Local Efficiency and Cross-Sectional Predictability: The Information Aggregation Wedge

Daniel Andrei*

May 20, 2026

Abstract

Local efficiency need not imply cross-sectional efficiency. I study economies where each firm is priced rationally under local information and public cross-sectional signals. Unless aggregation is perfect, dispersed learning about a common state leaves residual Bayesian filter errors: an information aggregation wedge. Observable characteristics cannot predict returns in a firm's time series, but by proxying for this wedge, they predict cross-sectional returns, neither as risk factors nor behavioral mistakes. This single wedge explains why anomalies have a strong factor structure, why firms migrate across value and growth portfolios, why value operates primarily within industries, why anomaly returns concentrate on news days, and why the Security Market Line appears flat. Characteristic arbitrage cannot make prices cross-sectionally efficient; only perfect aggregation can—but perfect aggregation destroys the dispersed learning that generates predictability.

JEL Classification Codes: G12, G14, D83
Keywords: Information aggregation, Market efficiency, Cross-sectional predictability, Bayesian learning

---

I gratefully acknowledge financial support from McGill University.

*McGill University, Desautels Faculty of Management, 1001 Sherbrooke Street West, Office 549, Montréal, Québec H3A 1G5, Canada daniel.andrei@mcgill.ca, danielandrei.info.

1 Introduction

Financial markets display substantial local efficiency: an individual stock's price reflects the information held by the investors who trade it.[1] In no contradiction to the previous sentence, this paper argues that markets can display substantial cross-sectional *in*efficiency: the cross-section of firm prices and fundamentals reveals information absent from any individual pricing filtration.

How can a market be efficient stock by stock, yet inefficient across the cross-section? The answer traces to Hayek (1945): information is dispersed across local price-setters, and no individual holds what the cross-section as a whole reveals. Consequently, local efficiency does not imply cross-sectional efficiency.

I study this distinction in a model where local efficiency holds by construction. The economy features many firms, each priced by investors who specialize in local fundamentals. Every firm's profitability loads on a common aggregate state that no single agent observes directly. Investors and managers learn about this state through Bayesian updating of their local information. Because the state is common, the cross-section of valuations and fundamentals encodes information no local filtration contains. I vary the degree to which this information is extracted and made public. At one extreme, firms are priced under purely local information. At the other, public aggregation reveals the common state, and local efficiency becomes cross-sectional efficiency. Between these limits, Hayek's mechanism operates: dispersed information is only partially aggregated, leaving a residual *information aggregation wedge*. I characterize asset pricing through this wedge.

To isolate this mechanism, the baseline economy is risk-neutral. Each firm is priced to earn exactly the risk-free rate under its local filtration; there is no risk compensation, and no local trading rule can predict excess returns. Valuation, however, requires a local belief about the common state. Unless public aggregation is perfect, this belief diverges from the

[1] I define "local efficiency" as efficiency with respect to a stock's own pricing filtration. Accounting evidence shows prices anticipate fundamentals (Ball and Brown, 1968; Beaver, 1968; Collins, Kothari, and Rayburn, 1987); microstructure research measures how informed trading drives price discovery (Hasbrouck, 1991; Easley, Kiefer, O'Hara, and Paperman, 1996); and firm-specific price variance captures decentralized information that guides corporate investment (Roll, 1988; Morck, Yeung, and Yu, 2000; Vuolteenaho, 2002; Chen, Goldstein, and Jiang, 2007).

truth. The information aggregation wedge is defined as this gap: the difference between the firm's local belief and the true macroeconomic state. It is an unobservable local forecast error that Bayesian learning eventually corrects. If a firm is priced under a belief below the true state, its incoming cash flows systematically exceed local expectations. As Bayesian updating incorporates these surprises, the local belief converges toward the truth and the firm's valuation rises. This error correction generates a capital gain invisible locally but predictable to an observer of the cross-section.

This predictability emerges only in the cross-section. For a single firm, predicting returns requires knowing the unobservable true state to measure the wedge. The cross-section removes this obstacle. Because the true state is common to all firms, it mathematically cancels out in cross-sectional deviations. To rank firms by their filter errors, one need only rank them by their relative local beliefs. The absolute truth is no longer necessary: dispersion *becomes* information. The joint configuration of firms therefore reveals the relative mispricing invisible to local price-setters—relying only on the latent beliefs firm by firm. While these latent beliefs escape direct observation, firm characteristics serve as their empirical proxies.

A high book-to-market ratio, conservative investment, and poor past returns all identify a firm priced under a low relative belief. Each characteristic captures a different empirical trace of the same underlying gap. The "factor zoo" (Cochrane, 2011) is therefore not a collection of independent risks or behavioral mistakes, but a family of projections of one information aggregation wedge onto observable coordinates.[2]

The information aggregation wedge is not a behavioral bias. Under incomplete information, rational Bayesian updating generates estimation errors that correct as new data arrives. Because local pricing is optimal, these errors can be exploited only in the cross-section. A trading rule buying firms with relatively low local beliefs and shorting those with relatively high local beliefs earns predictable returns by riding this mechanical convergence. Therefore,

[2]To build intuition, consider a "trip-let" (Hofstadter, 1979): a single three-dimensional solid whose projections along three mutually perpendicular axes form distinct letters. An observer restricted to studying these two-dimensional shadows might conclude they are observing three fundamentally different objects. In this model, the information aggregation wedge is the trip-let, and the distinct anomalies are its shadows. The factor zoo proliferates because the econometrician observes only the lower-dimensional projections, not the single latent object casting them.

local efficiency does not imply cross-sectional efficiency.

This predictability cannot be eliminated by trading away a single anomaly. Characteristic arbitrage targets observable projections, not the underlying informational gap. Arbitrageurs can compress these premia, but unless their trading perfectly aggregates all dispersed information—a self-defeating outcome that destroys the information the market relies upon (Grossman and Stiglitz, 1980)—the residual wedge survives and predictability rotates into other projections. Anomalies persist, or learning collapses. There is no third option.

Incomplete aggregation yields immediate implications for empirical asset pricing. Characteristics do not predict returns in the time series (Welch and Goyal, 2008)—consistent with local efficiency—but do in the cross-section (Fama and French, 1992). Because these characteristics project the same latent wedge, the cross-section of expected returns is low-dimensional (Kozak, Nagel, and Santosh, 2020). Since local expected returns equal the risk-free rate, no firm is a permanent high-return type. Return predictability and style migration (Fama and French, 2007) are therefore inseparable: firms move across value and growth portfolios as learning incorporates new information. The model provides the structural microfoundation for the interaction between value and profitability (Novy-Marx, 2013): controlling for profitability strengthens value strategies because the two characteristics jointly isolate the informational wedge. This mechanism extends to industry-relative valuations (Cohen, Polk, and Vuolteenaho, 2003; Kim, Kim, and Lee, 2025): controlling for the industry strips away fundamental variation, exposing firms' relative beliefs. Ultimately, each of these empirical phenomena is the footprint of decentralized Bayesian learning and imperfect information aggregation.

The information aggregation wedge governs returns across time and risk exposures. Discrete public signals collapse the latent wedge, explaining why anomaly returns concentrate on news days (Engelberg, McLean, and Pontiff, 2018). Under aggregate risk pricing, an unmeasured local-learning premium flattens the empirical Security Market Line. The gap between the information sets of pricers and the econometrician therefore establishes a duality: when estimating betas from realized aggregate returns, the econometrician observes *less* information than the market, mismeasuring betas (Andrei, Cujean, and Wilson, 2023);

when aggregating the cross-section for characteristic sorts, the econometrician observes *more* information than any individual local price-setter, uncovering anomalies.

This paper contributes to the literature on market efficiency and the origins of cross-sectional anomalies. Traditional limits to arbitrage rely on institutional frictions, funding constraints, or behavioral noise traders (Shleifer and Vishny, 1997). I isolate an informational limit driven by decentralized Bayesian learning. This boundary complements the statistical limits of Da, Nagel, and Xiu (2024), who show high-dimensional noise prevents arbitrageurs from fully extracting weak pricing errors. The paper formalizes a cross-sectional counterpart to Samuelson's dictum that individual stocks are micro-efficient even if the aggregate market is macro-inefficient (Jung and Shiller, 2005). While Samuelson alters the level of asset aggregation, I alter the level of information aggregation, contrasting local efficiency under firm-specific filtrations with cross-sectional inefficiency under the pooled filtration. In both frameworks, efficiency is not preserved across boundaries of aggregation.

Section 2 formalizes the information aggregation wedge. Section 3 illustrates the mechanism in a linear-Gaussian economy, and Section 4 explores its empirical implications. Section 5 introduces industry factors, discrete news, and aggregate risk pricing. Section 6 concludes.

2 Local and Cross-Sectional Efficiency

This section isolates the role of cross-sectional aggregation. I begin with a one-firm economy where the manager and investors price the firm using the same information set. I then replicate this local pricing problem across many firms. The pricing rule for each firm remains unchanged; the only difference is the information revealed by observing the entire cross-section simultaneously. This distinction separates local efficiency from cross-sectional efficiency.

2.1 One-Firm Economy

Consider first a one-firm economy. I retain the firm index n to keep notation consistent when introducing the cross-section later. The firm's operating profitability per unit of installed capital, $Z_{n,t}$, evolves as

$$dZ_{n,t} = (\mu_0 - \kappa_Z Z_{n,t} + \beta_n X_t)dt + \sigma_Z dW^Z_{n,t},$$

where $\mu_0 > 0$ sets the long-run level, $\kappa_Z > 0$ governs mean reversion, and X_t is the aggregate state. The constant $\beta_n > 0$ measures the firm's macroeconomic exposure, and the Brownian motion $W^Z_{n,t}$ represents idiosyncratic profitability risk.

The aggregate state follows a mean-reverting process:

$$dX_t = -\kappa_X X_t dt + \sigma_X dW^X_t.$$

The shocks W^X_t and $W^Z_{n,t}$ are mutually independent. The manager and investors know the parameters $(\mu_0, \kappa_Z, \beta_n, \sigma_Z, \kappa_X, \sigma_X)$ but do not observe X_t. They rely instead on the same local information: the profitability history $\{Z_{n,u}\}_{u\leq t}$ and a local signal about the aggregate state,

$$ds_{n,t} = X_t dt + \frac{1}{\sqrt{A}} dW^s_{n,t},$$

where $A > 0$ is the signal precision and $W^s_{n,t}$ is independent of all other shocks. Consequently, they share the local filtration,

$$\mathcal{F}_{n,t} = \sigma\{Z_{n,u}, s_{n,u} : u \leq t\}.$$

The firm is priced under this common filtration.

Installed capital generates operating cash flows

$$dC_{n,t} = Z_{n,t} K_{n,t} dt,$$

and accumulates according to

$$dK_{n,t} = (I_{n,t} - \delta K_{n,t})dt, \tag{1}$$

where $\delta > 0$ is the depreciation rate and $i_{n,t} \equiv I_{n,t}/K_{n,t}$ is the gross investment rate. Investment is irreversible ($i_{n,t} \geq 0$). The per-unit cost of investment is

$$\Psi(i, Z) = i + \frac{\theta_n}{2}(i - \delta)^2 + \frac{\phi_n}{2}(i - Z)^2_+,$$

where $\theta_n > 0$ governs convex adjustment costs around replacement investment, and $\phi_n > 0$ governs the shadow cost of external finance. These frictions enrich the firm's investment problem. Because locally efficient pricing holds regardless of them, they confirm that any return predictability must arise from information rather than technology. Furthermore, these firm-specific parameters introduce structural heterogeneity, which will later prevent the cross-section of observable characteristics from perfectly revealing the aggregate state.

The manager and investors are risk-neutral and discount cash flows at the constant rate $r > 0$. Given the local filtration $\mathcal{F}_{n,t}$, firm value is

$$V(K_{n,t}, Z_{n,t}, \hat{X}_{n,t}) = \max_{\{i_{n,u}\}_{u \geq t}} \mathbb{E}_t^{\mathcal{F}_n} \left[\int_t^{\infty} e^{-r(u-t)} K_{n,u} \left(Z_{n,u} - \Psi(i_{n,u}, Z_{n,u}) \right) du \right],$$

subject to the capital accumulation law (1), the non-negativity constraint $i_{n,u} \geq 0$, and the filtering problem for X_t. Here, the local belief shared by the manager and investors is $\hat{X}_{n,t} \equiv \mathbb{E}[X_t | \mathcal{F}_{n,t}]$. Because operating cash flows and capital accumulation scale linearly with installed capital, the model satisfies the conditions of Hayashi (1982). Firm value is therefore homogeneous of degree one: $V(K, Z, \hat{X}) = Kv(Z, \hat{X})$.

Let $i_{n,t}^*$ denote the optimal investment policy. The cum-dividend equity return is

$$dR_{n,t}^E = \frac{dV_{n,t} + K_{n,t}(Z_{n,t} - \Psi(i_{n,t}^*, Z_{n,t}))dt}{V_{n,t}}.$$

Because firm value is the expected present value of cash flows discounted at rate r, local pricing admits no arbitrage. The expected sum of capital gains and cash flows must exactly equal the time value of money, yielding the HJB equation:

$$\mathbb{E}_t^{\mathcal{F}_n}[dV_{n,t} + K_{n,t}(Z_{n,t} - \Psi(i_{n,t}^*, Z_{n,t}))dt] = rV_{n,t}dt.$$

Dividing by $V_{n,t}$ gives the local-efficiency condition:[3]

$$\mathbb{E}_t^{\mathcal{F}_n}[dR_{n,t}^E] = rdt. \tag{2}$$

Equation (2) implies that no trading rule based on the local filtration earns an expected

[3]Equation (2) is the continuous-time, risk-neutral counterpart of the producer Euler equation in Cochrane (1991), with the Hayashi conditions ($V = Kv$) ensuring that the investment return and the equity return coincide state by state (Restoy and Rockinger, 1994).

excess return. For any bounded position $\omega_{n,t}$ measurable with respect to $\mathcal{F}_{n,t}$,

$$\mathbb{E}_t^{\mathcal{F}_n}[\omega_{n,t}(dR_{n,t}^E - rdt)] = 0.$$

This applies to rules based on $Z_{n,t}$, $\hat{X}_{n,t}$, Tobin's q, book-to-market, or the firm's return history. Learning affects valuation and investment, but it creates no predictability under the pricing filtration.

2.2 Multi-Firm Economy

I now extend this local pricing problem to a multi-firm economy. Following Merton (1987), I consider a market with informational segmentation, where each firm is priced by investors who specialize in its local fundamentals. I replicate the previous structure across a large but finite set of firms indexed by $n \in \{1, 2, \ldots, N\}$. Firm n has its own profitability process, its own private signal, and its own local filtration $\mathcal{F}_{n,t} = \sigma\{Z_{n,u}, s_{n,u} : u \leq t\}$.

This extension changes the information problem. In the one-firm economy, the firm's valuation was not part of a cross-section. With many firms, it is. Since each firm is priced under its own local filtration, the collection of firm valuations—the *cross-section*—becomes an additional source of information about the common fundamental X_t. This additional source of information is absent in the one-firm economy. I capture this cross-sectional information about X_t in an aggregation filtration $\mathcal{A}_t$, available at date t to all investors and firm managers. Each firm n is therefore priced under the augmented filtration

$$\mathcal{H}_{n,t}(\mathcal{A}) \equiv \mathcal{F}_{n,t} \vee \mathcal{A}_t.$$

I study the content of $\mathcal{A}_t$ in the next subsection. For now, the only point is that the pricing problem remains local once common public information enters each firm's information set. The manager and investors of firm n share the augmented filtration $\mathcal{H}_{n,t}(\mathcal{A})$, make decisions under it, and discount cash flows at rate r. The one-firm pricing of Section 2.1 therefore applies firm by firm. For each n,

$$\mathbb{E}_t^{\mathcal{H}_n(\mathcal{A})}[dR_{n,t}^E] = rdt. \tag{3}$$

Thus the multi-firm economy is locally efficient asset by asset. No trading rule based on firm

n's pricing filtration earns an expected excess return. The only novelty is informational: with many firms, the cross-section itself reveals information absent in the one-firm economy.

2.3 Information Aggregation and Cross-Sectional Efficiency

The previous subsection introduced $\mathcal{A}_t$ as the public information extracted from the cross-section. Its role is to locate the economy between two informational benchmarks. The first is strict segmentation: $\mathcal{A}_t$ is the trivial sigma-field, $\mathcal{A}_t = \{\varnothing, \Omega\}$ up to null sets, and firm n is priced only under its own local filtration $\mathcal{F}_{n,t}$. The second is full aggregation. Hayek (1945) argued that information is inherently dispersed—across histories, signals, and local knowledge—and cannot be centralized. Full aggregation is therefore a theoretical benchmark rather than an equilibrium outcome: if this dispersed information could nonetheless be collected and processed, the relevant filtration would be

$$\mathcal{G}_t \equiv \bigvee_{m=1}^{N} \mathcal{F}_{m,t}.$$

No single price-setter observes $\mathcal{G}_t$. Doing so would require the profitability histories, private signals, and local information of every firm in the economy. The pooled filtration is not a realistic individual information set; it is what a fictitious observer who could collect and process the entire cross-section would know. In this perfect-aggregation benchmark, the dispersed information reveals the common state:

$$\mathbb{E}[X_t|\mathcal{G}_t] = X_t.$$

The aggregation filtration $\mathcal{A}_t$ is the interior case: public information contains more than the trivial filtration but less than the full pooled filtration $\mathcal{G}_t$.

This notation makes the aggregation problem precise. The quantity of interest is the residual error imperfect aggregation leaves relative to the full-pooling benchmark $\mathcal{G}_t$. I define the *information aggregation wedge* for firm n as

$$\mathcal{W}_{n,t} \equiv \mathbb{E}[X_t|\mathcal{G}_t] - \mathbb{E}[X_t|\mathcal{H}_{n,t}(\mathcal{A})]. \tag{4}$$

Because the pooled filtration perfectly reveals the common state, this simplifies to

$$\mathcal{W}_{n,t} = X_t - \hat{X}^{\mathcal{H}}_{n,t},$$

where $\hat{X}^{\mathcal{H}}_{n,t} \equiv \mathbb{E}[X_t|\mathcal{H}_{n,t}(\mathcal{A})]$ is the augmented local belief. The wedge is therefore the residual filter error left after the market incorporates the public aggregate information $\mathcal{A}_t$.

