h)^{1-\gamma}}{1-\gamma}\mu_t^{w,n}+\frac{1}{2}q^{h\prime\prime}(w_t^n;\lambda,\delta\chi)\frac{(\hat{w}_t^h)^{1-\gamma}}{1-\gamma}(\sigma_t^{w,n})^\top\sigma_t^{w,n}\\
&\qquad+q^{h\prime}(w_t^n;\lambda,\delta,\chi)(\hat{w}_t^h)^{-\gamma}(\sigma_t^{\hat{w}^h})^\top\sigma_t^{w,n}\\
&\qquad+\lambda\Big[q^n(w_t^n)\frac{(\hat{w}_t^h+\frac{\chi}{\lambda}\hat{w}_t^h\hat{z}_t^{h\top}\epsilon)^{1-\gamma}}{1-\gamma}-q^h(w_t^n;\lambda,\delta,\chi)\frac{(\hat{w}_t^h)^{1-\gamma}}{1-\gamma}\Big]\Big\}.
\end{aligned}\tag{A.191}$$

Taking the first-order condition with respect to consumption yields the optimal consumption by the investor as

$$\hat{c}_t^h=q^h(w_t^n;\lambda,\delta,\chi)^{-\frac{1}{\gamma}}\hat{w}_t^h \tag{A.192}$$

Taking the first-order condition with respect to wealth-scaled asset holdings,

$$\begin{aligned}
0=&q^h(w_t^n;\lambda,\delta,\chi)\big[\bar{D}+\mu_{Pt}^n-rP_t^n+(1-\chi)\epsilon\big]\\
&-\gamma q^h(w_t^n;\lambda,\delta,\chi)(\sigma_{Pt}^n+\sigma)^\top(\sigma_{Pt}^n+\sigma)\hat{z}_t^h\\
&+q^{h\prime}(w_t^n;\lambda,\delta,\chi)(\sigma_{Pt}^n+\sigma)^\top(\sigma_{Pt}^n+\sigma)Y_t^n+\chi q^n(w_t^n)\Big(1+\frac{\chi}{\lambda}\hat{z}_t^{h\top}\epsilon\Big)^{-\gamma}\epsilon.
\end{aligned}\tag{A.193}$$

Then

$$\hat{z}_t^h=\frac{1}{\gamma}(\Pi_t^n)^{-1}\big[\bar{D}+\mu_{Pt}^n-rP_t^n+(1-\chi)\epsilon\big]+\frac{1}{\gamma}\frac{q^{h\prime}(w_t^n;\lambda,\delta,\chi)}{q^h(w_t^n;\lambda,\delta,\chi)}Y_t^n \tag{A.194}$$

$$+\frac{\chi}{\gamma}\frac{q^n(w_t^n)}{q^h(w_t^n;\lambda,\delta,\chi)}(\Pi_t^n)^{-1}\Big(1+\frac{\chi}{\lambda}\hat{z}_t^{h\top}\epsilon\Big)^{-\gamma}\epsilon. \tag{A.195}$$

Moreover, investors' aggregate consumption is given by

$$c_t^n=q^n(w_t^n)^{-\frac{1}{\gamma}}w_t^n. \tag{A.196}$$

Then we can rewrite the HJB equation as

$$\begin{aligned}
&\rho q^h(w_t^n;\lambda,\delta,\chi)\frac{(\hat{w}_t^h)^{1-\gamma}}{1-\gamma} = q^h(w,t^n;\lambda,\delta\chi)^{1-\frac{1}{\gamma}}\frac{(\hat{w}_t^h)^{1-\gamma}}{1-\gamma}\\
&+ q^h(w_t^n;\lambda,\delta,\chi)(\hat{w}_t^h)^{1-\gamma}\Big[r - q^h(w_t^n;\lambda,\delta,\chi)^{-\frac{1}{\gamma}} + \hat{z}_t^{h\top}\big(\bar{D}+\mu_{Pt}^n - rP_t^n + (1-\chi)\epsilon\big)\Big]\\
&- \frac{\gamma}{2}q^h(w_t^n;\lambda,\delta,\chi)(\hat{w}_t^h)^{1-\gamma}\hat{z}_t^{h\top}(\sigma_{Pt}^n+\sigma)^\top(\sigma_{Pt}^n+\sigma)\hat{z}_t^h\\
&+ w_t^n q^{h\prime}(w_t^n;\lambda,\delta,\chi)\frac{(\hat{w}_t^h)^{1-\gamma}}{1-\gamma}\Big[r - q^n(w_t^n)^{-\frac{1}{\gamma}} + z_t^{n\top}(\bar{D}+\mu_{Pt}^n - rP_t^n)\Big]\\
&+ \frac{1}{2}(w_t^n)^2 q^{h\prime\prime}(w_t^n;\lambda,\delta,\chi)\frac{(\hat{w}_t^h)^{1-\gamma}}{1-\gamma}z_t^{n\top}(\sigma_{Pt}^n+\sigma)^\top(\sigma_{Pt}^n+\sigma)z_t^n\\
&+ w_t^n q^{h\prime}(w_t^n;\lambda,\delta,\chi)(\hat{w}_t^h)^{1-\gamma}\hat{z}_t^{h\top}(\sigma_{Pt}^n+\sigma)^\top(\sigma_{Pt}^n+\sigma)z_t^n\\
&+ \lambda\Big[q^n(w_t^n)\frac{(\hat{w}_t^h)^{1-\gamma}(1+\frac{\chi}{\lambda}\hat{z}_t^{h\top}\epsilon)^{1-\gamma}}{1-\gamma} - q^h(w_t^n;\lambda,\delta,\chi)\frac{(\hat{w}_t^h)^{1-\gamma}}{1-\gamma}\Big].
\end{aligned} \tag{A.197}$$

Dividing both sides by $q^h(w_t^n;\lambda,\delta,\chi)\frac{(\hat{w}_t^h)^{1-\gamma}}{1-\gamma}$ yields

$$\begin{aligned}
&q^h(w_t^n;\lambda,\delta,\chi)^{-\frac{1}{\gamma}} + (1-\gamma)\Big[r - q^h(w_t^n;\lambda,\delta,\chi)^{-\frac{1}{\gamma}} + \hat{z}_t^{h\top}\big(\bar{D}+\mu_{Pt}^n - rP_t^n + (1-\chi)\epsilon\big)\Big]\\
&+ w_t^n\frac{q^{h\prime}(w_t^n;\lambda,\delta,\chi)}{q^h(w_t^n;\lambda,\delta,\chi)}\Big[r - q^n(w_t^n)^{-\frac{1}{\gamma}} + z_t^{n\top}(\bar{D}+\mu_{Pt}^n - rP_t^n)\Big]\\
&- \frac{\gamma(1-\gamma)}{2}\hat{z}^{h\top}(\sigma_{Pt}^n+\sigma)^\top(\sigma_{Pt}^n+\sigma)\hat{z}_t^h\\
&+ \frac{1}{2}(w_t^n)^2\frac{q^{h\prime\prime}(w_t^n;\lambda,\delta,\chi)}{q^h(w_t^n;\lambda,\delta,\chi)}z_t^{n\top}(\sigma_{Pt}^n+\sigma)^\top(\sigma_{Pt}^n+\sigma)z_t^n\\
&+ (1-\gamma)w_t^n\frac{q^{h\prime}(w_t^n;\lambda,\delta,\chi)}{q^h(w_t^n;\lambda,\delta,\chi)}\hat{z}^{h\top}(\sigma_{Pt}^n+\sigma)^\top(\sigma_{Pt}^n+\sigma)z_t^n\\
&+ \lambda\Big[\frac{q^n(w_t^n)}{q^h(w_t^n;\lambda,\delta,\chi)}\Big(1+\frac{\chi}{\lambda}\hat{z}_t^{h\top}\epsilon\Big)^{1-\gamma} - 1\Big] = \rho
\end{aligned} \tag{A.198}$$