With no public aggregation, $\mathcal{H}_{n,t}(\mathcal{A}) = \mathcal{F}_{n,t}$, so the wedge is the ordinary local filter error: the gap between the true state and firm n's local belief. With perfect aggregation, $\mathcal{H}_{n,t}(\mathcal{A}) = \mathcal{G}_t$, so the wedge is zero for every firm. The relevant case is the interior one,

$$\mathcal{F}_{n,t} \subsetneq \mathcal{H}_{n,t}(\mathcal{A}) \subsetneq \mathcal{G}_t.$$

Here, the relevant information is dispersed across the economy—no price-setter holds it in full. Public aggregation moves part of it into the common domain but leaves the rest local. The wedge measures what stays behind. Cross-sectional efficiency requires this wedge to vanish; local efficiency does not.

2.4 Information Aggregation as the Origin of Anomalies

The wedge in Equation (4) is a purely informational object, defined before any statement about returns. It is not a violation of local efficiency: firms are priced under the augmented filtration $\mathcal{H}_{n,t}(\mathcal{A})$, and no rule measurable with respect to that filtration earns an expected excess return. The wedge matters only when returns are evaluated under the pooled filtration $\mathcal{G}_t$. Under $\mathcal{G}_t$, the true state X_t is known, and so is the residual filter error $\mathcal{W}_{n,t}$. Because cash flows and belief revisions are generated by X_t, this error predicts how local beliefs and valuations will adjust. The following proposition maps this informational wedge into expected returns.

Proposition 1 (Cross-Sectional Predictability)**.** *Under the pooled filtration $\mathcal{G}_t$, the expected return contains a predictable component proportional to the information aggregation wedge:*

$$\frac{\mathbb{E}^{\mathcal{G}}_t[dR^E_{n,t}]}{dt} = r + \Gamma_{n,t}(\mathcal{A})\mathcal{W}_{n,t}, \tag{5}$$

where the loading $\Gamma_{n,t}(\mathcal{A}) > 0$ is given by

$$\Gamma_{n,t}(\mathcal{A}) = \frac{\beta_n v_Z}{v_n} + \frac{\lambda^{\mathcal{H}}_{n,t} v_{\hat{X}}}{v_n}. \tag{6}$$

Here, v_n is normalized firm value, $v_Z \equiv \frac{\partial v_n}{\partial Z} > 0$ and $v_{\hat{X}} \equiv \frac{\partial v_n}{\partial \hat{X}} > 0$ are its partial derivatives with respect to fundamental profitability and the augmented belief. The parameter $\lambda^{\mathcal{H}}_{n,t} > 0$ is the optimal filter gain under $\mathcal{H}_{n,t}(\mathcal{A})$, which dictates the sensitivity of the market's belief revisions to all local and public innovations.

Return predictability appears under the pooled filtration $\mathcal{G}_t$, the information set of an observer who sees the whole cross-section. Equation (3) shows that under the pricing filtration $\mathcal{H}_{n,t}(\mathcal{A})$, expected excess returns remain zero.

The loading $\Gamma_{n,t}(\mathcal{A})$ has two parts. The first is a cash-flow effect: if $\mathcal{W}_{n,t} > 0$, the local belief understates the true state, so profitability grows faster under $\mathcal{G}_t$ than local price-setters expect, contributing $\beta_n v_Z/v_n$ to expected returns. The second is a revaluation effect: the same filter error makes future signals better than the local belief expects, so Bayesian updating pushes $\hat{X}^{\mathcal{H}}_{n,t}$ upward on average; since firm value rises with the belief, this contributes $\lambda^{\mathcal{H}}_{n,t} v_{\hat{X}}/v_n$. Both terms are positive. Expected excess returns under $\mathcal{G}_t$ therefore increase with the information aggregation wedge.

Can an econometrician—or an arbitrageur—exploit this predictability? The answer seems no: the wedge contains the true state X_t, which is unobservable. But the cross-section removes this obstacle. Since X_t is common to all firms, it cancels in cross-sectional deviations:

$$\mathcal{W}_{n,t} - \mathbb{E}_{CS}[\mathcal{W}_{n,t}] = -\left(\hat{X}^{\mathcal{H}}_{n,t} - \mathbb{E}_{CS}[\hat{X}^{\mathcal{H}}_{n,t}]\right), \tag{7}$$

where $\mathbb{E}_{CS}[\cdot]$ denotes the cross-sectional average across firms at date t. The truth X_t need not be observed. To rank firms by their relative wedges, one need only rank them by their relative local beliefs. This is where observable characteristics enter.

Observable characteristics matter when they proxy for relative local beliefs. A characteristic correlated with $\hat{X}^{\mathcal{H}}_{n,t} - \mathbb{E}_{CS}[\hat{X}^{\mathcal{H}}_{n,t}]$ is also correlated with the residual information aggregation wedge, and therefore predicts returns under $\mathcal{G}_t$. The next two assumptions impose minimal monotonicity conditions on the latent beliefs: they must move valuations, and they must command predictable returns.

Assumption 1 (Valuation monotonicity in beliefs)**.** *For any firm n with $\beta_n > 0$, holding fixed all known payoff-relevant observable states and parameters, firm value is strictly increasing in the local belief $B_{n,t} \equiv \hat{X}^{\mathcal{H}}_{n,t}$. That is, for any $B' > B$,*

$$V(K_{n,t}, Z_{n,t}, B') > V(K_{n,t}, Z_{n,t}, B).$$

Assumption 1 follows from primitives: a higher local belief raises expected future profitability and therefore the discounted value of optimally chosen cash flows. The unobservable belief thus *leaves a mark* on firm value even after controlling for all observable states and parameters.

Assumption 2 (Return monotonicity in beliefs)**.** *For any firm n with $\beta_n > 0$, holding fixed all known payoff-relevant observable states and parameters, the predictable return component $\Gamma_{n,t}(\mathcal{A})\mathcal{W}_{n,t}$ is strictly decreasing in the local belief $B_{n,t} \equiv \hat{X}^{\mathcal{H}}_{n,t}$. That is, for any $B' > B$,*

$$\Gamma_{n,t}(\mathcal{A}; B')(X_t - B') < \Gamma_{n,t}(\mathcal{A}; B)(X_t - B).$$

Assumption 2 refines Proposition 1, which shows that the predictable return component is $\Gamma_{n,t}(\mathcal{A})(X_t - B_{n,t})$ with $\Gamma_{n,t}(\mathcal{A}) > 0$. It requires that variation in $\Gamma_{n,t}(\mathcal{A}; B)$ does not reverse the strictly decreasing ordering the wedge $X_t - B_{n,t}$ induces.

These two monotonicity conditions are enough to turn latent belief variation into an observable predictive characteristic.

Proposition 2 (Predictive Residual Valuation Characteristic)**.** *Fix date t and sort firms into $M < \infty$ observable-characteristic portfolios using all known payoff-relevant observable states and parameters. Let $S_{n,t} \in \{1, \ldots, M\}$ denote the portfolio containing firm n at date t. Suppose Assumptions 1 and 2 hold. If, for at least one portfolio s^*, there is positive within-portfolio variation in $B_{n,t}$, then the observable residual valuation characteristic*

$$\chi^V_{n,t} \equiv \mathbb{1}_{\{S_{n,t}=s^*\}} \left(V_{n,t} - \bar{V}_t(s^*)\right),$$

where $\bar{V}_t(s^*) \equiv \frac{1}{N_t(s^*)} \sum_{m:S_{m,t}=s^*} V_{m,t}$*, predicts returns cross-sectionally. In particular,*

$$\mathrm{Cov}_{CS}\left(\chi^V_{n,t}, \frac{\mathbb{E}^{\mathcal{G}}_t[dR^E_{n,t}]}{dt} - r\right) < 0.$$

Thus high residual valuations predict lower returns.

Proposition 2 gives the empirical content of the wedge. If beliefs move valuations, an unobserved belief error cannot remain hidden: it changes the price investors are willing to pay. If the same belief commands predictable returns, the valuation residual becomes a return predictor. Within a portfolio controlling for observable states and parameters, high residual valuations identify firms priced under high local beliefs—firms with low relative wedges and low expected returns under $\mathcal{G}_t$. Low residual valuations identify the reverse: high relative wedges and high expected returns.

This provides a foundation for valuation characteristics, especially book-to-market. A high book-to-market ratio is a low market valuation relative to book value. In the model, low valuation reflects a low local belief, holding fixed the relevant observable states. By Equation (7), a low relative belief is a high relative wedge. The value premium is therefore an observable projection of the information aggregation wedge.

The point is not limited to book-to-market. Any characteristic that covaries with relative beliefs also covaries with the residual wedge and predicts returns under $\mathcal{G}_t$. Profitability, investment, reversal, size, and beta are not independent primitives; they are different observable coordinates of the same latent object. The factor zoo is then a family of projections of one information aggregation wedge, not a list of unrelated risks or mistakes. The origin of anomalies is incomplete cross-sectional aggregation: locally efficient prices need not exhaust the information contained in the cross-section.

2.5 Equilibrium Persistence of Characteristic Predictability

Proposition 2 turns the information aggregation wedge into a characteristic premium, raising an arbitrage objection: why do investors not trade until the premium disappears? Exploiting this predictability is in fact *risky* arbitrage. While this objection becomes less stark with aggregate risk pricing (Section 5.3), a frictionless risk-neutral economy poses a stricter challenge: a predictable return is an arbitrage unless trading changes the information set that generates the prediction. The question is therefore whether characteristic trading eliminates the wedge itself, or only one observable projection of it.

The answer turns on the gap between the predictable-return vector and its observable projections. Let

$$\mu_{n,t}(\mathcal{A}) \equiv \frac{\mathbb{E}_t^{\mathcal{G}}[dR_{n,t}^E]}{dt} - r = \Gamma_{n,t}(\mathcal{A})\mathcal{W}_{n,t}$$

denote the predictable excess-return component under the pooled filtration. A characteristic strategy based on an observable characteristic $\chi_{n,t}$ does not trade $\mu_{n,t}(\mathcal{A})$ directly; it earns a premium through the cross-sectional projection $\mathrm{Cov}_{CS}(\chi_{n,t}, \mu_{n,t}(\mathcal{A}))$. Driving this covariance to zero imposes one orthogonality condition on $\mu_{n,t}(\mathcal{A})$; it does not make $\mu_{n,t}(\mathcal{A}) = 0$ pointwise, nor does it eliminate the wedge. Trading a single characteristic can therefore eliminate one anomaly while leaving the predictable-return vector active in other coordinates.

This does not exhaust the objection. Arbitrageurs need not trade one characteristic at a time; they may trade multiple observable characteristics until their projections jointly span $\mu_{n,t}(\mathcal{A})$. The relevant question is therefore not whether one named anomaly can be arbitraged away, but whether all observable predictability can vanish while the information aggregation wedge remains non-zero.

Corollary 2.1 (Equilibrium Persistence of Predictability)**.** *An equilibrium in which trading has eliminated all cross-sectional predictability requires perfect public aggregation ($\mathcal{A}_t = \mathcal{G}_t$).*

The logic proceeds by contradiction. Suppose an equilibrium exists in which the wedge is non-zero, yet no observable characteristic predicts returns. Proposition 2 rules out this possibility: as long as the wedge is non-zero, the within-portfolio valuation residual forms an observable characteristic that covaries with expected returns. To eliminate all characteristic predictability, trading must therefore do more than compress observable projections. It must eliminate the wedge itself, so that $\mathcal{W}_{n,t} = 0$ for every firm. This requires public information to reveal the common state, which is the perfect-aggregation benchmark $\mathcal{A}_t = \mathcal{G}_t$.

Yet perfect aggregation is structurally self-undermining. If $\mathcal{A}_t$ perfectly reveals X_t, private signals lose their marginal value for learning the common state, and local price-setters place no weight on them. But those dispersed local signals are the very inputs from which public aggregation is constructed. If local valuations no longer move with local information, the cross-section ceases to encode dispersed learning, leaving the market with nothing to aggregate. This is the logic of Grossman and Stiglitz (1980) applied to the cross-section: full revelation destroys its own informational substrate. Arbitrageurs can rotate predictability across observable coordinates, but they cannot make the cross-section efficient without achieving perfect aggregation. Anomalies survive because learning must.

This impossibility result does not leave the analysis without a subject—it *locates* it. By identifying perfect aggregation as the only way out, it says that the relevant economics lives in the interior, where $\mathcal{A}_t$ is informative but incomplete. The question shifts from "Can the wedge be eliminated?" to "How large is the wedge, and what does it imply for the cross-section?" That is precisely what the next section studies.

3 A Linear-Gaussian Economy

The previous section developed the information aggregation wedge under general filtrations. I now study it in a linear-Gaussian economy, treating the aggregation filtration $\mathcal{A}_t$ as given. Observable firm characteristics—valuations, investment rates, profitability—depend on the local belief $\hat{X}^{\mathcal{H}}_{n,t}$ but also on firm-specific parameters β_n, θ_n, and ϕ_n, so the cross-section is not an invertible map to X_t: no aggregation technology recovers the common state exactly. Rather than model the aggregation problem directly, I parameterize what it extracts. I assume that conditioning on $\mathcal{A}_t$ gives a Gaussian posterior for X_t.

3.1 The Filtering Problem

The aggregation filtration $\mathcal{A}_t$ may be complex, but the object of learning is the scalar state X_t. For tractability, I impose a Gaussian posterior:

$$X_t \mid \mathcal{A}_t \sim N(\hat{X}^{\mathcal{A}}_t, \nu^{\mathcal{A}}_t).$$

Here $\hat{X}_t^{\mathcal{A}} \equiv \mathbb{E}[X_t \mid \mathcal{A}_t]$ is the aggregate posterior mean, and $\nu_t^{\mathcal{A}} \equiv \mathrm{Var}(X_t \mid \mathcal{A}_t)$ is the aggregate posterior variance.

Given this Gaussian posterior, let $\tau_{\mathcal{A}} \geq 0$ denote the aggregate precision, defined by the quadratic variation of aggregate belief revisions:

$$\frac{d\langle \hat{X}^{\mathcal{A}} \rangle_t}{dt} = \tau_{\mathcal{A}}(\nu_t^{\mathcal{A}})^2. \tag{8}$$

The parameter $\tau_{\mathcal{A}}$ measures the precision of information extracted from the cross-section. It nests the two limits from the previous section: $\tau_{\mathcal{A}} = 0$ gives strict informational segmentation, while $\tau_{\mathcal{A}} \to \infty$ gives the perfect-aggregation limit in which $\mathcal{A}_t$ reveals the common state.

Under the pricing filtration $\mathcal{H}_{n,t} = \mathcal{F}_{n,t} \vee \mathcal{A}_t$, firm n's manager and investors combine the aggregate posterior with its own local information: the cash-flow history $Z_{n,t}$ and the private signal $s_{n,t}$. I assume that the innovations in $\mathcal{A}_t$ are conditionally orthogonal to the innovations in the two local channels. Under this assumption, precisions add, and the total precision in firm n's filtering problem is

$$\Lambda_n = \frac{\beta_n^2}{\sigma_Z^2} + A + \tau_{\mathcal{A}}. \tag{9}$$

The next lemma gives the steady-state filter implied by these assumptions.

Lemma 1 (Firm-level filtering)**.** *Under the linear-Gaussian specification and the orthogonality assumption above, firm n's posterior variance under $\mathcal{H}_{n,t}$ converges to*

$$\bar{\nu}_n = \frac{\sigma_X^2}{\kappa_X + \sqrt{\kappa_X^2 + \Lambda_n \sigma_X^2}},$$

where Λ_n is defined in (9). The posterior mean evolves as

$$d\hat{X}_{n,t}^{\mathcal{H}} = -\kappa_X \hat{X}_{n,t}^{\mathcal{H}} dt + \frac{\beta_n \bar{\nu}_n}{\sigma_Z} d\widetilde{W}_{n,t}^Z + \bar{\nu}_n \sqrt{A} d\widetilde{W}_{n,t}^s + \bar{\nu}_n \sqrt{\tau_{\mathcal{A}}} d\widetilde{W}_{n,t}^{\mathcal{A}}, \tag{10}$$

where $\widetilde{W}_{n,t}^Z$, $\widetilde{W}_{n,t}^s$, and $\widetilde{W}_{n,t}^{\mathcal{A}}$ are the orthogonal innovation Brownian motions associated with the cash-flow channel, the private-signal channel, and the aggregation channel under $\mathcal{H}_{n,t}$.

The posterior variance $\bar{\nu}_n$ decreases with each component of Λ_n. Each innovation enters the posterior mean with gain equal to $\bar{\nu}_n$ times the square root of that channel's precision. Thus the belief inherits the mean reversion of the common state and updates through the three information channels available to firm n.

3.2 Firm Value, Investment Policy, and Expected Returns

Under $\mathcal{H}_{n,t}$, the manager and investors share the posterior belief $\hat{X}^{\mathcal{H}}_{n,t}$. The manager chooses investment to maximize firm value, and investors discount the resulting cash flows at rate r. Linear homogeneity in capital implies $V = Kv$, so normalized value depends only on profitability Z and the belief $\hat{X}^{\mathcal{H}}$. At the steady-state posterior variance $\bar{\nu}_n$, v solves

$$rv = \max_{i \geq 0} \Big\{ Z - \Psi(i, Z) + v(i - \delta) + (\mu_0 - \kappa_Z Z + \beta_n \hat{X}^{\mathcal{H}}) v_Z + \frac{1}{2}\sigma_Z^2 v_{ZZ} \\ - \kappa_X \hat{X}^{\mathcal{H}} v_{\hat{X}} + \frac{1}{2}\bar{\sigma}_n^2 v_{\hat{X}\hat{X}} + \beta_n \bar{\nu}_n v_{Z\hat{X}} \Big\},$$

where the variance of the belief diffusion is

$$\bar{\sigma}_n^2 \equiv \sigma_X^2 - 2\kappa_X \bar{\nu}_n,$$

as implied by the steady-state Riccati equation. The cross-term $\beta_n \bar{\nu}_n v_{Z\hat{X}}$ appears because profitability and the belief share the cash-flow innovation, so their instantaneous covariation is $\beta_n \bar{\nu}_n$. Thus the HJB depends on information precision through $\bar{\nu}_n$ and $\bar{\sigma}_n^2$, both pinned down by Λ_n. As $\tau_{\mathcal{A}} \to \infty$, $\bar{\nu}_n \to 0$ and $\bar{\sigma}_n^2 \to \sigma_X^2$, so the HJB converges to the perfect-information benchmark. The next proposition gives the value decomposition and the investment rule.