Rearranging yields

$$\begin{aligned}
&q^h(w_t^n;\lambda,\delta,\chi)^{-\frac{1}{\gamma}}\\
&+ (1-\gamma)\Big[r - q^h(w_t^n;\lambda,\delta,\chi)^{-\frac{1}{\gamma}} + \hat{z}_t^{h\top}\big(\bar{D}+\mu_{Pt}^n - rP_t^n + (1-\chi)\epsilon\big)\Big]\\
&+ w_t^n\frac{q^{h\prime}(w_t^n;\lambda,\delta,\chi)}{q^h(w_t^n;\lambda,\delta,\chi)}\Big[r - q^n(w_t^n)^{-\frac{1}{\gamma}} + z_t^{n\top}(\bar{D}+\mu_{Pt}^n - rP_t^n) + (1-\gamma)\hat{z}_t^{h\top}\Pi_t^n z_t^n\Big]\\
&- \frac{\gamma(1-\gamma)}{2}\hat{z}_t^{h\top}\Pi_t^n\hat{z}_t^h + \frac{1}{2}(w_t^n)^2\frac{q^{h\prime\prime}(w_t^n;\lambda,\delta,\chi)}{q^h(w_t^n;\lambda,\delta,\chi)}z_t^{n\top}\Pi_t^n z_t^n\\
&+ \lambda\Big[\frac{q^n(w_t^n)}{q^h(w_t^n;\lambda,\delta,\chi)}\Big(1+\frac{\chi}{\lambda}\hat{z}_t^{h\top}\epsilon\Big)^{1-\gamma} - 1\Big] = \rho.
\end{aligned} \tag{A.199}$$

Since

$$\bar{D} + \mu_{Pt}^n - rP_t^n = \frac{\Gamma^n(w_t^n)\hat{\sigma}_t^{n\top}\sigma z_t^n}{[1+\pi^{n\prime}(w_t^n)^\top z_t^n w_t^n]^2}, \tag{A.200}$$

$$z_t^{n\top}(\bar{D} + \mu_{Pt}^n - rP_t^n) = \frac{\Gamma^n(w_t^n) z_t^{n\top}\Sigma z_t^n}{[1+\pi^{n\prime}(w_t^n)^\top z_t^n w_t^n]^2}, \tag{A.201}$$

$$\sigma_{Pt}^n + \sigma = \frac{\hat{\sigma}_t^n}{1+\pi^{n\prime}(w_t^n)^\top z_t^n w_t^n}, \tag{A.202}$$

where

$$\hat{\sigma}_t^n \equiv -w_t^n \sigma z_t^n \pi^{n\prime}(w_t^n)^\top + [1 + w_t^n \pi^{n\prime}(w_t^n)^\top z_t^n]\sigma. \tag{A.203}$$

Hence, the ODE becomes

$$\begin{aligned}
&q^h(w_t^n;\lambda,\delta,\chi)^{-\frac{1}{\gamma}} + (1-\gamma)\Big[r - q^h(w_t^n;\lambda,\delta,\chi)^{-\frac{1}{\gamma}} + \hat{z}^{h\top}\Big(\frac{\Gamma^n(w_t^n)\hat{\sigma}_t^{n\top}\sigma z_t^n}{[1+\pi^{n\prime}(w_t^n)^\top z_t^n w_t^n]^2} + (1-\chi)\epsilon\Big)\Big] \\
&+ w_t^n \frac{q^{h\prime}(w_t^n;\lambda,\delta,\chi)}{q^h(w_t^n;\lambda,\delta,\chi)}\Big[r - q^n(w_t^n)^{-\frac{1}{\gamma}} + \frac{\Gamma^n(w_t^n) z_t^{n\top}\Sigma z_t^n}{[1+\pi^{n\prime}(w_t^n)^\top z_t^n w_t^n]^2} + (1-\gamma)\frac{\hat{z}_t^{h\top}\hat{\sigma}_t^{n\top}\sigma z_t^n}{[1+\pi^{n\prime}(w_t^n)^\top z_t^n w_t^n]^2}\Big] \\
&- \frac{\gamma(1-\gamma)}{2}\frac{\hat{z}_t^{h\top}\hat{\sigma}_t^{n\top}\hat{\sigma}_t^n\hat{z}_t^h}{[1+\pi^{n\prime}(w_t^n)^\top z_t^n w_t^n]^2} + \frac{1}{2}(w_t^n)^2\frac{q^{h\prime\prime}(w_t^n;\lambda,\delta,\chi)}{q^h(w_t^n;\lambda,\delta,\chi)}\frac{z_t^{n\top}\Sigma z_t^n}{[1+\pi^{n\prime}(w_t^n)^\top z_t^n w_t^n]^2} \\
&+ \lambda\Big[\frac{q^n(w_t^n)}{q^h(w_t^n;\lambda,\delta,\chi)}\Big(1 + \frac{\chi}{\lambda}\hat{z}^{h\top}\epsilon\Big)^{1-\gamma} - 1\Big] = \rho.
\end{aligned} \tag{A.204}$$

Moreover,

$$\begin{aligned}
\hat{z}_t^h =& \frac{1}{\gamma}(\hat{\sigma}_t^{n\top}\hat{\sigma}_t^n)^{-1}\Big\{\Gamma^n(w_t^n)\hat{\sigma}_t^{n\top}\sigma z_t^n + [1+\pi^{n\prime}(w_t^n)^\top z_t^n w_t^n]^2(1-\chi)\epsilon\Big\} \\
&+ \frac{q^{h\prime}(w_t^n;\lambda,\delta,\chi,)}{\gamma q^h(w_t^n;\lambda,\delta\chi)} z_t^n w_t^n \\
&+ \frac{\chi}{\gamma}\frac{q^n(w_t^n)}{q^h(w_t^n;\lambda,\delta,\chi)}[1+\pi^{n\prime}(w_t^n)^\top z_t^n w_t^n]^2(\hat{\sigma}_t^{n\top}\hat{\sigma}_t^n)^{-1}\Big(1 + \frac{\chi}{\lambda}\hat{z}_t^{h\top}\epsilon\Big)^{-\gamma}\epsilon.
\end{aligned} \tag{A.205}$$

That is,

$$\begin{aligned}
\hat{z}_t^h =& \Big[\frac{1}{\gamma}\Gamma^n(w_t^n) + \frac{q^{h\prime}(w_t^n;\lambda,\delta,\chi)}{\gamma q^h(w_t^n;\lambda,\delta,\chi)}w_t^n\Big] z_t^n + \frac{1-\chi}{\gamma}[1+\pi^{n\prime}(w_t^n)^\top z_t^n w_t^n]^2(\hat{\sigma}_t^{n\top}\hat{\sigma}_t^n)^{-1}\epsilon \\
&+ \frac{\chi}{\gamma}\frac{q^n(w_t^n)}{q^h(w_t^n;\lambda,\delta,\chi)}[1+\pi^{n\prime}(w_t^n)^\top z_t^n w_t^n]^2(\hat{\sigma}_t^{n\top}\hat{\sigma}_t^n)^{-1}\Big(1 + \frac{\chi}{\lambda}\hat{z}_t^{h\top}\epsilon\Big)^{-\gamma}\epsilon.
\end{aligned} \tag{A.206}$$

B Connection to HHHKL

As explained in Section 6.2, we have

$$\sigma_{Pt}^{\xi} = \frac{\bar{\sigma}}{1 - P^{\xi\prime}(w_t^{\xi})^{\top} Y_t^{\xi}} Y_t^{\xi} P^{\xi\prime}(w_t^{\xi})^{\top}. \tag{B.207}$$

Then

$$\sigma_{Pt}^{\xi} + \sigma = \bar{\sigma} \left[\mathbf{I} + \frac{Y_t^{\xi} P^{\xi\prime}(w_t^{\xi})^{\top}}{1 - P^{\xi\prime}(w_t^{\xi})^{\top} Y_t^{\xi}} \right]. \tag{B.208}$$

Hence, the endogenous variance-covariance matrix of dollar returns can be written as

$$\begin{aligned} \Pi_t^{\xi} &\equiv (\sigma_{Pt}^{\xi} + \sigma)^{\top} (\sigma_{Pt}^{\xi} + \sigma) \\ &= \bar{\sigma}^2 \left[\mathbf{I} + \frac{Y_t^{\xi} P^{\xi\prime}(w_t^{\xi})^{\top}}{1 - P^{\xi\prime}(w_t^{\xi})^{\top} Y_t^{\xi}} \right]^{\top} \left[\mathbf{I} + \frac{Y_t^{\xi} P^{\xi\prime}(w_t^{\xi})^{\top}}{1 - P^{\xi\prime}(w_t^{\xi})^{\top} Y_t^{\xi}} \right] \\ &= \bar{\sigma}^2 \left[\mathbf{I} + \frac{P^{\xi\prime}(w_t^{\xi}) Y_t^{\xi\top}}{1 - P^{\xi\prime}(w_t^{\xi})^{\top} Y_t^{\xi}} \right] \left[\mathbf{I} + \frac{Y_t^{\xi} P^{\xi\prime}(w_t^{\xi})^{\top}}{1 - P^{\xi\prime}(w_t^{\xi})^{\top} Y_t^{\xi}} \right]. \end{aligned} \tag{B.209}$$

Taking the inverse of Π_t^{ξ} and using the Sherman-Morrison formula, we get

$$\begin{aligned} (\Pi_t^{\xi})^{-1} &= \frac{1}{\bar{\sigma}^2} \left[\mathbf{I} + \frac{Y_t^{\xi} P^{\xi\prime}(w_t^{\xi})^{\top}}{1 - P^{\xi\prime}(w_t^{\xi})^{\top} Y_t^{\xi}} \right]^{-1} \left[\mathbf{I} + \frac{P^{\xi\prime}(w_t^{\xi}) Y_t^{\xi\top}}{1 - P^{\xi\prime}(w_t^{\xi})^{\top} Y_t^{\xi}} \right]^{-1} \\ &= \frac{1}{\bar{\sigma}^2} \left[\mathbf{I} - Y_t^{\xi} P^{\xi\prime}(w_t^{\xi})^{\top} \right] \left[\mathbf{I} - P^{\xi\prime}(w_t^{\xi}) Y_t^{\xi\top} \right] \\ &= \frac{1}{\bar{\sigma}^2} \left[\mathbf{I} - P^{\xi\prime}(w_t^{\xi}) Y_t^{\xi\top} - Y_t^{\xi} P^{\xi\prime}(w_t^{\xi})^{\top} + P^{\xi\prime}(w_t^{\xi})^{\top} P^{\xi\prime}(w_t^{\xi}) Y_t^{\xi} Y_t^{\xi\top} \right]. \end{aligned} \tag{B.210}$$

We can define

$$\boldsymbol{\beta} \equiv \begin{bmatrix} Y_t^{\xi} & P^{\xi\prime}(w_t^{\xi}) \end{bmatrix}_{I \times 2} \tag{B.211}$$

and

$$\mathcal{E}_{sub} \equiv \begin{bmatrix} P^{\xi\prime}(w_t^{\xi})^{\top} P^{\xi\prime}(w_t^{\xi}) & -1 \\ -1 & 0 \end{bmatrix}_{2 \times 2} \tag{B.212}$$

Then it is easy to verify that

$$(\Pi_t^{\xi})^{-1} = \frac{1}{\bar{\sigma}^2} \left(\mathbf{I} + \boldsymbol{\beta} \mathcal{E}_{sub} \boldsymbol{\beta}^{\top} \right) \tag{B.213}$$

Substituting into the investor's individual demand (46), we get

$$\begin{aligned}\hat{Y}_t^\xi =& \frac{1}{\gamma^A(w_t^\xi)\bar{\sigma}^2}(\mathbf{I}+\boldsymbol{\beta}\mathcal{E}_{sub}\boldsymbol{\beta}^\top)(\bar{D}+\mu_{Pt}^\xi - rP_t^\xi) + \frac{1}{\gamma^A(w_t^\xi)}\frac{q^{\xi\prime}(w_t^\xi)}{q^\xi(w_t^\xi)}Y_t^\xi \\ &+ \frac{\lambda^\xi}{\gamma^A(w_t^\xi)}(\Pi_t^\xi)^{-1}\frac{q^{-\xi}(w_t^\xi + Y_t^{\xi\top}J_t^\xi)}{q^\xi(w_t^\xi)\left(1+\frac{1}{\hat{w}_t^\xi}\hat{Y}_t^{\xi\top}J_t^\xi\right)^\gamma}J_t^\xi.\end{aligned} \tag{B.214}$$

B.1 Factor Structure for Variance-Covariance Matrix

We now add another factor structure for the variance-covariance matrix of cash flows. Specifically, we suppose

$$\Sigma \equiv \sigma^\top\sigma = \bar{\sigma}^2(\mathbf{I}+\psi\psi^\top), \tag{B.215}$$

where $\bar{\sigma}$ is a scalar, $\mathbf{I}$ is an $I\times I$ identity matrix and ψ is $I\times K$. By eigen-decomposition, the diffusion σ can be written as

$$\sigma = \bar{\sigma}(\mathbf{I}+\nu\Lambda\nu^\top), \tag{B.216}$$

where ν are the eigenvectors of $\psi\psi^\top$ corresponding to its non-zero eigenvalues (size $I\times K$), and $\Lambda = \text{diag}(\lambda_1,\ldots,\lambda_K)$ where $\lambda_i = -\bar{\sigma}+\sqrt{\bar{\sigma}^2+s_i^2}$ where s_i^2 denotes eigenvalues of $\psi\psi^\top$.

Then the endogenous variance-covariance matrix of dollar returns can be written as

$$\begin{aligned}\Pi_t^\xi &\equiv (\sigma_{Pt}^\xi+\sigma)^\top(\sigma_{Pt}^\xi+\sigma) \\ &= \bar{\sigma}^2\left[\mathbf{I}+\nu\Lambda\nu^\top+\frac{Y_t^\xi P^{\xi\prime}(w_t^\xi)^\top}{1-P^{\xi\prime}(w_t^\xi)^\top Y_t^\xi}\right]^\top\left[\mathbf{I}+\nu\Lambda\nu^\top+\frac{Y_t^\xi P^{\xi\prime}(w_t^\xi)^\top}{1-P^{\xi\prime}(w_t^\xi)^\top Y_t^\xi}\right] \\ &= \bar{\sigma}^2\left[\mathbf{I}+\nu\Lambda\nu^\top+\frac{P^{\xi\prime}(w_t^\xi)Y_t^{\xi\top}}{1-P^{\xi\prime}(w_t^\xi)^\top Y_t^\xi}\right]\left[\mathbf{I}+\nu\Lambda\nu^\top+\frac{Y_t^\xi P^{\xi\prime}(w_t^\xi)^\top}{1-P^{\xi\prime}(w_t^\xi)^\top Y_t^\xi}\right].\end{aligned} \tag{B.217}$$

Let $A\equiv\mathbf{I}+\nu\Lambda\nu^\top$. Taking the inverse of Π_t^ξ and using the Sherman-Morrison formula, we have

$$\begin{aligned}(\Pi_t^\xi)^{-1} &= \frac{1}{\bar{\sigma}^2}\left[\mathbf{I}+\nu\Lambda\nu^\top+\frac{Y_t^\xi P^{\xi\prime}(w_t^\xi)^\top}{1-P^{\xi\prime}(w_t^\xi)^\top Y_t^\xi}\right]^{-1}\left[\mathbf{I}+\nu\Lambda\nu^\top+\frac{P^{\xi\prime}(w_t^\xi)Y_t^{\xi\top}}{1-P^{\xi\prime}(w_t^\xi)^\top Y_t^\xi}\right]^{-1} \\ &= \frac{1}{\bar{\sigma}^2}\left[A+\frac{Y_t^\xi P^{\xi\prime}(w_t^\xi)^\top}{1-P^{\xi\prime}(w_t^\xi)^\top Y_t^\xi}\right]^{-1}\left[A+\frac{P^{\xi\prime}(w_t^\xi)Y_t^{\xi\top}}{1-P^{\xi\prime}(w_t^\xi)^\top Y_t^\xi}\right]^{-1} \\ &= \frac{1}{\bar{\sigma}^2}\left[A^{-1}-\frac{A^{-1}Y_t^\xi P^{\xi\prime}(w_t^\xi)^\top A^{-1}}{1-P^{\xi\prime}(w_t^\xi)^\top Y_t^\xi+P^{\xi\prime}(w_t^\xi)^\top A^{-1}Y_t^\xi}\right]\left[A^{-1}-\frac{A^{-1}P^{\xi\prime}(w_t^\xi)Y_t^{\xi\top}A^{-1}}{1-P^{\xi\prime}(w_t^\xi)^\top Y_t^\xi+Y_t^{\xi\top}A^{-1}P^{\xi\prime}(w_t^\xi)}\right].\end{aligned} \tag{B.218}$$

By Woodbury identity,

$$A^{-1} = \mathbf{I} - \nu(\Lambda^{-1} + \nu^\top \nu)^{-1}\nu^\top. \tag{B.219}$$