Proposition 3 (Value decomposition and optimal investment)**.** *The normalized value function decomposes as*

$$v(Z, \hat{X}^{\mathcal{H}}) = v^{AIP}(Z, \hat{X}^{\mathcal{H}}) + g(Z, \hat{X}^{\mathcal{H}}), \tag{11}$$

where the assets-in-place value is affine,

$$v^{AIP}(Z, \hat{X}^{\mathcal{H}}) = A_0 + A_Z Z + A_X \hat{X}^{\mathcal{H}},$$

with constants A_0, A_Z, and A_X given in the appendix. For $Z \geq 0$, the optimal investment *rule is*

$$i^*(v, Z) = \begin{cases} 0 & v \le 1 - \theta_n \delta, \\ \delta + \dfrac{v - 1}{\theta_n} & 1 - \theta_n \delta < v \le 1 + \theta_n (Z - \delta), \\ \dfrac{v - 1 + \theta_n \delta + \phi_n Z}{\theta_n + \phi_n} & v > 1 + \theta_n (Z - \delta). \end{cases} \tag{12}$$

The option-to-scale component $g(Z, \hat{X}^{\mathcal{H}})$ solves a PDE given in the appendix.

The decomposition separates assets in place from the option to scale. The investment rule has two kinks. At $v = 1 - \theta_n \delta$, irreversibility binds and the firm does not invest. At $v = 1 + \theta_n(Z - \delta)$, investment exceeds current profitability and the financing penalty ϕ_n binds. Information affects investment only indirectly through v: the belief $\hat{X}^{\mathcal{H}}$ does not enter i^* directly.

3.3 Numerical Implementation

I solve the model on a two-dimensional grid for $(Z, \hat{X}^{\mathcal{H}})$. The baseline parameters reflect standard values from the macro-finance literature. I set the discount rate to $r = 0.05$ and depreciation to $\delta = 0.10$ (Bolton, Chen, and Wang, 2011; Andrei, Mann, and Moyen, 2019). Profitability parameters $\mu_0 = 0.12$ and $\kappa_Z = 0.80$ yield a 15% steady state (Eberly, Rebelo, and Vincent, 2012). Volatilities are $\sigma_Z = 0.05$ and $\sigma_X = 0.04$, the aggregate mean reversion is $\kappa_X = 0.50$, and macroeconomic exposure is $\beta_n = 0.70$ (Gomes, 2001; Zhang, 2005). The investment frictions—adjustment cost $\theta_n = 2$ and external-finance penalty $\phi_n = 25$—discipline growth and equity issuance (Whited, 1992; Altinkilic and Hansen, 2000; Hennessy and Whited, 2007). Finally, the signal precisions $A = 10^4$ and $\tau_{\mathcal{A}} = 100$ fix the baseline of imperfect information aggregation.

The numerical method is described in Appendix B. I factor out the affine value of assets in place and solve the PDE for the option-to-scale component $g(Z, \hat{X}^{\mathcal{H}})$ by finite differences. The investment rule has two kinks, from irreversibility and external finance, so I use an implicit false-transient scheme rather than a global polynomial approximation.

Figure 1 plots the resulting investment and value policies. Higher beliefs raise normalized value at each level of profitability. The flat portion of the investment schedule is the

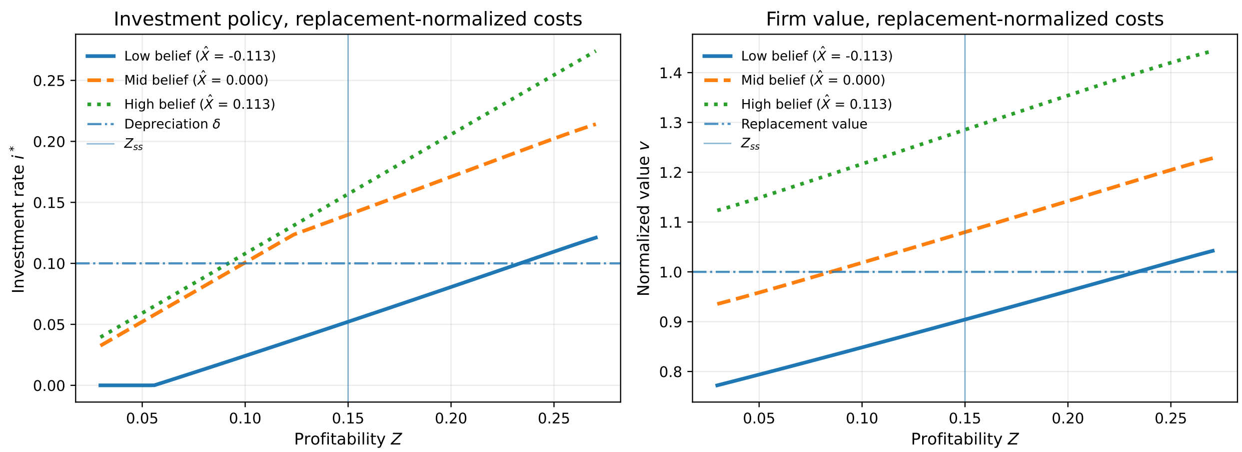

Figure 1: Investment and Firm Value. The left panel plots the optimal investment rate i^*, and the right panel plots normalized firm value v, both as functions of current profitability Z. The three curves correspond to low, medium, and high levels of the posterior belief $\hat{X}^{\mathcal{H}}$.

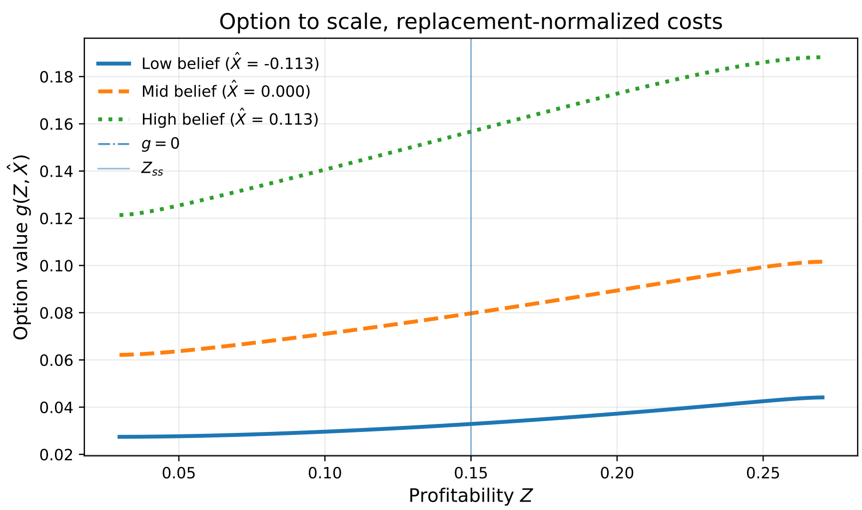

Figure 2: Option to Scale. The figure plots the option-to-scale component $g(Z, \hat{X}^{\mathcal{H}})$ as a function of current profitability Z. The three curves correspond to low, medium, and high levels of the posterior belief $\hat{X}^{\mathcal{H}}$.

irreversibility region; the kink at higher profitability marks the point where desired investment exceeds current profitability and the external-finance penalty binds.

Figure 2 plots the option-to-scale component. The option value rises with profitability and with the posterior belief. The figure shows the channel through which learning affects investment: beliefs do not enter the investment rule directly, but they move the continuation value that determines when the firm invests.

4 Implications for Asset Pricing

I now study the model's asset-pricing implications using simulated data. Because the value function is numerical, simulations offer a practical way to evaluate the model's mechanics. They also generate panels for standard predictive regressions. I simulate firms daily, update beliefs through the filtering equations, and sample the data quarterly. Each quarter, I record valuations, profitability, investment, and beliefs, along with one-quarter-ahead excess returns.

I then ask when firm characteristics predict returns. In this risk-neutral economy, no individual firm earns an unconditional excess return in long samples.[4] Characteristic premia arise only through dynamic rebalancing: the sorting rule selects firms when their current state loads on the information wedge. The simulation therefore tests whether characteristics predict returns across states, not whether firms are permanent high-return types. Table 1 reports results across four information structures over 400 quarters. Column (1) shows time-series averages across 500 independent firms. Columns (2) through (4) present Fama-MacBeth estimates for 500-firm panels under idiosyncratic fundamentals, incomplete learning about a common fundamental, and perfect information aggregation.

Column (1) establishes the time-series benchmark for an economy with a common state X_t, using cross-sectional averages of 500 independent firm-by-firm regressions. Within an individual firm's history, regressing future excess returns on observable characteristics (Panel A) or the local belief (Panel B) yields no predictability. This lack of explanatory power confirms local efficiency. Predictability emerges only in Panel C, which regresses returns on the *oracle filter error*—the gap between the local belief and the unobservable true state. Because the true state is hidden, the filter error drives realized returns without creating an exploitable time-series premium.

Column (2) shifts to cross-sectional estimation but removes the common state, simulating 500 firms with idiosyncratic fundamentals $X_{n,t}$. This isolates the effect of the cross-section from that of a shared latent state. Because each firm learns about a different target, the cross-section contains unrelated filter errors rather than noisy signals about a single object.

[4] This follows from the law of iterated expectations. Local efficiency implies $\mathbb{E}_t^{\mathcal{H}_n(\mathcal{A})}[dR^E_{n,t} - rdt] = 0$ for any firm n. Hence $\mathbb{E}[dR^E_{n,t} - rdt] = \mathbb{E}[\mathbb{E}_t^{\mathcal{H}_n(\mathcal{A})}[dR^E_{n,t} - rdt]] = 0$.

Table 1: Return Predictability Across Information Structures. This table reports coefficient estimates from regressions of one-quarter-ahead excess returns on characteristics (Panel A), the local belief (Panel B), and the oracle filter error (Panel C). Column (1) reports cross-sectional averages of 500 firm-by-firm time-series regressions. Columns (2) through (4) report Fama-MacBeth estimates for 500-firm panels under idiosyncratic fundamentals (2), a common fundamental with incomplete information aggregation (3), and a common fundamental with perfect learning (4). Newey-West HAC t-statistics are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Independent Variable	(1) 500 Firms Time-Series	(2) 500 Firms Idiosyncratic X	(3) 500 Firms Common X	(4) 500 Firms Perfect Learning
Panel A: Physical Characteristics				
Intercept	-0.0512	-0.0002	-0.7467***	-0.1553***
	(-0.64)	(-0.06)	(-166.32)	(-5.50)
Book/Market	0.0503	-0.0006	0.6972***	0.1465***
	(0.73)	(-0.17)	(102.03)	(5.61)
Profitability (Z)	0.0407	0.0081*	0.7156***	0.1533***
	(0.41)	(1.66)	(106.11)	(5.73)
Investment	0.0099	-0.0162	0.0334	0.0090
	(0.01)	(-0.59)	(1.26)	(0.41)
Average R^2	0.009	0.006	0.112	0.007
Panel B: Local Information State				
Intercept	0.0019	-0.0001	0.0072	–
	(0.90)	(-1.27)	(1.59)	–
Local Belief ($\hat{X}_{n,t}$)	-0.0750	0.0048	-1.0097***	–
	(-1.16)	(1.61)	(-165.83)	–
Average R^2	0.004	0.002	0.110	–
Panel C: Oracle Filter Error				
Intercept	0.0009	-0.0000	0.0008	–
	(0.51)	(-0.41)	(1.37)	–
Oracle Filter Error	-1.0076***	-1.0069***	-1.0097***	–
	(-10.08)	(-202.30)	(-165.83)	–
Average R^2	0.197	0.197	0.110	–
Firm-Quarters	199,000	199,000	199,000	199,000
Time Periods	398	398	398	398

This placebo confirms the logic of equation (7): demeaning beliefs identifies relative filter errors *only* when the latent state has a common component. Consequently, observable characteristics (Panel A) and local beliefs (Panel B) generate no meaningful predictability; although some slopes are statistically significant, their explanatory power is negligible. Strong predictability appears only when regressing on the oracle filter error (Panel C). Yet, because these true idiosyncratic states are hidden, this mechanical premium cannot be exploited.

Column (3) replaces idiosyncratic fundamentals with a common state X_t. This shared target transforms the cross-section. Meaningful predictability emerges across observable characteristics (Panel A) because they proxy for relative local beliefs, whose direct predictive power is confirmed in Panel B. The cross-sectional slope on the local belief $\hat{X}_{n,t}$ (Panel B) perfectly matches the slope on the oracle filter error, $\hat{X}_{n,t} - X_t$ (Panel C). As equation (7) shows, cross-sectional demeaning mathematically cancels the unobservable truth X_t. Comparing the time series in Column (1) to the cross-section in Column (3) establishes the core mechanism: what is noise locally operates as a signal cross-sectionally. When firms share a common state, the equivalence of Panels B and C demonstrates that the absolute truth is no longer required to identify relative filter errors. Truth cancels in the mean; dispersion remains the signal.

Column (4) retains the common state X_t but removes the aggregation friction by setting public aggregation precision to infinity ($\tau_{\mathcal{A}} \to \infty$). This shuts down the information aggregation wedge. Under perfect aggregation, investors observe X_t without error. Because the local belief perfectly matches the true state across all firms, the filter error vanishes. Cross-sectional regressions on these metrics are therefore unidentified, and Panels B and C are omitted. For the physical characteristics in Panel A, although some coefficients are statistically significant (discretization artifacts), explanatory power collapses. The slopes therefore provide no economically meaningful evidence of return predictability. Once aggregation is perfect, neither physical characteristics nor local beliefs explain cross-sectional returns.

The predictability in Column (3) is a premium on a transient state, not on a permanent firm type. Consistent with local efficiency, no individual firm earns an unconditional excess return. The premium arises only through dynamic rebalancing: a firm enters a high-expected-return portfolio when a low local belief depresses its valuation and raises its book-to-market ratio. As Bayesian updating incorporates new cash flows and the local belief converges toward the common state, the valuation recovers and the firm exits the portfolio. This mechanical correction links characteristic predictability directly to the style migration documented by Fama and French (2007), who show that realized returns drive firms to move across value and growth portfolios over time.

Table 2 reports the 1-year transition matrix for book-to-market quintiles, mapping each firm from its origin portfolio at quarter t to its destination at $t+4$. Extreme value and growth portfolios are persistent but not permanent. After one year, nearly two-thirds of value and growth firms leave their initial portfolios. Migration is largely gradual: firms move first toward adjacent portfolios and then toward the middle quintiles.

Table 2: 1-Year Style Transition Matrix. This table reports the percentage of firms migrating across book-to-market quintiles. Quintiles are formed cross-sectionally each quarter.

	Destination Portfolio at $t+4$				
Origin at t	Growth	Q2	Neutral	Q4	Value
Growth	36.4%	24.5%	18.4%	13.0%	7.6%
Q2	24.1%	23.1%	21.1%	18.6%	13.1%
Neutral	18.4%	21.1%	21.4%	20.9%	18.2%
Q4	13.4%	18.2%	20.9%	23.3%	24.3%
Value	7.7%	13.1%	18.2%	24.2%	36.8%

These transition rates align with the empirical evidence of Fama and French (2007) and Broussard, Mikkonen, and Puttonen (2016), who document extensive style migration. The model provides a rational foundation for this movement: firms lack fixed value or growth types. A firm enters an extreme book-to-market portfolio when its local belief diverges from the common state. As Bayesian learning corrects this error, the firm's valuation reverts toward the cross-sectional average, driving it into a different style portfolio. Style migration is simply the observable footprint of decentralized Bayesian learning. Because local efficiency precludes unconditional excess returns, characteristic premia exist strictly through portfolio turnover. Return predictability and style migration are therefore inseparable: they are two empirical reflections of imperfect information aggregation.

This mechanism also explains why controlling for profitability strengthens value strategies (Novy-Marx, 2013). In the model, this complementarity is not a separate source of risk compensation. Instead, book-to-market and profitability serve as two observable projections of the same latent belief error. Table 3 tests this implication using Fama-MacBeth regressions on value, profitability, and both jointly.

Table 3: Value and Profitability as Complementary Signals. This table reports Fama-MacBeth regressions of one-quarter-ahead excess returns on book-to-market and profitability. Newey-West HAC t-statistics are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Independent Variable	(1) Value	(2) Profitability	(3) Joint
Intercept	-0.1131***	0.0059***	-0.7449***
	(-55.94)	(4.01)	(-184.49)
Book/Market	0.1260***		0.6960***
	(43.95)		(102.23)
Profitability (Z)		-0.0283***	0.7194***
		(-13.45)	(117.23)
Average R^2	0.025	0.003	0.110
Firm-Quarters	199,500	199,500	199,500
Time Periods	399	399	399

Evaluated alone, book-to-market predicts higher returns, while profitability yields a negative sign. Evaluated jointly, both coefficients become positive and explanatory power increases. The sign reversal is informative. Because profitable firms tend to have low book-to-market ratios, an unconditional profitability regression mixes the profitability signal with a growth signal. Conditional on book-to-market, however, profitability isolates firms with strong cash flows relative to their valuation. These firms must be priced under a low relative belief, earning higher expected returns as learning corrects the error.

The baseline regressions in Table 1 show that investment lacks predictive power when evaluated alongside valuation and profitability. Yet, an extensive empirical literature documents a robust negative investment premium (Titman, Wei, and Xie, 2004; Cooper, Gulen, and Schill, 2008; Fama and French, 2015; Hou, Xue, and Zhang, 2015). Table 4 resolves this tension by unpacking the investment signal sequentially. It evaluates investment in isolation, tests its interaction with profitability, and ultimately pits these physical fundamentals against the local belief.

Investment is a valid but easily subsumed proxy for local beliefs. Evaluated alone, high

Table 4: The Investment Factor: Information vs. Fundamentals. This table reports Fama-MacBeth regressions of one-quarter-ahead excess returns on investment, profitability, and the local information state. Newey-West HAC t-statistics are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Independent Variable	(1) Investment	(2) With Z	(3) With $\hat{X}$	(4) Joint
Intercept	0.0127***	0.0151***	0.0066	0.0067
	(9.62)	(11.72)	(1.42)	(1.43)
Investment	-0.3427***	-1.1071***	0.0155	-0.0142
	(-17.73)	(-15.72)	(1.29)	(-0.49)
Profitability (Z)		0.1514***		0.0075
		(14.81)		(1.30)
$\hat{X}$ (Information)			-1.0119***	-1.0121***
			(-175.03)	(-159.95)
Average R^2	0.007	0.014	0.112	0.114
Firm-Quarters	199,000	199,000	199,000	199,000
Time Periods	398	398	398	398

investment predicts lower returns (Column 1). Because firms invest when marginal value is high, high investment identifies firms priced under a high relative belief. As Bayesian updating corrects this positive belief error, these firms earn lower realized returns. Controlling for profitability strengthens this negative signal (Column 2). When the local belief enters the regression, however, the predictive power of both investment and profitability vanishes (Columns 3 and 4). These physical characteristics provide no independent pricing information; they predict cross-sectional returns only because they capture the transient information aggregation wedge.