Substituting into above, we get

$$\begin{aligned}(\Pi_t^\xi)^{-1} = \frac{1}{\bar{\sigma}^2}\Bigg[\mathbf{I} - \nu(\Lambda^{-1} + \nu^\top\nu)^{-1}\nu^\top - \frac{A^{-1}Y_t^\xi P^{\xi\prime}(w_t^\xi)^\top A^{-1}}{1 - P^{\xi\prime}(w_t^\xi)^\top Y_t^\xi + P^{\xi\prime}(w_t^\xi)^\top A^{-1}Y_t^\xi}\Bigg] \\ \Bigg[\mathbf{I} - \nu(\Lambda^{-1} + \nu^\top\nu)^{-1}\nu^\top - \frac{A^{-1}P^{\xi\prime}(w_t^\xi)Y_t^{\xi\top} A^{-1}}{1 - P^{\xi\prime}(w_t^\xi)^\top Y_t^\xi + Y_t^{\xi\top}A^{-1}P^{\xi\prime}(w_t^\xi)}\Bigg]\end{aligned} \tag{B.220}$$

Let $d_1 \equiv 1 - P^{\xi\prime}(w_t^\xi)^\top Y_t^\xi + P^{\xi\prime}(w_t^\xi)^\top A^{-1}Y_t^\xi$ and $d_2 \equiv 1 - P^{\xi\prime}(w_t^\xi)^\top Y_t^\xi + Y_t^{\xi\top}A^{-1}P^{\xi\prime}(w_t^\xi)$. Then,

$$\begin{aligned}(\Pi_t^\xi)^{-1} = \frac{1}{\bar{\sigma}^2}\Big\{&\mathbf{I} - 2\nu(\Lambda^{-1} + \nu^\top\nu)^{-1}\nu^\top - \frac{A^{-1}P^{\xi\prime}(w_t^\xi)Y_t^{\xi\top}A^{-1}}{d_2} \\ &+ \nu(\Lambda^{-1} + \nu^\top\nu)^{-1}\nu^\top\nu(\Lambda^{-1} + \nu^\top\nu)^{-1}\nu^\top + \frac{\nu(\Lambda^{-1} + \nu^\top\nu)^{-1}\nu^\top A^{-1}P^{\xi\prime}(w_t^\xi)Y_t^{\xi\top}A^{-1}}{d_2} \\ &- \frac{A^{-1}Y_t^\xi P^{\xi\prime}(w_t^\xi)^\top A^{-1}}{d_1} + \frac{A^{-1}Y_t^\xi P^{\xi\prime}(w_t^\xi)^\top A^{-1}\nu(\Lambda^{-1} + \nu^\top\nu)^{-1}\nu^\top}{d_1} \\ &+ \frac{P^{\xi\prime}(w_t^\xi)^\top A^{-1}A^{-1}P^{\xi\prime}(w_t^\xi)}{d_1 d_2}A^{-1}Y_t^\xi Y^{\xi\top}A^{-1}\Big\}.\end{aligned} \tag{B.221}$$

Let us define

$$\boldsymbol{\beta}^{full} \equiv \Big[\nu \quad A^{-1}Y_t^\xi \quad A^{-1}P^{\xi\prime}(w_t^\xi)\Big]_{I\times(K+2)} \tag{B.222}$$

and

$$\mathcal{E}^{full} \equiv \begin{bmatrix} (\Lambda^{-1} + \nu^\top\nu)^{-1}\nu^\top\nu(\Lambda^{-1} + \nu^\top\nu)^{-1} - 2(\Lambda^{-1} + \nu^\top\nu)^{-1} & \frac{(\Lambda^{-1}+\nu^\top\nu)^{-1}\nu^\top A^{-1}P^{\xi\prime}(w_t^\xi)}{d_2} & 0 \\ \frac{P^{\xi\prime}(w_t^\xi)^\top A^{-1}\nu(\Lambda^{-1}+\nu^\top\nu)^{-1}}{d_1} & \frac{P^{\xi\prime}(w_t^\xi)^\top A^{-1}A^{-1}P^{\xi\prime}(w_t^\xi)}{d_1 d_2} & -\frac{1}{d_1} \\ 0 & -\frac{1}{d_2} & 0 \end{bmatrix} \tag{B.223}$$

which is of size $(K+2)\times(K+2)$. Then

$$(\Pi_t^\xi)^{-1} = \mathbf{I} + \boldsymbol{\beta}^{full}\mathcal{E}^{full}\boldsymbol{\beta}^{full\top}. \tag{B.224}$$

C The CARA case

We also consider the case with CARA investors. For these investors, their value functions are independent of the investors' total wealth w_t. Thus, w_t is not a state variable in this case. An investor maximizes

$$-\mathbb{E}_t\Big[\int_t^\infty e^{-\rho(\tau-t)}e^{-\frac{\gamma}{r}\hat{c}_\tau}d\tau\Big], \tag{C.225}$$

subject to the wealth dynamic

$$d\hat{w}_t^\xi = r\hat{w}_t^\xi dt + \hat{Y}_t^{\xi\top}(dD_t - rP_t^\xi dt) - \hat{c}_t^\xi dt. \tag{C.226}$$

Note that the asset prices depend only on the state of the economy ξ. That is, in a given state of the economy, the asset prices are constant. Thus, the investor's dollar excess returns from investing in the risky assets become $dD_t - rP_t^\xi dt$. Substituting (3) into the investor's wealth dynamic (C.226) yields

$$d\hat{w}_t^\xi = \Big[r\hat{w}_t^\xi - \hat{c}_t^\xi + \hat{Y}_t^{\xi\top}(\bar{D} - rP_t^\xi)\Big]dt + \hat{Y}_t^{\xi\top}\sigma^\top dB_t. \tag{C.227}$$

The investor's problem is to choose the optimal consumption and asset holdings $\{\hat{c}_t^\xi, \hat{Y}_t^\xi\}$ to maximize their utility (C.225) subject to the wealth dynamic (C.227) while taking prices P_t^ξ as given. Let $V^\xi(\hat{w}_t^\xi)$ denote the investor's value function in state ξ. The investor's HJB equation can be written as

$$\rho V^\xi = \max_{\hat{c}_t^\xi, \hat{Y}_t^\xi}\Big\{-e^{-\frac{\gamma}{r}\hat{c}_t^\xi} + V_{\hat{w}}^\xi \mu_t^{\hat{w},\xi} + \frac{1}{2}V_{\hat{w}\hat{w}}^\xi(\sigma_t^{\hat{w},\xi})^\top\sigma_t^{\hat{w},\xi} + \lambda^\xi(V^{-\xi} - V^\xi)\Big\}, \tag{C.228}$$

where

$$\mu_t^{\hat{w},\xi} = r\hat{w}_t^\xi - \hat{c}_t^\xi + \hat{Y}_t^{\xi\top}(\bar{D} - rP_t^\xi), \tag{C.229}$$

$$\sigma_t^{\hat{w},\xi} = \sigma\hat{Y}_t^\xi. \tag{C.230}$$

We conjecture (and later verify) that the investor's value function takes the form

$$V^\xi(\hat{w}_t^\xi) = -e^{-\gamma\hat{w}_t^\xi}q^\xi, \tag{C.231}$$

where q^ξ is a constant. The long-term CARA investor's value function over wealth has the same negative-exponential form as the utility function over consumption, with the risk-aversion coefficient being γ rather than γ/r. Substituting the conjectured value function form into the HJB equation, we can rewrite the HJB equation as

$$\begin{aligned}-\rho e^{-\gamma\hat{w}_t^\xi}q^\xi = \max_{\hat{c}_t^\xi, \hat{Y}_t^\xi}\Big\{&-e^{-\frac{\gamma}{r}\hat{c}_t^\xi} + \gamma e^{-\gamma\hat{w}_t^\xi}q^\xi\Big[r\hat{w}_t^\xi - \hat{c}_t^\xi + \hat{Y}_t^{\xi\top}(\bar{D} - rP_t^\xi)\Big]\\ &- \frac{1}{2}\gamma^2 e^{-\gamma\hat{w}_t^\xi}q^\xi\hat{Y}_t^{\xi\top}\Sigma\hat{Y}_t^\xi + \lambda^\xi\Big[-e^{-\gamma[\hat{w}_t^\xi + \hat{Y}_t^{\xi\top}(P_t^{-\xi} - P_t^\xi)]}q^{-\xi} + e^{-\gamma\hat{w}_t^\xi}q^\xi\Big]\Big\}.\end{aligned} \tag{C.232}$$