This logic also explains the reversal anomaly. Table 5 tests whether value and reversal are separate patterns or two measures of the same belief error. In the model, a low valuation indicates a low local belief. A low past return captures this exact error: the firm was priced under a low belief, and its price has not fully adjusted as learning proceeds. Both characteristics should therefore predict returns when evaluated alone, but neither should retain power once the local belief is observed.

Evaluated alone, book-to-market predicts higher returns (Column 1), while past returns predict lower returns (Column 2). When the local belief enters the regression, however,

Table 5: Value, Reversal, and the Local Belief. This table reports Fama-MacBeth regressions of one-quarter-ahead excess returns on book-to-market, past-quarter returns, and the local belief. Newey-West HAC t-statistics are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Independent Variable	(1) Value	(2) Reversal	(3) With $\hat{X}$
Intercept	-0.1132***	0.0034**	0.0091**
	(-56.58)	(2.28)	(2.01)
Book/Market	0.1260***		-0.0019
	(43.90)		(-0.87)
Past Return (1-Quarter)		-0.1358***	0.0009
		(-60.17)	(0.35)
$\hat{X}$ (Information)			-1.0140***
			(-167.90)
Average R^2	0.025	0.020	0.114
Firm-Quarters	199,000	199,000	199,000
Time Periods	398	398	398

the predictive power of both value and reversal vanishes (Column 3). Value and reversal represent no independent sources of predictability. Like investment and profitability, they predict returns only because they track the latent belief error.

These simulations confirm that local efficiency does not imply cross-sectional efficiency. In this risk-neutral economy, value, profitability, investment, and reversal predict returns. Yet they are not independent anomalies. They are observable projections of a single Bayesian filter error, rendering the cross-section of expected returns low-dimensional (Kozak et al., 2020). This error persists because locally efficient prices imperfectly aggregate the common state. Cross-sectional efficiency therefore demands more than asset-by-asset efficiency—it requires the market to perfectly aggregate dispersed local information. What the time series hides, the cross-section reveals.

5 Extensions

The baseline model isolates the information aggregation wedge in a simplified setting. This section extends the framework in three directions: multiple industry factors, discrete public news, and aggregate risk pricing.

5.1 Multiple Common Factors and Within-Industry Value

Proposition 2 implies that residual valuation predicts returns within any observable comparison group that leaves room for latent belief variation. Industries serve as natural groups: they absorb structural differences in accounting, capital intensity, and growth opportunities while exposing firm-level belief errors. The empirical literature reflects this distinction. Standard sorts measure book-to-market across the entire cross-section (Fama and French, 1992, 1993). In contrast, Cohen et al. (2003) demonstrate that the value premium operates primarily as an intra-industry effect, and Kim et al. (2025) attribute its recent aggregate weakness to persistent cross-industry valuation differences. I therefore extend the baseline model to incorporate multiple industry-level factors, testing whether the information aggregation wedge generates stronger predictability within industries than in the aggregate cross-section.

Suppose the economy contains I industries, each with a latent state $X_{i,t}$. For firm n in industry i, profitability $Z_{n,t}$ loads on $X_{i,t}$ rather than a single economy-wide state. The firm combines its local information and public industry signals to form the augmented belief $\hat{X}^{\mathcal{H}}_{n,t}$.

Because the pooled filtration $\mathcal{G}_t$ reveals the true industry state $X_{i,t}$, Proposition 1 applies industry by industry. For firm n in industry i,

$$\frac{\mathbb{E}^{\mathcal{G}}_t[dR^E_{n,t}]}{dt} - r = \Gamma_{n,t}(\mathcal{A})(X_{i,t} - \hat{X}^{\mathcal{H}}_{n,t}).$$

An industry-specific information wedge therefore drives the predictable return.

An aggregate value sort ranks firms by raw book-to-market across all industries. This measure conflates two components: the firm-level belief error, $X_{i,t} - \hat{X}^{\mathcal{H}}_{n,t}$, and true variation in industry states. Firms in low-state industries naturally exhibit higher book-to-market ratios—even if their local beliefs contain no relative error. The aggregate sort therefore loads on cross-industry fundamentals as well as within-industry belief errors.

To isolate the latent belief error, the model demands the empirical adjustment proposed by Kim et al. (2025). Their book-share-to-market-share ratio (BS/MS) nets the industry's aggregate log book-to-market ratio out of the firm's raw ratio. This transformation mechanically strips away cross-industry fundamentals, isolating the firm's relative valuation within its industry.

This ratio precisely isolates the information wedge. Let $\mathbb{E}_i[\cdot]$ denote the cross-sectional average within industry i. Because $X_{i,t}$ is common to all firms in the industry,

$$(X_{i,t} - \hat{X}^{\mathcal{H}}_{n,t}) - \mathbb{E}_i[X_{i,t} - \hat{X}^{\mathcal{H}}_{n,t}] = -(\hat{X}^{\mathcal{H}}_{n,t} - \mathbb{E}_i[\hat{X}^{\mathcal{H}}_{n,t}]).$$

Within-industry demeaning removes the true industry state, exposing the firm's relative belief. An industry-adjusted value characteristic therefore provides a cleaner proxy for the belief-correction channel than raw aggregate book-to-market.

I simulate this multi-factor economy using 2,000 firms across ten industries, each exposed to its own latent state. The local filtering and valuation equations remain identical to the baseline model. Table 6 reports Fama-MacBeth regressions of one-quarter-ahead excess returns on aggregate and industry-adjusted book-to-market ratios. Column (1) runs a standard aggregate regression using raw $\log(B/M)$. Column (2) substitutes $\log(BS/MS)$ to measure value relative to industry peers. Column (3) follows the residual horse race of Kim et al. (2025), adding the component of $\log(BS/MS)$ that is orthogonal to raw $\log(B/M)$.

The results mirror the empirical evidence of Cohen et al. (2003) and Kim et al. (2025). Raw book-to-market fails to predict returns in the aggregate cross-section, whereas industry-adjusted measures capture the predictive content. The within-industry measure predicts returns independently, and its residualized component remains significant even after controlling for raw book-to-market. The simulated economy therefore reproduces the core empirical regularity: value is not an aggregate book-to-market effect. Its predictability derives from ranking firms relative to peers sharing the same industry-level state.[5]

5.2 Anomalies and News

Although the baseline model updates beliefs continuously, its core mechanism extends directly to discrete news days. Suppose firm-level public signals arrive at Poisson times τ. At each

[5]This mirrors the principal-component rotation in Kozak et al. (2020): industry adjustment strips shared variation to isolate a sparse latent wedge. I use 2,000 firms here so industry averages precisely capture the latent industry states.

Table 6: Aggregate and Within-Industry Value Predictability. This table reports Fama-MacBeth regressions of one-quarter-ahead excess returns on raw book-to-market, industry-adjusted book-to-market, and a residual horse race following Kim et al. (2025). Newey-West HAC t-statistics are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Independent Variable	(1) Raw B/M	(2) BS/MS	(3) Residual Horse Race
Intercept	0.0001	-0.0001	0.0001
	(0.12)	(-0.24)	(0.13)
$\log(B/M)$	0.0034	–	0.0034
	(0.74)	–	(0.74)
$\log(BS/MS)$	–	0.1125***	–
	–	(119.73)	–
Residualized BS/MS	–	–	0.1615***
	–	–	(30.78)
Average R^2	0.016	0.019	0.048
Firm-Quarters	798,000	798,000	798,000
Time Periods	399	399	399

announcement, firm n releases the signal

$$Y_{n,\tau} = X_\tau + \varepsilon_{n,\tau},$$

where $\varepsilon_{n,\tau}$ is independent noise with variance $\sigma^2_{\varepsilon,n}$. Let $\hat{X}^{\mathcal{H}}_{n,\tau-}$ and $\nu_{n,\tau-}$ denote the pre-announcement belief and posterior variance. The Gaussian update is

$$\Delta\hat{X}^{\mathcal{H}}_{n,\tau} = G_{n,\tau}(Y_{n,\tau} - \hat{X}^{\mathcal{H}}_{n,\tau-}), \qquad G_{n,\tau} = \frac{\nu_{n,\tau-}}{\nu_{n,\tau-} + \sigma^2_{\varepsilon,n}}.$$

The announcement pulls the local belief toward the public signal. Because the signal centers on the true state, the expected revision under the pooled filtration is proportional to the pre-announcement wedge.

Lemma 2 (News-Day Belief Revisions)**.** *Suppose a public signal $Y_{n,\tau} = X_\tau + \varepsilon_{n,\tau}$ arrives for firm n at time τ, with $\varepsilon_{n,\tau}$ independent of $\mathcal{G}_{\tau-}$ and mean zero. Then*

$$\mathbb{E}^{\mathcal{G}}_{\tau-}[\Delta\hat{X}^{\mathcal{H}}_{n,\tau}] = G_{n,\tau}(X_\tau - \hat{X}^{\mathcal{H}}_{n,\tau-}).$$

If firm value is differentiable in the local belief, the expected announcement return satisfies

$$\mathbb{E}^{\mathcal{G}}_{\tau-}\left[\frac{\Delta V_{n,\tau}}{V_{n,\tau-}}\right] \approx \frac{v_{\hat{X},n,\tau-}}{v_{n,\tau-}} G_{n,\tau}(X_\tau - \hat{X}^{\mathcal{H}}_{n,\tau-}).$$

Substituting $Y_{n,\tau} = X_\tau + \varepsilon_{n,\tau}$ into the announcement update yields

$$\Delta\hat{X}^{\mathcal{H}}_{n,\tau} = G_{n,\tau}(X_\tau - \hat{X}^{\mathcal{H}}_{n,\tau-}) + G_{n,\tau}\varepsilon_{n,\tau}.$$

Conditioning on $\mathcal{G}_{\tau-}$ eliminates the mean-zero noise term, leaving the expected belief revision proportional to the pre-announcement information wedge. A first-order expansion of firm value around the pre-announcement belief generates the return expression:

$$\Delta V_{n,\tau} \approx V_{\hat{X},n,\tau-}\Delta\hat{X}^{\mathcal{H}}_{n,\tau}.$$

Lemma 2 provides the mechanical link to the empirical findings of Engelberg et al. (2018). Between announcements, observable characteristics predict returns by proxying for the latent wedge, $X_t - \hat{X}^{\mathcal{H}}_{n,t}$, generating a gradual price drift as local beliefs update. On news days, the arrival of a discrete public signal forces an immediate Bayesian revision. Because this discrete jump is mathematically proportional to the pre-announcement wedge, the model structurally requires characteristic predictability to concentrate on news days. Any characteristic that sorts firms by their latent filter error inherently sorts them by the magnitude of their expected news-day revaluation. In this framework, the underlying forecast error is not a behavioral mistake; it is the rational consequence of an agent processing strictly local information. Public news partially aggregates the cross-section, directly correcting the identical wedge that anomalies proxy for between announcements.

5.3 Aggregate Risk Pricing and the SML

The baseline economy is strictly risk-neutral, isolating the information wedge but precluding risk compensation. Introducing aggregate risk pricing into the segmented-information structure addresses two questions: Does the predictive wedge survive standard risk compensation? And does this friction explain the flat empirical security market line?

A full general-equilibrium derivation of risk premia would require tracking wealth across all investor groups. This would make firm value depend on the distribution of investors' wealth and would break the scale independence of the investment problem. I therefore use a no-arbitrage approach. I specify a global stochastic discount factor on the pooled filtration $\mathcal{G}_t$ with constant prices of risk, then project it onto each local pricing filtration $\mathcal{H}_{n,t}(\mathcal{A})$. This uses the conditioning-information logic of Hansen and Richard (1987): asset-pricing restrictions are defined relative to the information set under which payoffs are priced.

The pooled-information stochastic discount factor M_t is strictly positive, adapted to $\mathcal{G}_t$, and satisfies

$$\frac{dM_t}{M_t} = -rdt - \lambda_X dW_t^X,$$

where λ_X is the price of latent macroeconomic risk. Firm n, however, is priced under $\mathcal{H}_{n,t}(\mathcal{A})$, not under $\mathcal{G}_t$. Its local pricing kernel is therefore the conditional projection

$$\widehat{M}_{n,t} \equiv \mathbb{E}[M_t \mid \mathcal{H}_{n,t}(\mathcal{A})].$$

Because $e^{rt}M_t$ is a $\mathcal{G}_t$-martingale and r is deterministic, $e^{rt}\widehat{M}_{n,t}$ is an $\mathcal{H}_{n,t}(\mathcal{A})$-martingale. By the filtering equation (10), the local belief is driven by the cash-flow, private-signal, and public-aggregation innovations. The martingale part of $\widehat{M}_{n,t}$ can therefore be written as a stochastic integral with respect to these three Brownian innovations, following the approach of Fujisaki, Kallianpur, and Kunita (1972).

Lemma 3 (Projected Local Stochastic Discount Factor)**.** *The projected local pricing kernel admits the representation*

$$\frac{d\widehat{M}_{n,t}}{\widehat{M}_{n,t}} = -rdt + \frac{\beta_n}{\sigma_Z}\gamma_n d\widetilde{W}_{n,t}^Z + \sqrt{A}\gamma_n d\widetilde{W}_{n,t}^s + \sqrt{\tau_{\mathcal{A}}}\gamma_n d\widetilde{W}_{n,t}^{\mathcal{A}}, \tag{13}$$

where γ_n is the steady-state difference between the local conditional mean of X_t under the global risk-neutral measure $\mathbb{Q}$ and under the physical measure $\mathbb{P}$, both conditioned on $\mathcal{H}_{n,t}(\mathcal{A})$:

$$\gamma_n = \frac{-\sigma_X \lambda_X}{\kappa_X + \bar{\nu}_n \Lambda_n}.$$

This projected kernel defines the local risk-neutral measure $\mathbb{Q}^n$, under which local investors *price the firm by discounting expected cash flows at the risk-free rate.*

As $\tau_{\mathcal{A}} \to \infty$, the local pricing kernel converges to the pooled-information pricing kernel. In this limit, $\gamma_n \to 0$, idiosyncratic risk prices vanish ($\frac{\beta_n}{\sigma_Z}\gamma_n \to 0$ and $\sqrt{A}\gamma_n \to 0$), and the aggregate risk price converges to the macroeconomic risk premium ($\sqrt{\tau_{\mathcal{A}}}\gamma_n \to -\lambda_X$). Away from this limit, learning innovations carry risk prices because they update beliefs about the priced fundamental X_t. The local risk prices are nonzero whenever $\gamma_n \neq 0$ and the corresponding information channel is active. Because the steady-state filtering variance fixes these risk prices as constants, the change of measure shifts the drifts of $Z_{n,t}$ and the physical belief $\hat{X}^{\mathbb{P}}_{n,t}$ by constants. Risk pricing therefore adds no new state variables to the local valuation problem.

Proposition 4 (Local Valuation under Risk Aversion)**.** *Under $\mathbb{Q}^n$, the local state evolves as*

$$dZ_{n,t} = \left(\mu_0 - \kappa_Z Z_{n,t} + \beta_n \hat{X}^{\mathcal{H}}_{n,t} + \beta_n \gamma_n\right) dt + \sigma_Z d\widetilde{W}^{Z,\mathbb{Q}}_{n,t},$$

$$d\hat{X}^{\mathcal{H}}_{n,t} = \left(-\kappa_X \hat{X}^{\mathcal{H}}_{n,t} + \bar{\nu}_n \Lambda_n(\mathcal{A})\gamma_n\right) dt + \frac{\beta_n \bar{\nu}_n}{\sigma_Z} d\widetilde{W}^{Z,\mathbb{Q}}_{n,t} + \bar{\nu}_n \sqrt{A} d\widetilde{W}^{s,\mathbb{Q}}_{n,t} + \bar{\nu}_n \sqrt{\tau_{\mathcal{A}}} d\widetilde{W}^{\mathcal{A},\mathbb{Q}}_{n,t}.$$

Firm value remains homogeneous in capital, $V = K v_n$, and the normalized value function solves

$$\begin{aligned} r v_n = \max_{i \geq 0} \Big\{ & Z - \Psi(i, Z) + (i - \delta) v_n + \left(\mu_0 - \kappa_Z Z + \beta_n \hat{X}^{\mathcal{H}} + \beta_n \gamma_n\right) v_Z \\ & + \left(-\kappa_X \hat{X}^{\mathcal{H}} + \bar{\nu}_n \Lambda_n(\mathcal{A})\gamma_n\right) v_{\hat{X}} + \frac{1}{2}\sigma_Z^2 v_{ZZ} + \beta_n \bar{\nu}_n v_{Z\hat{X}} + \frac{1}{2}\bar{\nu}_n^2 \Lambda_n(\mathcal{A}) v_{\hat{X}\hat{X}} \Big\}. \end{aligned}$$

The optimal investment rule retains the three-region form of Proposition 3, and the decomposition $v = v^{AIP} + g$ continues to hold. Risk pricing enters only through the constant drift adjustments $\beta_n \gamma_n$ and $\bar{\nu}_n \Lambda_n(\mathcal{A})\gamma_n$, so it does not add a new state variable to the local valuation problem.

With the value function characterized by Proposition 4, we can apply Itô's lemma to write the cum-dividend equity return as

$$dR^E_{n,t} = \mu^R_{n,t} dt + a_{n,Z,t} d\widetilde{W}^Z_{n,t} + a_{n,s,t} d\widetilde{W}^s_{n,t} + a_{n,\mathcal{A},t} d\widetilde{W}^{\mathcal{A}}_{n,t}, \tag{14}$$

where the return exposures are

$$a_{n,Z,t} = \sigma_Z \frac{v_Z}{v_n} + \frac{\beta_n \bar{\nu}_n}{\sigma_Z} \frac{v_{\hat{X}}}{v_n}, \qquad a_{n,s,t} = \bar{\nu}_n \sqrt{A} \frac{v_{\hat{X}}}{v_n}, \qquad a_{n,\mathcal{A},t} = \bar{\nu}_n \sqrt{\tau_{\mathcal{A}}} \frac{v_{\hat{X}}}{v_n}. \tag{15}$$

By standard no-arbitrage pricing, the local expected excess return equals the negative covariance of the return with the projected local pricing kernel. Using the kernel representation from Lemma 3, this yields:

$$\frac{\mathbb{E}_t^{\mathcal{H}_n(\mathcal{A})}[dR^E_{n,t}]}{dt} - r = -\frac{\beta_n}{\sigma_Z} \gamma_n a_{n,Z,t} - \sqrt{A} \gamma_n a_{n,s,t} - \sqrt{\tau_{\mathcal{A}}} \gamma_n a_{n,\mathcal{A},t}. \tag{16}$$

To study the empirical cross-section, I evaluate this locally efficient return under the pooled filtration $\mathcal{G}_t$. Conditioning on $\mathcal{G}_t$ reveals the true latent macroeconomic state X_t. Under this pooled perspective, the local learning innovations are no longer mean-zero; they acquire a conditional physical drift driven by the information aggregation wedge, $\mathcal{W}_{n,t} = X_t - \hat{X}^{\mathcal{H}}_{n,t}$. Taking the expectation of the local return diffusion (14) under $\mathcal{G}_t$ projects this predictable forecast error directly into the asset's expected return, yielding the central decomposition.