Taking the first-order condition with respect to $\hat{c}_t$ yields

$$\frac{\gamma}{r}e^{-\frac{\gamma}{r}\hat{c}_t^{\xi}} = \gamma e^{-\gamma\hat{w}_t^{\xi}}q^{\xi}, \tag{C.233}$$

which implies that the investor's optimal consumption $\hat{c}_t^{\xi}$ at time t if the economy is in state ξ is given by

$$\hat{c}_t^{\xi} = r\hat{w}_t^{\xi} - \frac{r}{\gamma}(\log q^{\xi} + \log r). \tag{C.234}$$

Taking the first-order condition with respect to $\hat{Y}_t^{\xi}$ yields

$$\begin{aligned}&\gamma e^{-\gamma\hat{w}_t^{\xi}}q^{\xi}(\bar{D} - rP_t^{\xi}) - \gamma^2 e^{-\gamma\hat{w}_t^{\xi}}q^{\xi}\Sigma\hat{Y}_t^{\xi} + \lambda^{\xi}\gamma e^{-\gamma[\hat{w}_t^{\xi}+\hat{Y}_t^{\xi\top}(P_t^{-\xi}-P_t^{\xi})]}q^{-\xi}(P_t^{-\xi} - P_t^{\xi})\\ &= 0.\end{aligned} \tag{C.235}$$

Rearranging, we obtain the investor's optimal asset holdings $\hat{Y}_t^{\xi}$ as

$$\hat{Y}_t^{\xi} = \frac{1}{\gamma}\Sigma^{-1}(\bar{D} - rP_t^{\xi}) + \frac{\lambda^{\xi}}{\gamma}\frac{q^{-\xi}}{q^{\xi}}e^{-\gamma\hat{Y}_t^{\xi\top}(P_t^{-\xi}-P_t^{\xi})}\Sigma^{-1}(P_t^{-\xi} - P_t^{\xi}). \tag{C.236}$$

The investor's individual demand can be written as

$$\begin{aligned}\hat{Y}_t^{\xi} =&\frac{1}{\gamma}\Sigma^{-1}[\bar{D} - rP_t^{\xi} + \lambda^{\xi}(P_t^{-\xi} - P_t^{\xi})]\\ &+ \frac{\lambda^{\xi}}{\gamma}\Big[\frac{q^{-\xi}}{q^{\xi}}e^{-\gamma\hat{Y}_t^{\xi\top}(P_t^{-\xi}-P_t^{\xi})} - 1\Big]\Sigma^{-1}(P_t^{-\xi} - P_t^{\xi}),\end{aligned} \tag{C.237}$$

where $\bar{D} - rP_t^{\xi} + \lambda^{\xi}(P_t^{-\xi} - P_t^{\xi})$ corresponds to the investor's expected dollar excess returns at time t when the economy is in state ξ.

Substituting the investor's optimal consumption $\hat{c}_t^{\xi}$ given by (C.234) and the investor's optimal asset holdings $\hat{Y}_t^{\xi}$ given by (C.236) into the HJB equation, we have

$$\begin{aligned}-\rho =&r(\log q^{\xi} + \log r - 1) + \gamma\hat{Y}_t^{\xi\top}(\bar{D} - rP^{\xi}) - \frac{1}{2}\gamma^2\hat{Y}_t^{\xi\top}\Sigma\hat{Y}_t^{\xi}\\ &+ \lambda^{\xi}\Big[1 - \frac{q^{-\xi}}{q^{\xi}}e^{-\gamma\hat{Y}_t^{\xi\top}(P_t^{-\xi}-P_t^{\xi})}\Big]\end{aligned} \tag{C.238}$$

In equilibrium, because there is a unit measure of symmetric investors, the aggregate asset holdings by investors $Y_t^{\xi} = \hat{Y}_t^{\xi}$. Market clearing requires that

$$Q_t^{\xi} + Y_t^{\xi} = S \tag{C.239}$$

where Q_t^{ξ} equals hedgers' aggregate asset holdings, and is given by (A.6). From (A.6) and

(C.236),

$$\begin{aligned}&\Big[\Sigma+\lambda^{\xi}(P_t^{-\xi}-P_t^{\xi})(P_t^{-\xi}-P_t^{\xi})^{\top}\Big]^{-1}\Big[\frac{\bar{D}-rP_t^{\xi}+\lambda^{\xi}(P_t^{-\xi}-P_t^{\xi})}{\alpha}-\Sigma u^{\xi}\Big]\\&+\frac{1}{\gamma}\Sigma^{-1}(\bar{D}-rP_t^{\xi})+\frac{\lambda^{\xi}}{\gamma}\frac{q^{-\xi}}{q^{\xi}}e^{-\gamma\hat{Y}_t^{\xi\top}(P_t^{-\xi}-P_t^{\xi})}\Sigma^{-1}(P_t^{-\xi}-P_t^{\xi})=S\end{aligned}\tag{C.240}$$

As explained earlier, the risky assets' prices depend only on the state of the economy ξ. Hence, we can define the state-dependent constant

$$\pi^{\xi}\equiv\frac{\bar{D}}{r}-P_t^{\xi}\tag{C.241}$$

Substituting (C.241) into (C.236), (C.238) and (C.240), and noting that $\hat{Y}_t^{\xi}=Y_t^{\xi}$, we have the following system of equations

$$Y_t^{\xi}=\frac{r}{\gamma}\Sigma^{-1}\pi^{\xi}+\frac{\lambda^{\xi}}{\gamma}\frac{q^{-\xi}}{q^{\xi}}e^{-\gamma Y_t^{\xi\top}(\pi^{\xi}-\pi^{-\xi})}\Sigma^{-1}(\pi^{\xi}-\pi^{-\xi})\tag{C.242}$$

and

$$-\rho=r(\log q^{\xi}+\log r-1)+r\gamma Y_t^{\xi\top}\pi^{\xi}-\frac{1}{2}\gamma^2Y_t^{\xi\top}\Sigma Y_t^{\xi}+\lambda^{\xi}\Big[1-\frac{q^{-\xi}}{q^{\xi}}e^{-\gamma Y_t^{\xi\top}(\pi^{\xi}-\pi^{-\xi})}\Big]\tag{C.243}$$

as well as

$$\begin{aligned}&\Big[\Sigma+\lambda^{\xi}(\pi^{\xi}-\pi^{-\xi})(\pi^{\xi}-\pi^{-\xi})^{\top}\Big]^{-1}\Big[\frac{r\pi^{\xi}+\lambda^{\xi}(\pi^{\xi}-\pi^{-\xi})}{\alpha}-\Sigma u^{\xi}\Big]\\&+\frac{r}{\gamma}\Sigma^{-1}\pi^{\xi}+\frac{\lambda^{\xi}}{\gamma}\frac{q^{-\xi}}{q^{\xi}}e^{-\gamma Y_t^{\xi\top}(\pi^{\xi}-\pi^{-\xi})}\Sigma^{-1}(\pi^{\xi}-\pi^{-\xi})=S\end{aligned}\tag{C.244}$$

The equilibrium prices P_t^{ξ} of the risky assets at time t when the economy is in state ξ are

$$P_t^{\xi}=\frac{\bar{D}}{r}-\pi^{\xi},\tag{C.245}$$

where $\xi\in\{n,s\}$. For all ξ, Y_t^{ξ}, π^{ξ} and q^{ξ} satisfy the sysem of equations

$$Y_t^{\xi}=\frac{r}{\gamma}\Sigma^{-1}\pi^{\xi}+\frac{\lambda^{\xi}}{\gamma}\frac{q^{-\xi}}{q^{\xi}}e^{-\gamma Y_t^{\xi\top}(\pi^{\xi}-\pi^{-\xi})}\Sigma^{-1}(\pi^{\xi}-\pi^{-\xi}),\tag{C.246}$$

and

$$-\rho=r(\log q^{\xi}+\log r-1)+r\gamma Y_t^{\xi\top}\pi^{\xi}-\frac{1}{2}\gamma^2Y_t^{\xi\top}\Sigma Y_t^{\xi}+\lambda^{\xi}\Big[1-\frac{q^{-\xi}}{q^{\xi}}e^{-\gamma Y_t^{\xi\top}(\pi^{\xi}-\pi^{-\xi})}\Big],\tag{C.247}$$

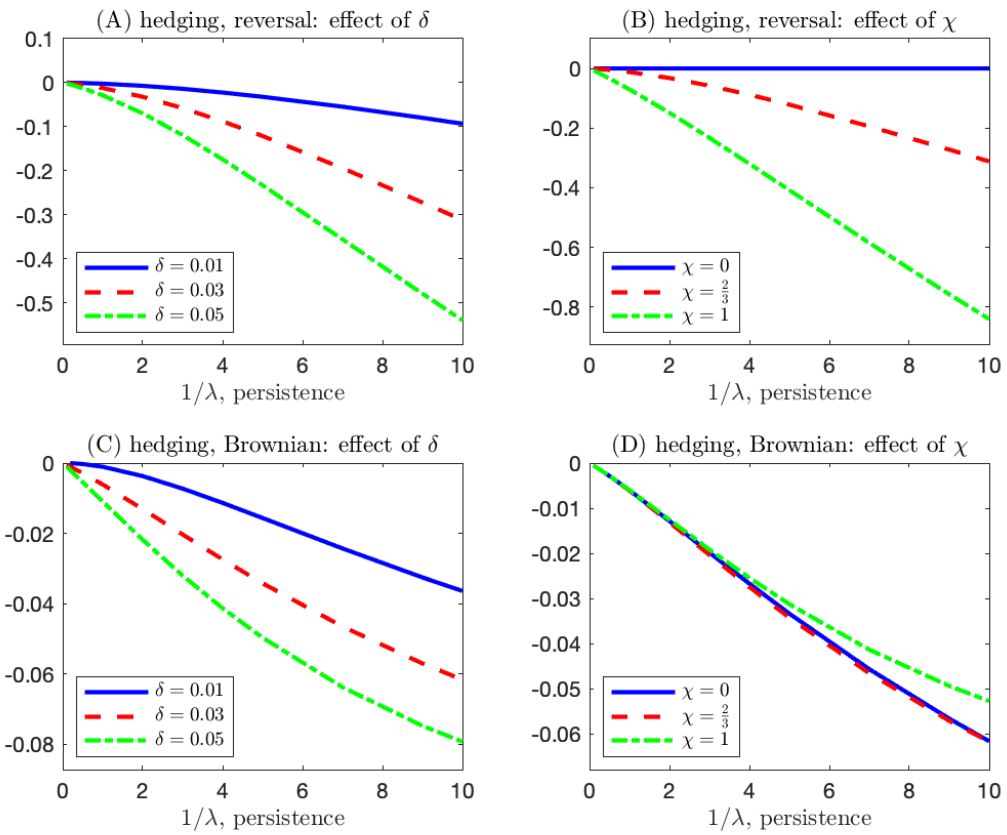


Figure D.1: Dynamic slope $\mathcal{C}$ and its decomposition for CRRA investors: myopic, Brownian hedging, and hedging against reversal components. We plot the decomposition for different size and persistence of the return shock. The parameter values are $\bar{D} = [0.27, 0.26]$, $\sigma = \text{diag}(0.485, 0.485)$, $\alpha = 1$, $\gamma = 3$, $r = 0.03$, $\rho = 0.08$, $s = [1, 1]$, and we consider $\delta = 0.01, 0.03, 0.05$ and $\chi = 0, \frac{2}{3}, 1$.

as well as

$$\begin{aligned}&\left[\Sigma + \lambda^{\xi}(\pi^{\xi} - \pi^{-\xi})(\pi^{\xi} - \pi^{-\xi})^{\top}\right]^{-1}\left[\frac{r\pi^{\xi} + \lambda^{\xi}(\pi^{\xi} - \pi^{-\xi})}{\alpha} - \Sigma u^{\xi}\right] \\ &+ \frac{r}{\gamma}\Sigma^{-1}\pi^{\xi} + \frac{\lambda^{\xi}}{\gamma}\frac{q^{-\xi}}{q^{\xi}}e^{-\gamma Y_t^{\xi\top}(\pi^{\xi} - \pi^{-\xi})}\Sigma^{-1}(\pi^{\xi} - \pi^{-\xi}) = S.\end{aligned} \tag{C.248}$$

D The economics of the Myopic Slope vs. the Dynamic Slope

Figure D.1 explores the comparative statics of the three parameters (return size δ, persistence $1/\lambda$, and reversal share χ) for the two hedging components. Panel A and B of Figure D.1 plot the intertemporal hedging with respect to the (Poisson) reversal, which is relatively easier to explain. Investors understand that at some point the shift in residual supply will be reversed; when this occurs they realize a higher return, but, at the same time, their investment opportunity set gets worse. Investors hedge this by cutting their asset exposure,

explaining the negative "reversal hedging." Finally, it is intuitive that this negative effect gets stronger for larger shocks (larger δ) and larger reversal share (χ). To the extreme of $\chi = 0$ there is no hedging component due to reversal (Panel B).

Panel C and D then illustrate the hedging demand with respect to Brownian shocks, with risk aversion $\gamma > 1$.[24] From (40), the sign of Brownian hedging depends on the sign

$$sgn\left\{\frac{q^{h\prime}\left(w_t^n\right)}{q^h\left(w_t^n\right)} - \frac{q^{n\prime}\left(w_t^n\right)}{q^n\left(w_t^n\right)}\right\} = sgn\left\{\left(\frac{q^h\left(w_t^n\right)}{q^n\left(w_t^n\right)}\right)'_{w^n}\right\}. \tag{D.249}$$

The simple expression in (D.249) captures the intuition of Brownian hedging. If the relative investment opportunities between two states $\frac{q^h(w_t^n)}{q^n(w_t^n)}$ decreases with the aggregate state w_t^n, then because the asset return is positively correlated with the Brownian shock, the investor should reduce her hedging demand of the asset with respect to (aggregate) Brownian shocks. Now, in line with Merton (1971) $q^n(w_t^n)$ is increasing in w_t^n for $\gamma > 1$,while the investment opportunity offered by the "deal" in Definition 2 features an extra expected return δ that is independent of w_t^n. To the extreme, when $\delta \to \infty$ the "deal" fully dominates the stochastic investment opportunity available in the economy and $q^h(w)$ becomes a constant. As a result, $\frac{q^h(w_t^n)}{q^n(w_t^n)}$ tends to be decreasing in w, giving rise to a negative Brownian hedging demand.

Several additional points are worth making on the comparative statics revealed in Figure D.1. First, it is intuitive that when the reversal share χ gets larger the Brownian component gets smaller. Second, when $\lambda \to \infty$, the investor is as if receiving a fixed extra return δdt on his wealth over the interval $[t, t + dt]$ and then immediately followed by the equilibrium investment opportunity; hence $\frac{q^h(w)}{q^n(w)}$ should be a constant independent of w, which explains zero hedging demand observed in all panels at the corner of when $1/\lambda \to 0$. Third, unlike the reversal component in Panel D, the Brownian component does not disappear as $\delta \to 0$ so that the dynamic slope stays strictly below the myopic one.

Finally, although in our later calibration in Section E.1 the effect of the (two) dynamic hedging components—which in sum is only about 1.4% of the dynamic slope—is quantitatively small,[25] from Figure D.1 they could be potentially large for a shock that is large and persistent.

E Details of Calibration and Sensitivity Analysis

In this part, we give details on the calibration excercises which are the basis of Section 6.1.

In our first calibration, we calibrate to the study of Ben-David et al. (2021) focusing on factor-level elasticities. In the second one, we are interested in an economy where the residual supply shock affects only very few assets out of many. Hence, we consider the case of index inclusion calibrating our model to the recent study of Greenwood and Sammon (2024).

[24] It is well known since Merton (1971) that all else equal the sign of this component flips if $\gamma < 1$.

[25] It is about 0.05 out of circa 3.60 in the factor elasticity exercise, with calibrated parameters $1/\lambda = 2$, $\delta = 0.03$ and $\chi = 2/3$. See more details in Section E.1.

Incr. Return ($\Delta\mathbb{E}_t(R)$)	0.01	0.03	0.05	0.08	0.10
Myopic Slope	3.21	3.21	3.21	3.22	3.22
Dynamic Slope	3.20	3.17	3.13	3.05	3.00
Observed Slope	1.39	1.24	1.10	0.93	0.83
Observed Slope/Dynamic Slope	43%	39%	35%	30%	28%
Implied Price Elasticity	1.50	1.42	1.35	1.24	1.17
Dynamic minus Observed (DmO)	1.81	1.93	2.02	2.13	2.17
Shock to risk/DmO	114%	112%	111%	111%	112%
Diff Hedging Brownian/DmO	-15%	-13%	-12%	-11%	-10%
Diff Hedging Reversal/DmO	1%	1%	1%	0%	-2%

Table E.1: Comparative statics with respect to expected incremental return $\Delta\mathbb{E}_t(R)$

E.1 Fund flows and the Factor-Elasticity

In this part, we calibrate our model to Ben-David et al. (2021) study on factor-level elasticities. Then, using the obtained set of parameters as our baseline, we assess how our measures respond to a change of various parameters.