Proposition 5 (Return Decomposition)**.** *Under the pooled filtration $\mathcal{G}_t$, the expected equity return decomposes into a local risk premium and an information aggregation wedge:*

$$\frac{\mathbb{E}_t^{\mathcal{G}}[dR^E_{n,t}]}{dt} - r = \underbrace{-\frac{\beta_n}{\sigma_Z} \gamma_n a_{n,Z,t} - \sqrt{A} \gamma_n a_{n,s,t} - \sqrt{\tau_{\mathcal{A}}} \gamma_n a_{n,\mathcal{A},t}}_{\text{local risk premium}} + \underbrace{\Gamma_{n,t}(\mathcal{A}) \mathcal{W}_{n,t}}_{\text{wedge}},$$

where the predictability loading is

$$\Gamma_{n,t}(\mathcal{A}) = \beta_n \frac{v_Z}{v_n} + \bar{\nu}_n \Lambda_n(\mathcal{A}) \frac{v_{\hat{X}}}{v_n}.$$

Proposition 5 provides the risk-averse counterpart to Proposition 1. While the risk-neutral baseline drives expected excess returns under $\mathcal{G}_t$ solely through the wedge, aggregate risk pricing introduces an additive local risk premium. The wedge remains mathematically independent because it arises from conditioning on a broader information set, not from risk compensation.

The local pricing kernel prices firm n under $\mathcal{H}_{n,t}(\mathcal{A})$, while the econometrician summarizes realized returns with a single market beta. To evaluate the empirical security market line, I assume constant return loadings and constant market weights. Let $(a_{n,Z}, a_{n,s}, a_{n,\mathcal{A}})$ denote the fixed return loadings in (15), and ω_n the fixed market weight. Because the information wedge in Proposition 5 has a zero unconditional mean, firm n's unconditional average excess return equals its local risk premium. I decompose this premium into an unmeasured local-learning component and a public-aggregation component:

$$\pi_n^{loc} \equiv -\gamma_n \left(\frac{\beta_n}{\sigma_Z} a_{n,Z} + \sqrt{A} a_{n,s} \right), \qquad \pi_n^{agg} \equiv -\gamma_n \sqrt{\tau_{\mathcal{A}}} a_{n,\mathcal{A}}.$$

The first term compensates cash-flow and private-signal innovations, while the second compensates the public-aggregation innovation. In a large cross-section, the market portfolio diversifies away firm-specific innovations but preserves the public-aggregation innovation. Market beta therefore measures the public component and misses π_n^{loc}. The label "unmeasured" reflects the econometrician's information set. The market counterparts are $\bar{\pi}_M^{loc} \equiv \sum_m \omega_m \pi_m^{loc}$ and $\bar{\pi}_M^{agg} \equiv \sum_m \omega_m \pi_m^{agg}$.

Proposition 6 (The Empirical Security Market Line)**.** *Let $dR_{M,t}^E \equiv \sum_m \omega_m dR_{m,t}^E$ denote the market return and let $\mu_n^e \equiv \mathbb{E}^{\mathbb{P}}[dR_{n,t}^E/dt - r]$ denote firm n's unconditional average excess return. Under the constant-loading SML benchmark, define*

$$\bar{a}_{\mathcal{A}} \equiv \sum_m \omega_m a_{m,\mathcal{A}}.$$

If $\bar{a}_{\mathcal{A}} \neq 0$, firm n's empirical market beta is

$$\hat{\beta}_n \equiv \frac{\mathrm{Cov}(dR_{n,t}^E, dR_{M,t}^E)}{\mathrm{Var}(dR_{M,t}^E)} = \frac{a_{n,\mathcal{A}}}{\bar{a}_{\mathcal{A}}}.$$

Let $\alpha_n \equiv \mu_n^e - \hat{\beta}_n \bar{\mu}_M^e$ denote the unconditional CAPM alpha, where $\bar{\mu}_M^e \equiv \sum_m \omega_m \mu_m^e$. Then

$$\alpha_n = \underbrace{\bar{\pi}_M^{loc}(1 - \hat{\beta}_n)}_{\textit{flattening effect}} + \underbrace{\sqrt{\tau_{\mathcal{A}}} a_{n,\mathcal{A}}(\bar{\gamma}_M - \gamma_n)}_{\textit{heterogeneous-pricing effect}} + \underbrace{(\pi_n^{loc} - \bar{\pi}_M^{loc})}_{\textit{premium dispersion}}, \tag{17}$$

where $\bar{\gamma}_M \equiv \sum_m \omega_m \gamma_m \frac{a_{m,\mathcal{A}}}{\bar{a}_{\mathcal{A}}}$ is the exposure-weighted market average local risk price.

The first term in (17) isolates the flattening effect. Because the market portfolio diversifies away cash-flow and private-signal innovations, the empirical market beta fails to price these local learning shocks. Assuming a uniform local-learning premium ($\pi_n^{loc} = \bar{\pi}_M^{loc}$), this unmeasured compensation generates a CAPM alpha of $\bar{\pi}_M^{loc}(1 - \hat{\beta}_n)$. Whenever the average local premium is positive ($\bar{\pi}_M^{loc} > 0$), this mechanism structurally flattens the empirical security market line, producing positive alphas for low-beta stocks ($\hat{\beta}_n < 1$) and negative alphas for high-beta stocks ($\hat{\beta}_n > 1$).

The second term isolates the heterogeneous-pricing effect. Although the market portfolio spans the public-aggregation innovation, the empirical CAPM prices this shock using the market-average local risk price, $\bar{\gamma}_M$. Because the local pricing kernel requires γ_n, this divergence generates alpha whenever $\gamma_n \neq \bar{\gamma}_M$. If high-beta firms hold more precise local information about X_t, their γ_n is less negative than the market average. This informational asymmetry dictates that $\bar{\gamma}_M - \gamma_n < 0$ for high-beta firms and $\bar{\gamma}_M - \gamma_n > 0$ for low-beta firms, causing the heterogeneous-pricing effect to structurally reinforce the flattening effect.

The third term isolates premium dispersion. This component generates positive alpha when firm n's unmeasured local premium exceeds the market average. If π_n^{loc} decreases with market beta, this dispersion reinforces the flattening effect; conversely, if π_n^{loc} increases with beta, it offsets the flattening. While the first two mechanisms structurally flatten the empirical security market line, this final term captures residual cross-sectional variation in locally priced learning risk.

Imperfect aggregation generates the two terms in Proposition 6. As public information becomes fully precise, the local filtration converges to the pooled filtration. Cash-flow and private-signal innovations lose their risk prices because they cease to reveal residual uncertainty about X_t, while the public-aggregation innovation becomes the macroeconomic shock. The unmeasured local-learning premium therefore vanishes, and local public-risk prices converge to the common macroeconomic risk price. This full-information limit restores the empirical security market line.

Proposition 7 (The Full-Information CAPM)**.** *As public information becomes perfectly precise, $\tau_{\mathcal{A}} \to \infty$, the flattening effect, the heterogeneous-pricing effect, and local heterogeneity* *all vanish. Hence the unconditional CAPM alpha converges to zero for every firm:*

$$\lim_{\tau_{\mathcal{A}} \to \infty} \alpha_n = 0.$$

The empirical Security Market Line is restored in the full-information limit.

This mechanism is the dual of the informational friction in Andrei et al. (2023). When estimating the empirical security market line, the econometrician conditions down—just as in their framework—observing less information than the market, which mismeasures beta. Conversely, when sorting on observable characteristics, the econometrician conditions across. By aggregating the cross-section, the econometrician effectively conditions up, extracting more information than individual local price-setters possess and generating characteristic predictability. Pricer and econometrician hold different information, and this gap cuts both ways: conditioning down hides risk compensation in beta tests; conditioning across creates apparent alpha in characteristic sorts. In both cases, better public information closes the gap and moves the data back toward the CAPM.

6 Conclusion

Does local efficiency imply cross-sectional efficiency? This paper shows it does not. When information is dispersed, prices are locally efficient under firm-specific filtrations yet leave a residual forecast error under the pooled cross-section. Characteristic premia—value, profitability, and investment—are not independent risks or behavioral mistakes, but observable projections of a single latent information aggregation wedge.

The framework rationalizes the empirical decay of characteristic premia. Real-time tracking shows the expected returns of prominent anomalies diminish significantly around publication (Marrow and Nagel, 2024), mirroring the broad post-publication decline in predictability (McLean and Pontiff, 2016). The fading of the value premium and other anomalies is often interpreted as evidence of statistical illusions. In this framework, their disappearance is the natural outcome of aggregation. Academic publication and characteristic arbitrage act as public aggregation technologies. By extracting the latent state and compressing the information wedge, they mechanically erode the predictable premium.

Yet, if arbitrageurs compress these premia to zero, it does not mean the anomalies were illusory. It means the wedge has been incorporated into prices. Information exists as an emergent property of the entire cross-section, largely invisible to local agents. The factor zoo is therefore not a catalog of human errors or unobservable risks, but the empirical footprint of a decentralized system attempting to process a macroeconomic state too complex to be instantly aggregated. The predictable residuals we identify as anomalies mark the exact mathematical boundary of the market's carrying capacity for information. Modern finance, driven by AI and algorithmic processing, is building technologies to breach this boundary. This presents a paradox: achieving perfect public aggregation would eliminate the decentralized learning that incentivizes trade.

References

Altinkilic, O. and R. S. Hansen (2000). Are there economies of scale in underwriting fees? evidence of rising external financing costs. *The Review of Financial Studies 13*(1), 191–218.

Andrei, D., J. Cujean, and M. Wilson (2023). The lost capital asset pricing model. *Review of Economic Studies 90*(6), 2703–2762.

Andrei, D., W. Mann, and N. Moyen (2019). Why did the q theory of investment start working? *Journal of Financial Economics 133*(2), 251–272.

Ball, R. and P. Brown (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research 6*(2), 159–178.

Beaver, W. H. (1968). The information content of annual earnings announcements. *Journal of Accounting Research 6*, 67–92.

Bolton, P., H. Chen, and N. Wang (2011). A unified theory of Tobin's q, corporate investment, financing, and risk management. *The Journal of Finance 66*(5), 1545–1578.

Broussard, J. P., J. Mikkonen, and V. Puttonen (2016). Style migration in europe. *European Financial Management 22*(5), 797–816.

Chen, Q., I. Goldstein, and W. Jiang (2007). Price informativeness and investment sensitivity to stock price. *Review of Financial Studies 20*(3), 619–650.

Cochrane, J. H. (1991). Production-based asset pricing and the link between stock returns and economic fluctuations. *The journal of finance 46*(1), 209–237.

Cochrane, J. H. (2011). Presidential address: Discount rates. *The Journal of finance 66*(4), 1047–1108.

Cohen, R. B., C. Polk, and T. Vuolteenaho (2003). The value spread. *The Journal of Finance 58*(2), 609–641.

Collins, D. W., S. P. Kothari, and J. D. Rayburn (1987). Firm size and the information content of prices with respect to earnings. *Journal of Accounting and Economics 9*(2), 111–138.

Cooper, M. J., H. Gulen, and M. J. Schill (2008). Asset growth and the cross-section of stock returns. *The Journal of Finance 63*(4), 1609–1651.

Da, R., S. Nagel, and D. Xiu (2024). The statistical limit of arbitrage. Technical report, National Bureau of Economic Research.

Easley, D., N. M. Kiefer, M. O'Hara, and J. B. Paperman (1996). Liquidity, information, and infrequently traded stocks. *The Journal of Finance 51*(4), 1405–1436.

Eberly, J., S. Rebelo, and N. Vincent (2012). What explains the lagged-investment effect? *Journal of Monetary Economics 59*(4), 370–380.

Engelberg, J., R. D. McLean, and J. Pontiff (2018). Anomalies and news. *The Journal of Finance 73*(5), 1971–2001.

Fama, E. F. and K. R. French (1992). The cross-section of expected stock returns. *The Journal of Finance 47*(2), 427–465.

Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics 33*(1), 3–56.

Fama, E. F. and K. R. French (2007). Migration. *Financial Analysts Journal 63*(3), 48–58.

Fama, E. F. and K. R. French (2015). A five-factor asset pricing model. *Journal of financial economics 116*(1), 1–22.

Fujisaki, M., G. Kallianpur, and H. Kunita (1972). Stochastic differential equations for the non linear filtering problem. *Osaka Journal of Mathematics 9*(1), 19–40.

Gomes, J. F. (2001). Financing investment. *American Economic Review 91*(5), 1263–1285.

Grossman, S. J. and J. E. Stiglitz (1980). On the impossibility of informationally efficient markets. *American Economic Review 70*(3), 393–408.

Hansen, L. P. and S. F. Richard (1987). The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. *Econometrica: Journal of the Econometric Society*, 587–613.

Hasbrouck, J. (1991). Measuring the information content of stock trades. *Journal of Finance 46*(1), 179–207.

Hayashi, F. (1982). Tobin's marginal q and average q: A neoclassical interpretation. *Econometrica: Journal of the Econometric Society*, 213–224.

Hayek, F. A. (1945, 9). The use of knowledge in society. *The American Economic Review 35*(4), 519–530.

Hennessy, C. A. and T. M. Whited (2007). How costly is external financing? evidence from a structural estimation. *The Journal of Finance 62*(4), 1705–1745.

Hofstadter, D. R. (1979). *Gödel, Escher, Bach: An Eternal Golden Braid*. Basic Books.

Hou, K., C. Xue, and L. Zhang (2015). Digesting anomalies: An investment approach. *The Review of Financial Studies 28*(3), 650–705.

Jung, J. and R. J. Shiller (2005). Samuelson's dictum and the stock market. *Economic Inquiry 43*(2), 221–228.

Kim, J., Y. Kim, and S. Lee (2025). Industry-adjusted book-to-market ratio and value premium. *Finance Research Letters*, 108347.

Kozak, S., S. Nagel, and S. Santosh (2020). Shrinking the cross-section. *Journal of Financial Economics 135*(2), 271–292.

Liptser, R. S. and A. N. Shiryaev (2001). *Statistics of Random Processes II*. Springer Verlag, New York.

Marrow, B. and S. Nagel (2024). Real-time discovery and tracking of return-based anomalies. Technical report, Working Paper.

McLean, R. D. and J. Pontiff (2016). Does academic research destroy stock return predictability? *The Journal of Finance 71*(1), 5–32.

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *The Journal of Finance 42*(3), 483–510.

Morck, R., B. Yeung, and W. Yu (2000). The information content of stock markets: Why do emerging markets have synchronous stock price movements? *Journal of Financial Economics 58*(1–2), 215–260.

Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of financial economics 108*(1), 1–28.

Restoy, F. and G. M. Rockinger (1994). On stock market returns and returns on investment. *The Journal of Finance 49*(2), 543–556.

Roll, R. (1988). r^2. *Journal of Finance 43*(3), 541–566.

Shleifer, A. and R. W. Vishny (1997). The limits of arbitrage. *The Journal of finance 52*(1), 35–55.

Titman, S., K. J. Wei, and F. Xie (2004). Capital investments and stock returns. *Journal of Financial and Quantitative Analysis 39*(4), 677–700.

Vuolteenaho, T. (2002). What drives firm-level stock returns? *Journal of Finance 57*(1), 233–264.

Welch, I. and A. Goyal (2008). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies 21*(4), 1455–1508.

Whited, T. M. (1992). Debt, liquidity constraints, and corporate investment: Evidence from panel data. *The Journal of Finance 47*(4), 1425–1460.

Zhang, L. (2005). The value premium. *The Journal of Finance 60*(1), 67–103.

A Appendix: Proofs

A.1 Proof of Proposition 1

Let $v_n \equiv v(Z_{n,t}, \hat{X}^{\mathcal{H}}_{n,t})$ denote normalized firm value, where $v_Z \equiv \frac{\partial v_n}{\partial Z}$ and $v_{\hat{X}} \equiv \frac{\partial v_n}{\partial \hat{X}}$ are the partial derivatives of firm value with respect to its state variables. The total equity return is $dR^E_{n,t} = \frac{dv_n + D_{n,t}dt}{v_n}$.

By local efficiency, the expected return under the pricing filtration $\mathcal{H}_{n,t}(\mathcal{A})$ is the risk-free rate:

$$\mathbb{E}^{\mathcal{H}_n(\mathcal{A})}_t[dv_n + D_{n,t}dt] = rv_n dt. \tag{A1}$$

To find the expected return under the pooled filtration $\mathcal{G}_t$, we evaluate the difference in expected price changes across the two filtrations. By Itô's Lemma, because the quadratic-variation terms are identical under the two conditional expectations, the difference in the expected differential dv_n is driven entirely by the difference in the expected drifts of the state variables:

$$\mathbb{E}^{\mathcal{G}}_t[dv_n] - \mathbb{E}^{\mathcal{H}_n(\mathcal{A})}_t[dv_n] = v_Z(\mathbb{E}^{\mathcal{G}}_t[dZ_{n,t}] - \mathbb{E}^{\mathcal{H}_n(\mathcal{A})}_t[dZ_{n,t}]) + v_{\hat{X}}(\mathbb{E}^{\mathcal{G}}_t[d\hat{X}^{\mathcal{H}}_{n,t}] - \mathbb{E}^{\mathcal{H}_n(\mathcal{A})}_t[d\hat{X}^{\mathcal{H}}_{n,t}]). \tag{A2}$$

We evaluate these two drift differences independently. First, consider the fundamental profitability $Z_{n,t}$, which loads on the common state with coefficient β_n. Under the local pricing filtration, the expected state is $\hat{X}^{\mathcal{H}}_{n,t}$. Under the pooled filtration, the true state X_t is revealed. The difference in the expected drift is therefore identically proportional to the wedge:

$$\mathbb{E}^{\mathcal{G}}_t[dZ_{n,t}] - \mathbb{E}^{\mathcal{H}_n(\mathcal{A})}_t[dZ_{n,t}] = \beta_n X_t dt - \beta_n \hat{X}^{\mathcal{H}}_{n,t}dt = \beta_n \mathcal{W}_{n,t}dt. \tag{A3}$$

Second, consider the local belief $\hat{X}^{\mathcal{H}}_{n,t}$, which is updated via Bayesian filtering. Since X_t is mean-reverting, the belief inherits this drift under $\mathcal{H}_{n,t}(\mathcal{A})$, yielding $\mathbb{E}^{\mathcal{H}_n(\mathcal{A})}_t[d\hat{X}^{\mathcal{H}}_{n,t}] = -\kappa_X \hat{X}^{\mathcal{H}}_{n,t}dt$. The belief evolves dynamically as $d\hat{X}^{\mathcal{H}}_{n,t} = -\kappa_X \hat{X}^{\mathcal{H}}_{n,t}dt + \lambda^{\mathcal{H}}_{n,t}d\widetilde{W}^{\mathcal{H}}_{n,t}$, where $d\widetilde{W}^{\mathcal{H}}_{n,t}$ is the innovation process and $\lambda^{\mathcal{H}}_{n,t}$ is the optimal filter gain. Under the pooled filtration $\mathcal{G}_t$, the physical signals arriving at the firm are generated by the true state X_t. Consequently, the innovation process has a conditionally expected drift equal to the filter error $\mathcal{W}_{n,t}dt$. Thus, the expected belief revision under $\mathcal{G}_t$ is:

$$\mathbb{E}^{\mathcal{G}}_t[d\hat{X}^{\mathcal{H}}_{n,t}] = -\kappa_X \hat{X}^{\mathcal{H}}_{n,t}dt + \lambda^{\mathcal{H}}_{n,t}(X_t - \hat{X}^{\mathcal{H}}_{n,t})dt = -\kappa_X \hat{X}^{\mathcal{H}}_{n,t}dt + \lambda^{\mathcal{H}}_{n,t}\mathcal{W}_{n,t}dt.$$

Subtracting the expected drift under $\mathcal{H}_{n,t}(\mathcal{A})$ isolates the predictable learning component:

$$\mathbb{E}^{\mathcal{G}}_t[d\hat{X}^{\mathcal{H}}_{n,t}] - \mathbb{E}^{\mathcal{H}_n(\mathcal{A})}_t[d\hat{X}^{\mathcal{H}}_{n,t}] = \lambda^{\mathcal{H}}_{n,t}\mathcal{W}_{n,t}dt. \tag{A4}$$

Substituting Equations (A3) and (A4) into Equation (A2) yields the total difference in expected price changes:

$$\mathbb{E}_t^{\mathcal{G}}[dv_n] - \mathbb{E}_t^{\mathcal{H}_n(\mathcal{A})}[dv_n] = \left[\beta_n v_Z + \lambda_{n,t}^{\mathcal{H}} v_{\hat{X}}\right] \mathcal{W}_{n,t} dt.$$

Because the dividend rate $D_{n,t}$ is observable at date t under the local filtration, the expected dividend yield is identical under both filtrations and cancels out in the difference. Dividing the expected price difference by v_n therefore translates exactly into the expected return difference:

$$\mathbb{E}_t^{\mathcal{G}}[dR_{n,t}^E] - \mathbb{E}_t^{\mathcal{H}_n(\mathcal{A})}[dR_{n,t}^E] = \left[\frac{\beta_n v_Z + \lambda_{n,t}^{\mathcal{H}} v_{\hat{X}}}{v_n}\right] \mathcal{W}_{n,t} dt.$$

Using the local efficiency condition from Equation (A1) ($\mathbb{E}_t^{\mathcal{H}_n(\mathcal{A})}[dR_{n,t}^E] = rdt$), we obtain the final result:

$$\frac{\mathbb{E}_t^{\mathcal{G}}[dR_{n,t}^E]}{dt} = r + \left[\frac{\beta_n v_Z + \lambda_{n,t}^{\mathcal{H}} v_{\hat{X}}}{v_n}\right] \mathcal{W}_{n,t}.$$

This matches Equation (5) with $\Gamma_{n,t}(\mathcal{A})$ defined as in Equation (6). □

A.2 Proof of Proposition 2

By the proposition's non-degeneracy condition, there exists at least one observable-characteristic portfolio in which $B_{n,t}$ has positive within-portfolio variation. Fix such a portfolio and denote it by s^*. Since firms are sorted using all known payoff-relevant observable states and parameters, these objects are held fixed within s^*. Assumptions 1 and 2 therefore imply that, within s^*, $B \mapsto v(s^*, B)$ is strictly increasing and $B \mapsto \Gamma(s^*, B)(X_t - B)$ is strictly decreasing. Let

$$\mathcal{N}_t(s^*) \equiv \{n : S_{n,t} = s^*\}, \qquad N_t(s^*) \equiv |\mathcal{N}_t(s^*)|.$$

Define the average valuation in this portfolio by

$$\bar{V}_t(s^*) \equiv \frac{1}{N_t(s^*)} \sum_{m \in \mathcal{N}_t(s^*)} V_{m,t}.$$

Now define the observable residual valuation characteristic

$$\chi_{n,t}^V \equiv \mathbb{1}_{\{S_{n,t}=s^*\}} \left(V_{n,t} - \bar{V}_t(s^*)\right).$$

This characteristic is observable: it is the valuation residual inside the observable-characteristic portfolio s^* and is zero outside that portfolio.

Suppose, for contradiction, that no observable characteristic predicts returns. Since $\chi^V_{n,t}$ is an observable characteristic, this implies

$$\mathrm{Cov}_{CS}(\chi^V_{n,t}, \mu_{n,t}(\mathcal{A})) = 0. \tag{A5}$$

Here, Cov_{CS} denotes covariance across the full set of firms $n = 1, \ldots, N$ at date t, and Cov_{s^*} denotes covariance computed across firms in portfolio s^*.

Because $\chi^V_{n,t}$ has zero average within portfolio s^* and is zero outside that portfolio,

$$\mathrm{Cov}_{CS}(\chi^V_{n,t}, \mu_{n,t}(\mathcal{A})) = \frac{N_t(s^*)}{N} \mathrm{Cov}_{s^*}(V_{n,t}, \mu_{n,t}(\mathcal{A})).$$

Within portfolio s^*, observable characteristics are held fixed, so $V_{n,t} = v(s^*, B_{n,t})$ and

$$\mu_{n,t}(\mathcal{A}) = \Gamma(s^*, B_{n,t})(X_t - B_{n,t}).$$

Therefore

$$\mathrm{Cov}_{s^*}(V_{n,t}, \mu_{n,t}(\mathcal{A})) = \mathrm{Cov}_{s^*}(v(s^*, B_{n,t}), \Gamma(s^*, B_{n,t})(X_t - B_{n,t})).$$

By Assumption 1, $B \mapsto v(s^*, B)$ is strictly increasing. By Assumption 2, $B \mapsto \Gamma(s^*, B)(X_t - B)$ is strictly decreasing. Since $B_{n,t}$ has positive within-portfolio variation, Chebyshev's algebraic covariance inequality for oppositely monotone functions implies

$$\mathrm{Cov}_{s^*}(v(s^*, B_{n,t}), \Gamma(s^*, B_{n,t})(X_t - B_{n,t})) < 0.$$

Hence

$$\mathrm{Cov}_{CS}(\chi^V_{n,t}, \mu_{n,t}(\mathcal{A})) < 0.$$

This contradicts (A5). Therefore at least one observable valuation characteristic predicts returns cross-sectionally. □

A.3 Proof of Lemma 1

The filtering problem is scalar and linear-Gaussian. The total signal precision is Λ_n, so the posterior variance $\nu_{n,t}$ satisfies the Riccati equation

$$\dot{\nu}_{n,t} = \sigma_X^2 - 2\kappa_X \nu_{n,t} - \Lambda_n \nu_{n,t}^2. \tag{A6}$$

In steady state, $\dot{\nu}_{n,t} = 0$, so

$$\Lambda_n \bar{\nu}_n^2 + 2\kappa_X \bar{\nu}_n - \sigma_X^2 = 0.$$

Taking the positive root gives

$$\bar{\nu}_n = \frac{-\kappa_X + \sqrt{\kappa_X^2 + \Lambda_n \sigma_X^2}}{\Lambda_n} = \frac{\sigma_X^2}{\kappa_X + \sqrt{\kappa_X^2 + \Lambda_n \sigma_X^2}}.$$

The Kalman-Bucy update for the posterior mean has drift $-\kappa_X \hat{X}^{\mathcal{H}}_{n,t} dt$ and innovation loadings equal to posterior variance times signal loading divided by signal volatility. The three orthogonal channels therefore load on the posterior mean with coefficients $\beta_n \bar{\nu}_n / \sigma_Z$, $\bar{\nu}_n \sqrt{A}$, and $\bar{\nu}_n \sqrt{\tau_{\mathcal{A}}}$, which gives the stated law of motion. □

A.4 Proof of Proposition 3

Let v^{AIP} denote the expected present value of operating cash flows generated by existing capital. Since installed capital depreciates at rate δ, one unit of capital at time t produces $e^{-\delta s} Z_{t+s} ds$ at time $t+s$. Hence

$$v^{AIP}(Z, \hat{X}^{\mathcal{H}}) = \mathbb{E}_t^{\mathcal{H}} \left[\int_0^\infty e^{-(r+\delta)s} Z_{t+s} ds \right].$$

Under $\mathcal{H}_{n,t}$, the conditional means satisfy

$$\mathbb{E}_t^{\mathcal{H}}[X_{t+s}] = e^{-\kappa_X s} \hat{X}^{\mathcal{H}},$$

$$\mathbb{E}_t^{\mathcal{H}}[Z_{t+s}] = e^{-\kappa_Z s} Z + \frac{\mu_0}{\kappa_Z}(1 - e^{-\kappa_Z s}) + \beta_n \frac{e^{-\kappa_X s} - e^{-\kappa_Z s}}{\kappa_Z - \kappa_X} \hat{X}^{\mathcal{H}}.$$

In the knife-edge case $\kappa_Z = \kappa_X$, the final term is interpreted as its limit, $\beta_n s e^{-\kappa_X s} \hat{X}^{\mathcal{H}}$. Integrating term by term gives

$$v^{AIP}(Z, \hat{X}^{\mathcal{H}}) = A_0 + A_Z Z + A_X \hat{X}^{\mathcal{H}},$$

where

$$A_Z = \frac{1}{r + \delta + \kappa_Z},$$

$$A_X = \frac{\beta_n}{\kappa_Z - \kappa_X} \left(\frac{1}{r + \delta + \kappa_X} - \frac{1}{r + \delta + \kappa_Z} \right) = \frac{\beta_n}{(r + \delta + \kappa_X)(r + \delta + \kappa_Z)},$$

$$A_0 = \frac{\mu_0}{\kappa_Z} \left(\frac{1}{r + \delta} - \frac{1}{r + \delta + \kappa_Z} \right) = \frac{\mu_0}{(r + \delta)(r + \delta + \kappa_Z)}.$$

Thus v^{AIP} is affine in $(Z, \hat{X}^{\mathcal{H}})$. Define the option-to-scale component by

$$g(Z, \hat{X}^{\mathcal{H}}) = v(Z, \hat{X}^{\mathcal{H}}) - v^{AIP}(Z, \hat{X}^{\mathcal{H}}).$$

This gives the decomposition in (11).

We next derive the investment rule. The terms in the HJB that depend on i are

$$-\Psi(i, Z) + vi.$$

Using

$$\Psi(i, Z) = i + \frac{\theta_n}{2}(i - \delta)^2 + \frac{\phi_n}{2}(i - Z)_+^2,$$

the derivative with respect to i is

$$-1 - \theta_n(i - \delta) - \phi_n(i - Z)_+ + v.$$

The first-order condition is therefore

$$v = 1 + \theta_n(i - \delta) + \phi_n(i - Z)_+.$$

If $v \leq 1 - \theta_n\delta$, the constraint $i \geq 0$ binds, so $i^*(v, Z) = 0$. If $1 - \theta_n\delta < v \leq 1 + \theta_n(Z - \delta)$, the solution satisfies $i \leq Z$, so the external-finance term is zero and

$$i^*(v, Z) = \delta + \frac{v - 1}{\theta_n}.$$

If $v > 1 + \theta_n(Z - \delta)$, the solution satisfies $i > Z$, so the external-finance term binds and

$$v = 1 + \theta_n(i - \delta) + \phi_n(i - Z).$$

Solving gives

$$i^*(v, Z) = \frac{v - 1 + \theta_n\delta + \phi_n Z}{\theta_n + \phi_n}.$$

Combining the three regions yields (12).

It remains to state the PDE solved by g. Since v^{AIP} is affine, $v^{AIP}_{ZZ} = v^{AIP}_{\hat{X}\hat{X}} = v^{AIP}_{Z\hat{X}} = 0$. Substituting $v^{AIP}_Z = A_Z$ and $v^{AIP}_{\hat{X}} = A_X$ into the drift operator and using the coefficients above gives

$$(r + \delta)v^{AIP} = Z + (\mu_0 - \kappa_Z Z + \beta_n \hat{X}^{\mathcal{H}})v^{AIP}_Z - \kappa_X \hat{X}^{\mathcal{H}} v^{AIP}_{\hat{X}}.$$

Subtract this identity from the HJB, write $v = v^{AIP} + g$, and evaluate the maximand at $i^*(v, Z)$.

The option-to-scale component solves

$$rg = -\Psi(i^*, Z) + (v^{AIP} + g)i^* - \delta g + (\mu_0 - \kappa_Z Z + \beta_n \hat{X}^{\mathcal{H}})g_Z + \frac{1}{2}\sigma_Z^2 g_{ZZ}$$
$$- \kappa_X \hat{X}^{\mathcal{H}} g_{\hat{X}} + \frac{1}{2}\bar{\sigma}_n^2 g_{\hat{X}\hat{X}} + \beta_n \bar{\nu}_n g_{Z\hat{X}}.$$

This is the stated PDE for the option to scale. □

A.5 Proof of Lemma 3

To compute the projected local pricing kernel, we compare the filtering problem under the physical measure $\mathbb{P}$ and the global risk-neutral measure $\mathbb{Q}$ induced by the global stochastic discount factor M_t. Under the physical measure, the latent state follows

$$dX_t = -\kappa_X X_t dt + \sigma_X dW_t^X.$$

Let $\hat{X}_{n,t}^{\mathbb{P}} \equiv \mathbb{E}^{\mathbb{P}}[X_t | \mathcal{H}_{n,t}(\mathcal{A})]$ denote the local belief under this measure. By Lemma 1, the corresponding belief under the physical measure follows directly as

$$d\hat{X}_{n,t}^{\mathbb{P}} = -\kappa_X \hat{X}_{n,t}^{\mathbb{P}} dt + \frac{\beta_n \bar{\nu}_n}{\sigma_Z} d\widetilde{W}_{n,t}^{Z,\mathbb{P}} + \bar{\nu}_n \sqrt{A} d\widetilde{W}_{n,t}^{s,\mathbb{P}} + \bar{\nu}_n \sqrt{\tau_{\mathcal{A}}} d\widetilde{W}_{n,t}^{\mathcal{A},\mathbb{P}}. \tag{A7}$$

The global stochastic discount factor M_t defines a global risk-neutral measure $\mathbb{Q}$ on the pooled filtration $\mathcal{G}_t$. Because M_t prices the macroeconomic shock with risk price λ_X, the latent state under this global pricing measure absorbs this risk premium, yielding

$$dX_t = (-\kappa_X X_t - \sigma_X \lambda_X)dt + \sigma_X dW_t^{X,\mathbb{Q}}.$$

Let $\hat{X}_{n,t}^{\mathbb{Q}} \equiv \mathbb{E}^{\mathbb{Q}}[X_t | \mathcal{H}_{n,t}(\mathcal{A})]$ denote the corresponding risk-adjusted local conditional belief. The Riccati equation for the posterior variance (A6) depends on diffusion variances and signal loadings, not on constant drift shifts. Hence the $\mathbb{Q}$ filter has the same steady-state variance $\bar{\nu}_n$. Driven by the $\mathbb{Q}$-innovations and incorporating the macroeconomic risk price, the belief under the global pricing measure satisfies

$$d\hat{X}_{n,t}^{\mathbb{Q}} = (-\kappa_X \hat{X}_{n,t}^{\mathbb{Q}} - \sigma_X \lambda_X)dt + \frac{\beta_n \bar{\nu}_n}{\sigma_Z} d\widetilde{W}_{n,t}^{Z,\mathbb{Q}} + \bar{\nu}_n \sqrt{A} d\widetilde{W}_{n,t}^{s,\mathbb{Q}} + \bar{\nu}_n \sqrt{\tau_{\mathcal{A}}} d\widetilde{W}_{n,t}^{\mathcal{A},\mathbb{Q}}. \tag{A8}$$

The measure-change belief difference is defined as

$$\gamma_{n,t} \equiv \hat{X}_{n,t}^{\mathbb{Q}} - \hat{X}_{n,t}^{\mathbb{P}}.$$

Subtracting the physical filter (A7) from the pricing filter (A8) yields the dynamics of the measure-change belief difference:

$$\begin{aligned} d\gamma_{n,t} &= (-\kappa_X \gamma_{n,t} - \sigma_X \lambda_X)dt \\ &\quad + \frac{\beta_n \bar{\nu}_n}{\sigma_Z}(d\widetilde{W}^{Z,\mathbb{Q}}_{n,t} - d\widetilde{W}^{Z,\mathbb{P}}_{n,t}) + \bar{\nu}_n \sqrt{A}(d\widetilde{W}^{s,\mathbb{Q}}_{n,t} - d\widetilde{W}^{s,\mathbb{P}}_{n,t}) + \bar{\nu}_n \sqrt{\tau_{\mathcal{A}}}(d\widetilde{W}^{\mathcal{A},\mathbb{Q}}_{n,t} - d\widetilde{W}^{\mathcal{A},\mathbb{P}}_{n,t}). \end{aligned} \tag{A9}$$

Recall that a standard innovation is the realized differential minus its conditional expectation, scaled by the observation volatility. For the cash-flow channel, this yields

$$d\widetilde{W}^{Z,\mathbb{Q}}_{n,t} - d\widetilde{W}^{Z,\mathbb{P}}_{n,t} = \frac{1}{\sigma_Z}\left(\mathbb{E}^{\mathbb{P}}_t[dZ_{n,t}] - \mathbb{E}^{\mathbb{Q}}_t[dZ_{n,t}]\right).$$

Because the observation noise in this channel carries no direct price of risk under the global stochastic discount factor, expected cash flows differ across measures only through the difference in local conditional beliefs. Hence the difference in expectations is exactly $\beta_n(\hat{X}^{\mathbb{P}}_{n,t} - \hat{X}^{\mathbb{Q}}_{n,t})dt$. Substituting the definition of the measure-change belief difference gives

$$d\widetilde{W}^{Z,\mathbb{Q}}_{n,t} - d\widetilde{W}^{Z,\mathbb{P}}_{n,t} = -\frac{\beta_n}{\sigma_Z}\gamma_{n,t}dt.$$

Applying identical logic to the private-signal and public-aggregation channels gives

$$\begin{aligned} d\widetilde{W}^{s,\mathbb{Q}}_{n,t} - d\widetilde{W}^{s,\mathbb{P}}_{n,t} &= -\sqrt{A}\gamma_{n,t}dt, \\ d\widetilde{W}^{\mathcal{A},\mathbb{Q}}_{n,t} - d\widetilde{W}^{\mathcal{A},\mathbb{P}}_{n,t} &= -\sqrt{\tau_{\mathcal{A}}}\gamma_{n,t}dt. \end{aligned}$$

For the latter equation, defining the aggregate precision $\tau_{\mathcal{A}}$ via the quadratic variation of aggregate beliefs (8) establishes an informational equivalence to a continuous scalar signal with a unit loading on the latent state X_t and an effective precision $\tau_{\mathcal{A}}$. Thus, the difference in the expected signal drift under the physical and pricing measures is precisely the difference in local beliefs, $-(\hat{X}^{\mathbb{Q}}_{n,t} - \hat{X}^{\mathbb{P}}_{n,t})dt = -\gamma_{n,t}dt$. Scaling this expectation gap by the effective precision factor $\sqrt{\tau_{\mathcal{A}}}$ yields the difference in standardized innovations.