E.1.1 Baseline calibration

Our baseline parameters are obtained by calibrating our model to Ben-David et al. (2021). They show that when investment styles are highly correlated with Morningstar ratings, changes in fund ratings result in style-level fund flows which lead to price effect. In our calibration, we set $N = 2$. We interpret the two assets as two portfolios of different investment styles. In particular, we consider the first asset as the small value portfolio and the second asset as the large growth portfolio. The parameters $\bar{D}$ and σ pertain to the assets' fundamentals, and govern the first and second moments of portfolio returns. We set $\bar{D} = [0.27, 0.26]$ and $\sigma = diag(0.485, 0.485)$ to match the historical mean returns and return volatilities of the small value portfolio and the large growth portfolio as obtained from Ken French's data library.

We normalize the shocks in the normal state to zero, that is $u^n = [0, 0]$. We normalize the shock to the second asset to zero in the shock state and set $u^s = [0.15, 0]$ such that the model-implied difference in flows between the two assets (interpreted as the top and the bottom styles) is about 13%, consistent with the reported style-based fund flows in Ben-David et al. (2021). We set $\lambda^n = 0$, that is the economy transitions from the normal state to the shock state with intensity 0 and thus the normal state is the steady state in our calibration. Furthermore, we set $\lambda^s = 0.5$ (and hence a persistence 2 of the residual supply shock) to match the empirical half life of fund flows as reported by Ben-David et al. (2021).

We set $r = 0.03$, which roughly equals the historical average risk-free rate. The investors' time discount rate ρ governs the aggregate asset holdings by hedgers. We interpret the hedgers in our model to include households, mutual funds and ETFs, foreign investors, etc. We set $\rho = 0.08$, which implies that hedgers hold 59% of the assets. Finally, we set $\gamma = 3$ which is taken from the literature. We normalize the hedgers' absolute risk aversion α and the supply of each asset to one. Our baseline parameters are provided in Table 1.

Risk Aversion (γ)	0.7	1.5	2.0	3.0	3.5	log
Myopic Slope	9.59	5.37	4.34	3.21	2.84	7.28
Dynamic Slope	9.26	5.30	4.30	3.17	2.80	7.10
Observed Slope	1.43	1.48	1.41	1.24	1.15	1.49
Observed Slope/Dynamic Slope	15%	28%	33%	39%	41%	21%
Implied Price Elasticity	1.34	1.55	1.55	1.42	1.34	1.46
Dynamic minus Observed (DmO)	7.82	3.82	2.89	1.93	1.65	5.61
Shock to risk/DmO	97%	106%	109%	112%	112%	101%
Diff Hedging Brownian/DmO	4%	-6%	-9%	-13%	-13%	0%
Diff Hedging Reversal/DmO	-1%	0%	0%	1%	1%	-1%

Table E.2: Comparative statics with respect to risk aversion coefficient γ

Using the baseline parameters, we simulate the distribution of investors' aggregate wealth. In the baseline, the mean wealth of investors is 2.11. In our model, the hedgers' endowment shocks give rise to trading needs by the hedgers. As the economy switches from the normal state to the shock state, the hedgers wish to hold less of the assets leading to increased expected returns. At the mean investors' wealth level, our calibration implies an incremental return of $\Delta\mathbb{E}_t(R) = 3\%$, of which roughly 2/3 comes from the reversal.

It is worthwhile mapping our observed slope to the empirical demand elasticity à la Koijen and Yogo (2019), defined as the percentage of capitalization traded divided by the price impact. In Section E.2.2, we provide a detailed discussion of how our observed slope translates to demand elasticity. Using the baseline parameters, the corresponding demand elasticity is 1.4. In Ben-David et al. (2021), the empirical factor elasticity is around 0.2. While our seven-fold difference is inline with the general observation that frictionless asset pricing models struggle to rationalize the empirically small factor elasticity, our difference is smaller than it is perhaps previously thought.[26]

E.1.2 Results and slope analysis

Tables E.1-2 show the results of our baseline calibration and its slope to various parameters. In each table the colored column contains our benchmark parametrization, while the other columns show how our effects vary. For each set of parameters, we calculate the myopic slope, the dynamic slope and the observed slope, and decompose the difference to our three effects. Implied elasticity is a scaled version of our observed slope to make it comparable to the estimates collected by Gabaix and Koijen (2022) from the empirical literature.

Table E.1 focuses on the non-linear effect of the size of the expected return shocks implied by the shift in residual supply. Table E.2 illustrates the effect of varying relative-risk-aversion of investors. Table 2 concentrates on the effect of the persistence of the demand shift.[27]

There are a number of observations to make that are robust across our parameterizations.

[26] While the exact numbers widely vary, Gabaix and Koijen (2022) suggest that the empirical estimates are several order of magnitudes smaller than their conceptual counterpart.

[27] In Tables E.2-2, for each column we recalibrate the parameters u^ξ to ensure that the implied expected return shocks remain constant at $\Delta\mathbb{E}_t(R) = 0.025$.

First, consistent with the illustration in Panel A and B in Figure 2, we find that dynamic slope is always smaller than myopic slope. That is, Merton's intertemporal hedging tend to make the dynamic demand curve flatter compared to its static equivalent. However, this effect is small.

Second, in each of our parametrizations, observed slope significantly undershoots dynamic slope. In particular, observed slope is only 41% of demand slope in our baseline calibration. This fraction is shrinks to 15% when investors are less risk-averse than log. Larger persistence and smaller wealth monotonicly increases this fraction. This illustrates the main observation of this paper. Because demand shifts change the slope of the demand curve in equilibrium, the econometrician relying on these shifts for estimating the demand slope is shooting a moving target. This calibration suggests that this results in significant underestimation of the targeted concept.

Table $E.1$ illustrates the non-linearity of this effect by varying the size of the shock. Note that the fraction of observed slope to dynamic slope is shrinking when shocks are larger implying larger mismeasurement.

Third, the endogenous risk effect is responsible for the largest part of this mismeasurement. In fact, because the amplified intertemporal hedging component is going against the endogenous risk effect whenever $\gamma > 1$, in all these cases the endogenous risk effect is more than the 100% of the total under-measurement. This share is growing with risk-aversion and investor's wealth and shrinking with the persistence of the shock ranging from 102% (in our only parameterization with $\gamma < 1$) to 140%.

Fourth, while the GE amplified intertemporal hedging component is the same sign as Merton's intertemporal hedging effect, it is several magnitudes higher. In absolute terms, it can reach 34% of the total mismeasurement. Its absolute magnitude is larger for less persistent shocks, larger wealth and for greater risk-aversion.

Finally, the hedging effect against the reversal can be sizable reaching 31% for large shocks and 15% for persistence shocks relative to total mismeasurement.

E.2 Index Inclusion: Micro-Elasticity

In our second application, we are interested in an economy where the residual supply shock affects only very few assets. To make this point clear, we picture a market with a large number of fundamentally identical assets. That is, all I assets have identical average dividend $\bar{D}$, and $[\Sigma]_{ii} = \sigma^2$ and $[\Sigma]_{ij} = \kappa\sigma^2$ and their supply is $[s]_i = s$ for all i and $j \neq i$. The only difference across these assets is in hedgers' endowment. In particular, suppose that in the normal state, hedgers endowment for the first I_1 asset differs from the rest:

$$[u^n]_{i=1}^{I_1} = \upsilon_1,\ [u^n]_{i=I_1+1}^{I} = \upsilon_2.$$

where $\upsilon_1 \neq \upsilon_2$ are scalars.

Our interpretation is that the first I_1 asset is in an index. We are interested in a shock for which $[u^n]_{i=2}^{I_1} = [u^s]_{i=2}^{I_1}$ while $[u^s]_1 = \upsilon_2$, $[u^s]_I = \upsilon_1$. That is, the first asset is excluded from and the last asset is included in the index.

Note that we intentionally design this example, so that the shock itself has a minimal effect on the structure of the economy—effectively, we only swap the labels of the first and the last assets. Hence, one might expect that the shock should have a negligible effect on the

equilibrium outcome; hence, the informativeness of shifts to residual supply should perhaps be restored. We show that this is not the case.