Substituting these innovation differences into the dynamics for the measure-change belief difference (A9) eliminates all stochastic terms, yielding a deterministic ordinary differential equation

$$d\gamma_{n,t} = \left[-\kappa_X \gamma_{n,t} - \sigma_X \lambda_X - \bar{\nu}_n \gamma_{n,t}\left(\frac{\beta_n^2}{\sigma_Z^2} + A + \tau_{\mathcal{A}}\right)\right]dt.$$

Using the definition of total precision (9), $\Lambda_n = \frac{\beta_n^2}{\sigma_Z^2} + A + \tau_{\mathcal{A}}$, this simplifies to

$$d\gamma_{n,t} = \left[-(\kappa_X + \bar{\nu}_n \Lambda_n)\gamma_{n,t} - \sigma_X \lambda_X\right]dt.$$

The stationary value of this belief difference solves the algebraic equation

$$0 = -(\kappa_X + \bar{\nu}_n \Lambda_n)\gamma_n - \sigma_X \lambda_X,$$

which gives

$$\gamma_n = \frac{-\sigma_X \lambda_X}{\kappa_X + \bar{\nu}_n \Lambda_n}.$$

To establish the representation of the projected local pricing kernel $\widehat{M}_{n,t}$, we combine the preceding filtering analysis with the martingale representation theorem. Define the projected kernel as $\widehat{M}_{n,t} \equiv \mathbb{E}^{\mathbb{P}}[M_t \mid \mathcal{H}_{n,t}(\mathcal{A})]$. Because the global pricing kernel discounted at the risk-free rate, $e^{rt}M_t$, is a $\mathbb{P}$-martingale on the global filtration $\mathcal{G}_t$, the tower property ensures that its projection $e^{rt}\widehat{M}_{n,t}$ is a $\mathbb{P}$-martingale on the smaller local filtration $\mathcal{H}_{n,t}(\mathcal{A})$.[6]

By standard continuous-time filtering theory (Liptser and Shiryaev, 2001), the local filtration $\mathcal{H}_{n,t}(\mathcal{A})$ is generated by the three mutually orthogonal innovation processes $\widetilde{W}^{Z,\mathbb{P}}_{n,t}$, $\widetilde{W}^{s,\mathbb{P}}_{n,t}$, and $\widetilde{W}^{\mathcal{A},\mathbb{P}}_{n,t}$. By Lévy's characterization, these constitute a three-dimensional standard Brownian motion under $\mathbb{P}$.[7]

The martingale representation theorem therefore restricts the dynamics of $\widehat{M}_{n,t}$ to the form

$$\frac{d\widehat{M}_{n,t}}{\widehat{M}_{n,t}} = -r dt + \theta_Z d\widetilde{W}^{Z,\mathbb{P}}_{n,t} + \theta_s d\widetilde{W}^{s,\mathbb{P}}_{n,t} + \theta_{\mathcal{A}} d\widetilde{W}^{\mathcal{A},\mathbb{P}}_{n,t},$$

where the local risk prices θ_i remain to be determined.

Define the local pricing measure $\mathbb{Q}^n$ on $\mathcal{H}_{n,t}(\mathcal{A})$ through the projected kernel $\widehat{M}_{n,t}$; equivalently, its density process is the discounted projected kernel $e^{rt}\widehat{M}_{n,t}$, and $\mathbb{Q}^n$ is the restriction of the global risk-neutral measure $\mathbb{Q}$ to the local pricing filtration. Girsanov's theorem translates each loading into a deterministic drift shift on the corresponding physical innovation: $d\widetilde{W}^{i,\mathbb{Q}^n}_{n,t} - d\widetilde{W}^{i,\mathbb{P}}_{n,t} = -\theta_i dt$.

[6] Let $N_t \equiv e^{rt}M_t$. Because e^{rt} is deterministic, the discounted projected kernel is $\widehat{N}_t \equiv e^{rt}\widehat{M}_{n,t} = \mathbb{E}^{\mathbb{P}}[N_t \mid \mathcal{H}_{n,t}(\mathcal{A})]$. For $s \leq t$, sequentially applying the tower property to the nested filtrations $\mathcal{H}_{n,s}(\mathcal{A}) \subseteq \mathcal{H}_{n,t}(\mathcal{A})$ and $\mathcal{H}_{n,s}(\mathcal{A}) \subseteq \mathcal{G}_s$ yields $\mathbb{E}^{\mathbb{P}}[\widehat{N}_t \mid \mathcal{H}_{n,s}(\mathcal{A})] = \mathbb{E}^{\mathbb{P}}[N_t \mid \mathcal{H}_{n,s}(\mathcal{A})] = \mathbb{E}^{\mathbb{P}}[\mathbb{E}^{\mathbb{P}}[N_t \mid \mathcal{G}_s] \mid \mathcal{H}_{n,s}(\mathcal{A})]$. Because N is a $\mathbb{P}$-martingale on $\mathcal{G}$, this evaluates to $\mathbb{E}^{\mathbb{P}}[N_s \mid \mathcal{H}_{n,s}(\mathcal{A})] = \widehat{N}_s$. Integrability and adaptedness follow immediately from the properties of conditional expectation.

[7] Treating the public aggregation as a surrogate continuous signal, each innovation is the standardized prediction error for its observation channel. For example, $d\widetilde{W}^{Z,\mathbb{P}}_{n,t} \equiv \frac{1}{\sigma_Z}(dZ_{n,t} - \mathbb{E}^{\mathbb{P}}[dZ_{n,t} \mid \mathcal{H}_{n,t}(\mathcal{A})]) = \frac{\beta_n}{\sigma_Z}(X_t - \hat{X}^{\mathbb{P}}_{n,t})dt + dB^Z_{n,t}$, where $dB^Z_{n,t}$ denotes the Brownian noise in the orthogonalized cash-flow observation channel. Analogous expressions hold for the private-signal and public-aggregation channels. By definition of the conditional belief $\hat{X}^{\mathbb{P}}_{n,t}$, the local expectation of the drift $X_t - \hat{X}^{\mathbb{P}}_{n,t}$ is zero, verifying that each innovation is a continuous $\mathbb{P}$-martingale on $\mathcal{H}_{n,t}(\mathcal{A})$. The orthogonalized observation noises have quadratic variations dt and zero cross-variations, so the result follows directly from the multivariate Lévy characterization.

Because $\mathbb{Q}^n$ is the restriction of $\mathbb{Q}$ to $\mathcal{H}_{n,t}(\mathcal{A})$, the local innovation Brownian motions coincide on the local filtration, so $d\widetilde{W}_{n,t}^{i,\mathbb{Q}^n} = d\widetilde{W}_{n,t}^{i,\mathbb{Q}}$. The filtering analysis above characterized this same change of measure independently. Matching the Girsanov drift shifts term-by-term to the stationary expectation gaps computed previously yields

$$\theta_Z = \frac{\beta_n}{\sigma_Z}\gamma_n, \qquad \theta_s = \sqrt{A}\gamma_n, \qquad \theta_{\mathcal{A}} = \sqrt{\tau_{\mathcal{A}}}\gamma_n.$$

Substituting the uniquely identified loadings θ_i into the SDE for $\widehat{M}_{n,t}$ yields equation (13). $\square$

A.6 Proof of Proposition 4

By Lemma 3, the projected local pricing kernel satisfies

$$\frac{d\widehat{M}_{n,t}}{\widehat{M}_{n,t}} = -rdt + \frac{\beta_n}{\sigma_Z}\gamma_n d\widetilde{W}_{n,t}^{Z,\mathbb{P}} + \sqrt{A}\gamma_n d\widetilde{W}_{n,t}^{s,\mathbb{P}} + \sqrt{\tau_{\mathcal{A}}}\gamma_n d\widetilde{W}_{n,t}^{\mathcal{A},\mathbb{P}}.$$

Define the equivalent local risk-neutral measure $\mathbb{Q}^n$ using $\widehat{M}_{n,t}$. The drift term $-rdt$ prices the risk-free asset, while the martingale part determines the Brownian shifts. Girsanov's theorem gives

$$d\widetilde{W}_{n,t}^{Z,\mathbb{P}} = d\widetilde{W}_{n,t}^{Z,\mathbb{Q}} + \frac{\beta_n}{\sigma_Z}\gamma_n dt, \qquad d\widetilde{W}_{n,t}^{s,\mathbb{P}} = d\widetilde{W}_{n,t}^{s,\mathbb{Q}} + \sqrt{A}\gamma_n dt, \qquad d\widetilde{W}_{n,t}^{\mathcal{A},\mathbb{P}} = d\widetilde{W}_{n,t}^{\mathcal{A},\mathbb{Q}} + \sqrt{\tau_{\mathcal{A}}}\gamma_n dt.$$

Under $\mathbb{P}$, the local state satisfies

$$dZ_{n,t} = (\mu_0 - \kappa_Z Z_{n,t} + \beta_n \hat{X}_{n,t}^{\mathcal{H}})dt + \sigma_Z d\widetilde{W}_{n,t}^{Z,\mathbb{P}},$$

$$d\hat{X}_{n,t}^{\mathcal{H}} = -\kappa_X \hat{X}_{n,t}^{\mathcal{H}} dt + \frac{\beta_n \bar{\nu}_n}{\sigma_Z} d\widetilde{W}_{n,t}^{Z,\mathbb{P}} + \bar{\nu}_n\sqrt{A} d\widetilde{W}_{n,t}^{s,\mathbb{P}} + \bar{\nu}_n\sqrt{\tau_{\mathcal{A}}} d\widetilde{W}_{n,t}^{\mathcal{A},\mathbb{P}}.$$

Substituting the Girsanov shifts gives

$$dZ_{n,t} = \left(\mu_0 - \kappa_Z Z_{n,t} + \beta_n \hat{X}_{n,t}^{\mathcal{H}} + \beta_n\gamma_n\right) dt + \sigma_Z d\widetilde{W}_{n,t}^{Z,\mathbb{Q}}.$$

For the belief,

$$d\hat{X}_{n,t}^{\mathcal{H}} = \left(-\kappa_X \hat{X}_{n,t}^{\mathcal{H}} + \bar{\nu}_n \left(\frac{\beta_n^2}{\sigma_Z^2} + A + \tau_{\mathcal{A}}\right)\gamma_n\right) dt + \frac{\beta_n \bar{\nu}_n}{\sigma_Z} d\widetilde{W}_{n,t}^{Z,\mathbb{Q}} + \bar{\nu}_n\sqrt{A} d\widetilde{W}_{n,t}^{s,\mathbb{Q}} + \bar{\nu}_n\sqrt{\tau_{\mathcal{A}}} d\widetilde{W}_{n,t}^{\mathcal{A},\mathbb{Q}}.$$

Using $\Lambda_n(\mathcal{A}) = \frac{\beta_n^2}{\sigma_Z^2} + A + \tau_{\mathcal{A}}$ gives

$$d\hat{X}_{n,t}^{\mathcal{H}} = \left(-\kappa_X \hat{X}_{n,t}^{\mathcal{H}} + \bar{\nu}_n\Lambda_n(\mathcal{A})\gamma_n\right) dt + \frac{\beta_n \bar{\nu}_n}{\sigma_Z} d\widetilde{W}_{n,t}^{Z,\mathbb{Q}} + \bar{\nu}_n\sqrt{A} d\widetilde{W}_{n,t}^{s,\mathbb{Q}} + \bar{\nu}_n\sqrt{\tau_{\mathcal{A}}} d\widetilde{W}_{n,t}^{\mathcal{A},\mathbb{Q}}.$$

The capital law, cash-flow function, and adjustment-cost function are unchanged. Therefore the homogeneity argument from Proposition 3 applies directly, so $V = Kv_n$.

The equivalent change of measure changes drifts but not instantaneous quadratic variations. Hence, under $\mathbb{Q}^n$,

$$\frac{d\langle Z_n\rangle_t}{dt} = \sigma_Z^2, \qquad \frac{d\langle Z_n, \hat{X}_n^{\mathcal{H}}\rangle_t}{dt} = \beta_n\bar{\nu}_n, \qquad \frac{d\langle \hat{X}_n^{\mathcal{H}}\rangle_t}{dt} = \bar{\nu}_n^2\Lambda_n(\mathcal{A}).$$

Since under $\mathbb{Q}^n$ the firm is priced by discounting expected cash flows at rate r, the required return on normalized value equals normalized current cash flow plus the expected change in normalized continuation value, including the capital-growth term $(i-\delta)v_n$. This gives the HJB:

$$\begin{aligned} rv_n = \max_{i\geq 0}\Big\{&Z - \Psi(i,Z) + (i-\delta)v_n + \Big(\mu_0 - \kappa_Z Z + \beta_n\hat{X}^{\mathcal{H}} + \beta_n\gamma_n\Big)v_Z \\ &+ \Big(-\kappa_X\hat{X}^{\mathcal{H}} + \bar{\nu}_n\Lambda_n(\mathcal{A})\gamma_n\Big)v_{\hat{X}} + \frac{1}{2}\sigma_Z^2 v_{ZZ} + \beta_n\bar{\nu}_n v_{Z\hat{X}} + \frac{1}{2}\bar{\nu}_n^2\Lambda_n(\mathcal{A})v_{\hat{X}\hat{X}}\Big\}. \end{aligned}$$

The only terms in the HJB that depend on i remain $-\Psi(i,Z) + iv_n$, so the pointwise investment problem is unchanged and the three-region investment rule from Proposition 3 continues to hold, with v replaced by v_n.

Finally, the risk-adjusted drift system is affine in $(Z, \hat{X}^{\mathcal{H}})$, so the assets-in-place value under $\mathbb{Q}^n$ remains affine:

$$v^{AIP}(Z, \hat{X}^{\mathcal{H}}) = A_0^{\mathbb{Q}} + A_Z Z + A_X\hat{X}^{\mathcal{H}}.$$

Because v^{AIP} is linear in the state variables, its second derivatives vanish. Itô's lemma therefore requires only the first derivatives, $v_Z^{AIP} = A_Z$ and $v_{\hat{X}}^{AIP} = A_X$. Substituting the affine guess and these derivatives into the risk-adjusted pricing equation $(r+\delta)v^{AIP} = Z + \mathbb{E}_t^{\mathbb{Q}^n}[dv^{AIP}]/dt$, then matching the coefficients on Z, $\hat{X}^{\mathcal{H}}$, and the constants yields the system:

$$\begin{aligned} A_Z &= \frac{1}{r+\delta+\kappa_Z}, \\ A_X &= \frac{\beta_n}{(r+\delta+\kappa_Z)(r+\delta+\kappa_X)}, \\ A_0^{\mathbb{Q}} &= \frac{(\mu_0 + \beta_n\gamma_n)A_Z + \bar{\nu}_n\Lambda_n(\mathcal{A})\gamma_n A_X}{r+\delta}. \end{aligned}$$

Thus the slope coefficients are unchanged relative to Proposition 3, while the intercept is shifted by the constant risk corrections.