In this part, we calibrate to Greenwood and Sammon (2024), which studies the observed slope of index inclusion and deletion events. We set $I = 502$ and $I_1 = 501$ such that the index in our calibration corresponds to the S&P 500 Index. Asset 1 and asset 2 switch in and out of the S&P 500 Index, while assets 3 to 501 are always in the index and the last asset is always out of the index. $\bar{D}_i$ is the expected cash flow of asset i, where $i \in \{1, 2, \ldots, I\}$. σ_i is the asset's cash flow volatility. κ is the pairwise cash flow correlation between asset i and asset j, where $i \neq j \in \{1, 2, \ldots, I\}$. We calibrate $\bar{D}_i$ and σ_i to match historical return moments. We set $\kappa = 0.3$ such that roughly 75% of an asset's return volatility is idiosyncratic (see Campbell, Lettau, Malkiel and Xi (2001)).

As in the our calibration of the factor-elasticity, we normalize the hedgers' absolute risk aversion α to one. We set $r = 0.03$, which roughly equals the historical risk-free rate. We also set the investors' time discount rate $\rho = 0.08$, the same as in the factor-elasticity calibration. ρ affects the aggregate asset holdings by hedgers, which are interpreted as mutual funds and ETFs, pension funds, etc. In our calibration, hedgers hold 45% of index assets and 2% of non-index assets. We set $\gamma = 3$, which is taken from the literature.

We normalize the supply of each asset to $1/\sqrt{I} = 0.0446$. To match the fund flows due to index inclusion and deletion in Greenwood and Sammon (2024), we normalize hedgers' endowment shock v_1 to zero for index assets, and set their endowment shock for non-index assets $v_2 = 0.0107$. These give us a fund flow of roughly 13% which is in-line with industry research cited by Greenwood and Sammon (2024). Finally, we set the intensity of state transitions $\lambda = 0.2$, which targets the conditional probability of index inclusion or deletion. λ is computed as the number of stocks included (or excluded) in a given year, divided by the number of stocks within the band on watch for inclusion (or exclusion). We conduct sensitivity analysis with respect to λ and γ.

Using the baseline parameters, we simulate the distribution of investors' aggregate wealth. In the baseline, the mean wealth of investors is 4.75. In our model, the index inclusion/deletion changes hedgers' relative endowment shock across the affected assets, giving rise to trading needs by the hedgers. As expected for such a local shock, the implied mean incremental return is very small, $\Delta\mathbb{E}_t(R_1) \approx 0.01\%$. One of the main points of this calibration is that even if identification relies on a demand shock implying tiny effects on expected returns, the implied mismeasurement of demand slope due to general equilibrium effects is not diminishing.

E.2.1 Results and slope analysis

Tables E.3-E.4 show the results of our baseline calibration and its slope to various parameters. As before, in each table the colored column contains our benchmark parametrization, while the other columns show how our effects vary.

Table E.3 illustrates the effect of varying relative-risk-aversion of investors. Table 2 concentrates on the effect of the persistence of the demand shift. Table E.4 highlights the effect of state dependence: how our measures change for smaller than mean wealth of investors.[28]

The results are remarkably similar to our index inclusion experiment even in a quantitative sense. Dynamic Slope is somewhat smaller than myopic slope. Observed slope is a fraction

[28] In Tables E.3 and 2, for each column we recalibrate the parameters u^ξ to ensure that the implied expected return shocks remain constant at $\Delta\mathbb{E}_t(R) = 0.01\%$.

Risk Aversion (γ)	0.9	1.5	2.0	3.0	3.5	log
Myopic Slope	33.13	17.45	12.52	7.81	6.55	27.20
Dynamic Slope	33.13	17.45	12.52	7.81	6.55	27.20
Observed Slope	8.55	5.96	4.80	3.42	2.99	7.51
Observed Slope/Dynamic Slope	26%	34%	38%	44%	46%	28%
Implied Price Elasticity	719	569	500	399	357	648
Dynamic minus Observed (DmO)	24.58	11.49	7.72	4.39	3.55	19.68
Shock to risk/DmO	98%	107%	111%	116%	118%	100%
Diff Hedging Brownian/DmO	2%	-7%	-11%	-16%	-18%	0%
Diff Hedging Reversal/DmO	0%	0%	0%	0%	0%	0%

Table E.3: Comparative statics with respect to risk aversion coefficient γ

Investor Wealth (w)	1.00	2.50	3.50	4.25	4.75
Myopic Slope	3.20	5.11	6.66	7.81	8.51
Dynamic Slope	3.20	5.11	6.66	7.81	8.51
Observed Slope	1.48	2.08	2.65	3.42	4.87
Observed Slope/Dynamic Slope	46%	41%	40%	44%	57%
Implied Price Elasticity	79	218	322	399	443
Dynamic minus Observed (DmO)	1.71	3.03	4.01	4.39	3.64
Shock to risk/DmO	109%	114%	115%	116%	118%
Diff Hedging Brownian/DmO	-9%	-14%	-15%	-16%	-18%
Diff Hedging Reversal/DmO	0%	0%	0%	0%	0%

Table E.4: Comparative statics with respect to investor wealth w

of its theoretical counterpart (amounting only to 26% when risk-aversion is small). The bulk of the effect is due to the enodgenous change in the covariance matrix, but the GE amplified intertemporal effect remaining also significant.

In this calibration, we also present how our effects depend on the state of the economy, i.e. the level of aggregate wealth of investors, w_t on Table E.4. This illustrates that all our objects of interest, myopic slope, dynamic slope and observed slope are highly state dependent.

It is also useful to observe that the hedging effect against the reversal is the only one which is significantly smaller in index inclusion than in our factor-elasticity calibration. It is so, because this is the only one which diminishes with the size of the shock.

Under the baseline parameters, the index-inclusion calibration yields an implied price elasticity à la Koijen and Yogo (2019) of 399, which is substantially lower than the commonly cited value of around 6250 from Petajisto (2009). For comparison, the empirical micro elasticity estimated by Greenwood and Sammon (2024) is approximately 2.7.

E.2.2 Translating Observed Slope to Elasticity

As explained earlier, the demand elasticity is defined as the percentage of capitalization traded divided by the percentage change in price, that is

$$\mathcal{E}(w_t;\lambda,u^s)=\frac{\frac{w^s\hat{y}_i^s}{S_iP_i^s}-\frac{w^n\hat{y}_i^n}{S_iP_i^n}}{-\frac{P_i^s-P_i^n}{P_i^n}}, \tag{E.250}$$

where we have suppressed time index for simplicity. $\hat{y}_i^n$ and P_i^n are the investor's portfolio share and the equilibrium price of asset i in the normal state. $\hat{y}_i^s$ and P_i^s are portfolio share and price in the shock state. The investor has wealth w^n in the normal state and w^s in the shock state. Notice that (E.250) can be rewritten as

$$\mathcal{E}(w_t;\lambda,u^s)=\frac{1}{y_i^n}\frac{[w^n\,\mathrm{diag}(P^n)^{-1}y^n]_i}{S_i}\frac{\frac{P_i^nw^s}{P_i^sw^n}\hat{y}_i^s-\hat{y}_i^n}{\frac{P_i^n-P_i^s}{P_i^n}}. \tag{E.251}$$

The middle term on the right-hand side is the fraction of asset i's capitalization held by investors. The last term on the right-hand side is at the same order of magnitude as the observed slope. The equation thus suggests that the demand elasticity is inversely related to investors' portfolio share of asset i.

The implied demand elasticity $\mathcal{E}(w_t;\lambda,u^s)$ depends on the shock persistence $1/\lambda$. For our comparative statics, we fix the incremental expected return $\Delta\mathbb{E}_t(R)\equiv\mu_{Rt}^s-\mu_{Rt}^n+\lambda\,\mathrm{diag}(P_t^s)^{-1}(P_t^n-P_t^s)$ at the level implied in the baseline. Fixing $\Delta\mathbb{E}_t(R)$, the change in asset prices $\mathrm{diag}(P_t^s)^{-1}(P_t^n-P_t^s)$ is higher when λ is low, that is when the shock is more persistent. In other words, transitory shocks tend to have smaller price impact. Consequently, when $1/\lambda$ is small, the denominator in (E.250) is also small, magnifying the implied demand elasticity. Notice that our observed slope $\mathcal{O}(w_t;\lambda,u^s)$ adjusts for this horizon effect due to shock persistence.