Define the option-to-scale component by $g = v_n - v^{AIP}$. Subtracting the assets-in-place pricing equation from the HJB gives the risk-adjusted option-to-scale PDE:

$$rg = -\Psi(i^*, Z) + (v^{AIP} + g)i^* - \delta g + \left(\mu_0 - \kappa_Z Z + \beta_n \hat{X}^{\mathcal{H}} + \beta_n \gamma_n\right) g_Z + \frac{1}{2}\sigma_Z^2 g_{ZZ}$$
$$+ \left(-\kappa_X \hat{X}^{\mathcal{H}} + \bar{\nu}_n \Lambda_n(\mathcal{A})\gamma_n\right) g_{\hat{X}} + \frac{1}{2}\bar{\nu}_n^2 \Lambda_n(\mathcal{A}) g_{\hat{X}\hat{X}} + \beta_n \bar{\nu}_n g_{Z\hat{X}},$$

where $i^* = i^*(v^{AIP} + g, Z)$ is given by the same three-region investment rule as in Proposition 3. Since the risk corrections $\beta_n \gamma_n$ and $\bar{\nu}_n \Lambda_n(\mathcal{A})\gamma_n$ are constants at the steady-state filtering variance, this PDE remains two-dimensional in $(Z, \hat{X}^{\mathcal{H}})$ and no new state variable is introduced. □

A.7 Proof of Proposition 5

We now evaluate the same return under the pooled filtration $\mathcal{G}_t$. Under $\mathcal{H}_{n,t}(\mathcal{A})$, the innovations have zero conditional mean. Under $\mathcal{G}_t$, the true state X_t is known, so each innovation has a conditional drift proportional to the wedge $\mathcal{W}_{n,t} = X_t - \hat{X}^{\mathcal{H}}_{n,t}$:

$$\frac{\mathbb{E}_t^{\mathcal{G}}[d\widetilde{W}^Z_{n,t}]}{dt} = \frac{\beta_n}{\sigma_Z}\mathcal{W}_{n,t}, \qquad \frac{\mathbb{E}_t^{\mathcal{G}}[d\widetilde{W}^s_{n,t}]}{dt} = \sqrt{A}\mathcal{W}_{n,t}, \qquad \frac{\mathbb{E}_t^{\mathcal{G}}[d\widetilde{W}^{\mathcal{A}}_{n,t}]}{dt} = \sqrt{\tau_{\mathcal{A}}}\mathcal{W}_{n,t}.$$

Taking the $\mathcal{G}_t$-conditional expectation of the return diffusion (14) gives

$$\frac{\mathbb{E}_t^{\mathcal{G}}[dR^E_{n,t}]}{dt} - r = \mu^R_{n,t} - r + \left(a_{n,Z,t}\frac{\beta_n}{\sigma_Z} + a_{n,s,t}\sqrt{A} + a_{n,\mathcal{A},t}\sqrt{\tau_{\mathcal{A}}}\right)\mathcal{W}_{n,t}.$$

Substituting (16) yields

$$\begin{aligned}\frac{\mathbb{E}_t^{\mathcal{G}}[dR^E_{n,t}]}{dt} - r = &-\frac{\beta_n}{\sigma_Z}\gamma_n a_{n,Z,t} - \sqrt{A}\gamma_n a_{n,s,t} - \sqrt{\tau_{\mathcal{A}}}\gamma_n a_{n,\mathcal{A},t} \\ &+ \left(a_{n,Z,t}\frac{\beta_n}{\sigma_Z} + a_{n,s,t}\sqrt{A} + a_{n,\mathcal{A},t}\sqrt{\tau_{\mathcal{A}}}\right)\mathcal{W}_{n,t}.\end{aligned} \tag{A10}$$

Expanding the wedge coefficient using the return exposure formulas (15) yields

$$\begin{aligned}a_{n,Z,t}\frac{\beta_n}{\sigma_Z} + a_{n,s,t}\sqrt{A} + a_{n,\mathcal{A},t}\sqrt{\tau_{\mathcal{A}}} &= \left(\sigma_Z\frac{v_Z}{v_n} + \frac{\beta_n\bar{\nu}_n}{\sigma_Z}\frac{v_{\hat{X}}}{v_n}\right)\frac{\beta_n}{\sigma_Z} + \left(\bar{\nu}_n\sqrt{A}\frac{v_{\hat{X}}}{v_n}\right)\sqrt{A} + \left(\bar{\nu}_n\sqrt{\tau_{\mathcal{A}}}\frac{v_{\hat{X}}}{v_n}\right)\sqrt{\tau_{\mathcal{A}}} \\ &= \beta_n\frac{v_Z}{v_n} + \bar{\nu}_n\left(\frac{\beta_n^2}{\sigma_Z^2} + A + \tau_{\mathcal{A}}\right)\frac{v_{\hat{X}}}{v_n}.\end{aligned}$$

Using the precision identity (9), $\Lambda_n = \frac{\beta_n^2}{\sigma_Z^2} + A + \tau_{\mathcal{A}}$, the term inside the parentheses is exactly Λ_n.

The coefficient on the wedge therefore simplifies to

$$\beta_n \frac{v_Z}{v_n} + \bar{\nu}_n \Lambda_n \frac{v_{\hat{X}}}{v_n}.$$

This matches exactly the definition of the predictability loading $\Gamma_{n,t}(\mathcal{A})$ stated in Proposition 5. Substituting $\Gamma_{n,t}(\mathcal{A})$ back into (A10) gives

$$\frac{\mathbb{E}_t^{\mathcal{G}}[dR_{n,t}^E]}{dt} - r = -\frac{\beta_n}{\sigma_Z}\gamma_n a_{n,Z,t} - \sqrt{A}\gamma_n a_{n,s,t} - \sqrt{\tau_{\mathcal{A}}}\gamma_n a_{n,\mathcal{A},t} + \Gamma_{n,t}(\mathcal{A})\mathcal{W}_{n,t}.$$

This proves the proposition. □

A.8 Proof of Proposition 6

Throughout this proof, I use the constant-loading benchmark stated in the proposition: the return loadings $(a_{n,Z}, a_{n,s}, a_{n,\mathcal{A}})$, the risk-price terms γ_n, and the market weights ω_n are fixed.

Under the physical measure, write the return diffusion of firm n as

$$dR_{n,t}^E = \mu_{n,t}^R dt + a_{n,Z} d\widetilde{W}_{n,t}^Z + a_{n,s} d\widetilde{W}_{n,t}^s + a_{n,\mathcal{A}} d\widetilde{W}_t^{\mathcal{A}},$$

where $d\widetilde{W}_{n,t}^Z$ and $d\widetilde{W}_{n,t}^s$ are firm-specific innovations, orthogonal across firms, and $d\widetilde{W}_t^{\mathcal{A}}$ is the common public-aggregation innovation.

Defining $\bar{a}_{\mathcal{A}} \equiv \sum_m \omega_m a_{m,\mathcal{A}}$, the value-weighted market return is

$$dR_{M,t}^E = \sum_m \omega_m dR_{m,t}^E = \mu_{M,t}^R dt + \sum_m \omega_m a_{m,Z} d\widetilde{W}_{m,t}^Z + \sum_m \omega_m a_{m,s} d\widetilde{W}_{m,t}^s + \bar{a}_{\mathcal{A}} d\widetilde{W}_t^{\mathcal{A}},$$

In a large cross-section with bounded exposures and diversified weights, the firm-specific innovations diversify away, so only the common public-aggregation innovation contributes to market risk. Hence the relevant market diffusion is

$$dR_{M,t}^E = \mu_{M,t}^R dt + \bar{a}_{\mathcal{A}} d\widetilde{W}_t^{\mathcal{A}}.$$

If $\bar{a}_{\mathcal{A}} \neq 0$, firm n's empirical market beta is therefore

$$\hat{\beta}_n \equiv \frac{\mathrm{Cov}(dR_{n,t}^E, dR_{M,t}^E)}{\mathrm{Var}(dR_{M,t}^E)} = \frac{a_{n,\mathcal{A}}\bar{a}_{\mathcal{A}} dt}{\bar{a}_{\mathcal{A}}^2 dt} = \frac{a_{n,\mathcal{A}}}{\bar{a}_{\mathcal{A}}}.$$

Proposition 5 implies that, under the pooled filtration,

$$\frac{\mathbb{E}_t^{\mathcal{G}}[dR_{n,t}^E]}{dt} - r = \pi_n^{loc} + \pi_n^{agg} + \Gamma_{n,t}(\mathcal{A})\mathcal{W}_{n,t},$$

where, under the constant-loading benchmark,

$$\pi_n^{loc} \equiv -\gamma_n \left(\frac{\beta_n}{\sigma_Z} a_{n,Z} + \sqrt{A} a_{n,s} \right), \qquad \pi_n^{agg} \equiv -\gamma_n \sqrt{\tau_{\mathcal{A}}} a_{n,\mathcal{A}}.$$

The first component is the premium on the firm-specific local-learning innovations, and the second is the premium on the common public-aggregation innovation.

The wedge term has zero unconditional mean. Since $\mathcal{W}_{n,t} = X_t - \mathbb{E}[X_t | \mathcal{H}_{n,t}(\mathcal{A})]$ is a forecast error relative to $\mathcal{H}_{n,t}(\mathcal{A})$, $\mathbb{E}_t^{\mathcal{H}_n(\mathcal{A})}[\mathcal{W}_{n,t}] = 0$. Because $\Gamma_{n,t}(\mathcal{A})$ is measurable with respect to the local pricing filtration, the law of iterated expectations gives

$$\mathbb{E}[\Gamma_{n,t}(\mathcal{A}) \mathcal{W}_{n,t}] = \mathbb{E}[\Gamma_{n,t}(\mathcal{A}) \mathbb{E}[\mathcal{W}_{n,t} | \mathcal{H}_{n,t}(\mathcal{A})]] = 0.$$

Therefore firm n's unconditional average excess return is its local risk premium:

$$\mu_n^e \equiv \mathbb{E}^{\mathbb{P}} \left[\frac{dR_{n,t}^E}{dt} - r \right] = \pi_n^{loc} + \pi_n^{agg}. \tag{A11}$$

Aggregating across firms gives

$$\bar{\mu}_M^e = \sum_m \omega_m \mu_m^e = \bar{\pi}_M^{loc} + \bar{\pi}_M^{agg}, \tag{A12}$$

where

$$\bar{\pi}_M^{loc} \equiv \sum_m \omega_m \pi_m^{loc}, \qquad \bar{\pi}_M^{agg} \equiv \sum_m \omega_m \pi_m^{agg}.$$

The unconditional CAPM alpha is

$$\alpha_n \equiv \mu_n^e - \hat{\beta}_n \bar{\mu}_M^e.$$

Using the decompositions (A11) and (A12) of unconditional expected returns,

$$\alpha_n = (\pi_n^{loc} - \hat{\beta}_n \bar{\pi}_M^{loc}) + (\pi_n^{agg} - \hat{\beta}_n \bar{\pi}_M^{agg}). \tag{A13}$$

Adding and subtracting $\bar{\pi}_M^{loc}$ in the first bracket gives

$$\pi_n^{loc} - \hat{\beta}_n \bar{\pi}_M^{loc} = \bar{\pi}_M^{loc}(1 - \hat{\beta}_n) + (\pi_n^{loc} - \bar{\pi}_M^{loc}). \tag{A14}$$

The first term on the right-hand side is the flattening effect. The second term is local heterogeneity. The public component of firm n's local risk premium is

$$\pi_n^{agg} = -\gamma_n \sqrt{\tau_{\mathcal{A}}} a_{n,\mathcal{A}},$$

whereas the corresponding market component is

$$\bar{\pi}_M^{agg} = \sum_m \omega_m \pi_m^{agg} = -\sqrt{\tau_{\mathcal{A}}} \sum_m \omega_m \gamma_m a_{m,\mathcal{A}} = -\sqrt{\tau_{\mathcal{A}}} \bar{\gamma}_M \bar{a}_{\mathcal{A}},$$

where the last equality results from defining $\bar{\gamma}_M \equiv \sum_m \omega_m \gamma_m \frac{a_{m,\mathcal{A}}}{\bar{a}_{\mathcal{A}}}$. Using $\hat{\beta}_n = a_{n,\mathcal{A}}/\bar{a}_{\mathcal{A}}$,

$$\hat{\beta}_n \bar{\pi}_M^{agg} = -\sqrt{\tau_{\mathcal{A}}} \bar{\gamma}_M a_{n,\mathcal{A}}.$$

Therefore

$$\pi_n^{agg} - \hat{\beta}_n \bar{\pi}_M^{agg} = -\gamma_n \sqrt{\tau_{\mathcal{A}}} a_{n,\mathcal{A}} + \sqrt{\tau_{\mathcal{A}}} \bar{\gamma}_M a_{n,\mathcal{A}} = \sqrt{\tau_{\mathcal{A}}} a_{n,\mathcal{A}} (\bar{\gamma}_M - \gamma_n). \tag{A15}$$

This is the heterogeneous-pricing effect. Substituting (A14) and (A15) into (A13) gives

$$\alpha_n = \bar{\pi}_M^{loc}(1 - \hat{\beta}_n) + \sqrt{\tau_{\mathcal{A}}} a_{n,\mathcal{A}} (\bar{\gamma}_M - \gamma_n) + (\pi_n^{loc} - \bar{\pi}_M^{loc}).$$

□

A.9 Proof of Proposition 7

From Lemma 1, the steady-state filtering variance solves

$$\bar{\nu}_n = \frac{-\kappa_X + \sqrt{\kappa_X^2 + \Lambda_n \sigma_X^2}}{\Lambda_n}, \qquad \Lambda_n = \frac{\beta_n^2}{\sigma_Z^2} + A + \tau_{\mathcal{A}}.$$

As $\tau_{\mathcal{A}} \to \infty$, $\Lambda_n \to \infty$ with $\Lambda_n/\tau_{\mathcal{A}} \to 1$. Therefore

$$\lim_{\tau_{\mathcal{A}} \to \infty} \bar{\nu}_n = 0, \qquad \lim_{\tau_{\mathcal{A}} \to \infty} \sqrt{\tau_{\mathcal{A}}} \bar{\nu}_n = \sigma_X. \tag{A16}$$

From Lemma 3, the projected local risk price is $\gamma_n = \frac{-\sigma_X \lambda_X}{\kappa_X + \bar{\nu}_n \Lambda_n}$. Multiplying numerator and denominator by $1/\sqrt{\tau_{\mathcal{A}}}$ gives

$$\sqrt{\tau_{\mathcal{A}}} \gamma_n = \frac{-\sigma_X \lambda_X}{\kappa_X/\sqrt{\tau_{\mathcal{A}}} + \bar{\nu}_n \Lambda_n/\sqrt{\tau_{\mathcal{A}}}}.$$

Since

$$\frac{\bar{\nu}_n \Lambda_n}{\sqrt{\tau_{\mathcal{A}}}} = \sqrt{\tau_{\mathcal{A}}} \bar{\nu}_n \frac{\Lambda_n}{\tau_{\mathcal{A}}} \to \sigma_X,$$

it follows that $\gamma_n \to 0$ and

$$\lim_{\tau_{\mathcal{A}} \to \infty} \sqrt{\tau_{\mathcal{A}}} \gamma_n = -\lambda_X. \tag{A17}$$

The return loadings are

$$a_{n,Z} = \sigma_Z \frac{v_Z}{v_n} + \frac{\beta_n \bar{\nu}_n}{\sigma_Z} \frac{v_{\hat{X}}}{v_n}, \qquad a_{n,s} = \bar{\nu}_n \sqrt{A} \frac{v_{\hat{X}}}{v_n}, \qquad a_{n,\mathcal{A}} = \bar{\nu}_n \sqrt{\tau_{\mathcal{A}}} \frac{v_{\hat{X}}}{v_n}.$$

As $\tau_{\mathcal{A}} \to \infty$, the value function converges to its full-information limit $v_n \to v_n^\infty$ with finite derivatives. Substituting (A16) gives

$$a_{n,s} \to 0, \qquad a_{n,Z} \to \sigma_Z \frac{v_{n,Z}^\infty}{v_n^\infty}, \qquad a_{n,\mathcal{A}} \to \sigma_X \frac{v_{n,X}^\infty}{v_n^\infty}.$$

Thus the private-signal exposure vanishes, while the cash-flow and public-aggregation exposures converge to finite firm-specific constants.

The local-learning premium is

$$\pi_n^{loc} = -\gamma_n \left(\frac{\beta_n}{\sigma_Z} a_{n,Z} + \sqrt{A} a_{n,s} \right).$$

As $\tau_{\mathcal{A}} \to \infty$, the bracket converges to a finite firm-specific constant, and $\gamma_n \to 0$. Hence $\pi_n^{loc} \to 0$ for every firm n, and therefore $\bar{\pi}_M^{loc} \to 0$. Since $\hat{\beta}_n = a_{n,\mathcal{A}}/\bar{a}_{\mathcal{A}}$ converges to a finite limit when the limiting market exposure is nonzero, the flattening effect vanishes:

$$\lim_{\tau_{\mathcal{A}} \to \infty} \bar{\pi}_M^{loc}(1 - \hat{\beta}_n) = 0.$$

The same argument implies that local heterogeneity vanishes:

$$\lim_{\tau_{\mathcal{A}} \to \infty} (\pi_n^{loc} - \bar{\pi}_M^{loc}) = 0.$$

The scaling relation (A17) holds firm by firm with the same limit $-\lambda_X$. Since $\bar{\gamma}_M = \sum_m \omega_m \gamma_m \frac{a_{m,\mathcal{A}}}{\bar{a}_{\mathcal{A}}}$ is an exposure-weighted market average with weights that sum to one, the same scaling gives

$$\lim_{\tau_{\mathcal{A}} \to \infty} \sqrt{\tau_{\mathcal{A}}} \bar{\gamma}_M = -\lambda_X.$$

Subtracting (A17),

$$\lim_{\tau_{\mathcal{A}} \to \infty} \sqrt{\tau_{\mathcal{A}}} (\bar{\gamma}_M - \gamma_n) = 0.$$

Because $a_{n,\mathcal{A}}$ converges to a finite limit, the heterogeneous-pricing effect vanishes:

$$\lim_{\tau_{\mathcal{A}} \to \infty} \sqrt{\tau_{\mathcal{A}}} a_{n,\mathcal{A}} (\bar{\gamma}_M - \gamma_n) = 0.$$

Combining the three limits in Proposition 6 gives $\lim_{\tau_{\mathcal{A}} \to \infty} \alpha_n = 0$. □

B Appendix: Numerical Solution Method

I compute the normalized firm value $v(Z, \hat{X}^{\mathcal{H}}) = v^{AIP}(Z, \hat{X}^{\mathcal{H}}) + g(Z, \hat{X}^{\mathcal{H}})$ on a discrete two-dimensional state space. The grid covers ± 2.5 unconditional standard deviations for profitability Z and ± 4.0 standard deviations for the local belief $\hat{X}^{\mathcal{H}}$. I calculate the affine assets-in-place value v^{AIP} directly and solve the remaining partial differential equation for the option-to-scale component $g(Z, \hat{X}^{\mathcal{H}})$ using finite differences.

I approximate first derivatives using a first-order upwind scheme, applying forward or backward differences depending on the sign of the local drift to guarantee numerical stability. Second derivatives rely on standard central differences. To preserve the sparsity of the transition matrix, I compute the cross-derivative $g_{Z\hat{X}}$ explicitly as a source term using the candidate value function from the previous iteration.

Because the optimal investment rule features kinks from irreversibility and external financing constraints, direct nonlinear root-finding is unstable. I instead apply an implicit false-transient algorithm. I introduce an artificial time step Δ and evaluate the optimal investment policy using the value function from the prior step. This yields a sequence of linear sparse-matrix systems. To ensure global convergence, I use a multi-gear schedule: the algorithm starts with a small pseudo-time step ($\Delta = 0.05$) to absorb initial gradients and progressively increases it (up to $\Delta = 5.0$) as the solution approaches the steady state. I apply an explicit relaxation parameter to update the value function, dampening oscillations across steps. The algorithm iterates until the maximum absolute change in the grid falls below a tolerance of 10^{-6}.