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

Daniel Andrei*

November 8, 2026

Abstract

Local efficiency need not imply cross-sectional efficiency. I develop a model in which investors specialize in individual firms and learn from local information about fundamentals that are correlated across firms. Investors are risk-neutral and rational Bayesians. Each stock's return is unpredictable from its own history, yet comparing firms (e.g., ranking them on characteristics) predicts returns. The reason is an information aggregation wedge: the gap between what all firms' information implies about a firm's fundamentals and what its own investors believe. The wedge survives arbitrage because processing information across firms is costly. The model explains why value and profitability complement each other, why firms migrate across value and growth portfolios, and why anomaly returns concentrate on announcement days. U.S. stock and analyst data confirm these predictions: firms that are cheap and profitable relative to industry peers (but not to their own history) earn higher returns, especially in announcement months. These firms have lower growth forecasts, which analysts later revise upward.

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

I thank participants at the University of Lausanne, the University of Neuchâtel, and the SAFE Asset Pricing Workshop for helpful comments. I am also grateful to Lorenzo Bretscher, Michael Hasler, and Christian Heyerdahl-Larsen (discussant) for their comments and suggestions.

*HEC Lausanne, UNIL, Extranef 231, CH-1015 Lausanne, daniel.andrei@unil.ch, danielandrei.info.

1 Introduction

Financial markets display substantial local efficiency: each stock's price reflects the information held by the investors who trade it.[1] I argue that the same markets can display substantial cross-sectional *in*efficiency: each stock's price fails to reflect information contained in other firms' prices and fundamentals.

How can a market be efficient stock by stock, yet inefficient in the cross-section? The answer goes back to Hayek (1945): information is dispersed across firms, and each firm's investors price its stock using local information—yet no local investor observes the information across all firms.

Think of an economy with many firms whose fundamentals are related. Investors do not observe these fundamentals directly, nor do they know the links across firms. Instead, they specialize in individual firms and learn from noisy local information. Learning leaves unavoidable forecast errors, even for rational Bayesians. Because investors use their local information efficiently, they cannot predict how those errors will be corrected. The paper's main point is that comparing firms (e.g., ranking them on observable characteristics) supplies additional information. One firm's local information helps assess another's prospects because their fundamentals are related—even if no one knows the exact link. Comparisons can therefore predict which firms will receive favorable or unfavorable surprises relative to their investors' expectations. As new information arrives, investors revise their assessments and prices adjust. Returns are thus predictable in the cross-section, even though each firm's return is unpredictable from its own information.

For example, consider two firms, L and H, whose fundamentals tend to move together. Suppose firm L's investors observe low signals and form a low assessment, while firm H's investors observe high signals and form a high assessment. Taken separately, each assessment is unbiased: firm L's low signals could reflect weak fundamentals or negative noise, and local information alone cannot separate the two. Side by side, however, firm

[1] I define "local efficiency" as efficiency under a stock's own pricing filtration. Stock prices incorporate firm-specific information from disclosures and order flow that anticipates fundamentals and guides corporate investment (Ball and Brown, 1968; Roll, 1988; Hasbrouck, 1991; Chen, Goldstein, and Jiang, 2007; Bai, Philippon, and Savov, 2016). Predicting returns in the time series is difficult: predictable variation in individual stock returns is small relative to firm-specific shocks (Fama, 1991; Vuolteenaho, 2002; Gu, Kelly, and Xiu, 2020). Time-series predictors of market returns perform poorly out of sample (Goyal and Welch, 2008; Engelberg, McLean, Pontiff, and Ringgenberg, 2023; Goyal, Welch, and Zafirov, 2024).

H's high signals suggest negative noise at firm L, because the two firms' fundamentals are related. By the same logic, firm L's low signals suggest positive noise at firm H. Given both firms' information, firm L's assessment is too low relative to firm H's. Neither firm's investors can detect this, because they observe only their own firm. But as new information arrives, investors revise firm L's assessment upward relative to firm H's on average. Buying firm L's stock and selling firm H's earns a positive expected return.

I formalize this point in a model where local efficiency holds by construction: investors who specialize in a firm use all their information about it efficiently. Investors are risk-neutral, so each stock is priced to earn the risk-free rate under its local filtration. Risk compensation therefore cannot explain cross-sectional variation in expected returns. The economy has N firms. Local investors observe their firm's profitability and noisy signals and price its stock. Each firm's profitability has a persistent unobservable drift. These drifts are positively correlated across firms (e.g., through macroeconomic or industry links).

Now consider pooling local information across all N firms. No local investor observes all this information, yet pooling it lets me define a theoretical gap: the *information aggregation wedge*. For each firm, this wedge equals the pooled belief about its drift minus its own investors' belief. Conditional on pooled information, each firm's expected excess return is proportional to its wedge. The wedges themselves are unobservable. Yet because drifts are positively correlated, firms with lower local beliefs than their peers' tend to have higher wedges and earn higher future returns. Comparing local beliefs across firms thus predicts returns without observing the wedges or knowing how drifts are correlated, even under arbitrary firm heterogeneity.

In practice, local beliefs are not directly observable either, but firm characteristics proxy for them. A low market-to-book ratio or conservative investment can proxy for low local beliefs relative to other firms. Sorting on either characteristic thus captures the same source of predictability. This suggests that the "factor zoo" (Cochrane, 2011) consists of different proxies for the same information aggregation wedge, rather than a collection of separate risks or behavioral mistakes.[2]

[2]To build intuition, consider a "trip-let" (Hofstadter, 1979): a 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 different objects. In this model, the information aggregation wedge is the trip-let, and the distinct anomalies are its shadows.

Why does this predictability survive arbitrage? I introduce arbitrageurs who pay an information-processing cost to observe local information across all firms and trade in all of them. Local investors observe these trades only through noisy order flow. As more arbitrageurs enter, order flow becomes more informative, local prices move toward pooled valuations, and trading profits per arbitrageur shrink toward zero. For any positive information-processing cost, free entry stops short of full aggregation: dispersed information is only partially revealed, and a residual wedge survives in equilibrium. If trading eliminated the wedge, arbitrageurs would earn zero gross profits and could not recover their information-processing cost (Grossman and Stiglitz, 1980).

Arbitrageurs cannot eliminate the wedge for free by conditioning only on public characteristics. Public characteristics mix local beliefs with other fundamentals, such as current profitability. Even observing every firm's current belief cannot replace filtering the information history across all firms (Timmermann, 2006). Strategies using public characteristics therefore capture only a projection of the wedge and leave the rest intact.

The model yields several predictions for the cross-section of returns. First, it explains why value and profitability complement each other (Novy-Marx, 2013). In the model, a firm's low market-to-book ratio can reflect weak current profitability or a low local belief about its drift. Controlling for current profitability separates these two forces and isolates the local belief. Value and profitability therefore predict cross-sectional returns better together than either alone. An alternative argument is that a profitable firm's low valuation implies a high discount rate (Novy-Marx, 2013). That channel is absent here because investors are risk-neutral. Instead, the complementarity comes from learning and from comparing firms.

Second, cross-sectional return patterns depend on how closely firms are linked (Cohen and Frazzini, 2008). Within an industry, firms' drifts are more strongly correlated than across industries. Thus, comparing a firm with its industry peers reveals more about its wedge. Value and profitability therefore work mainly within industries (Novy-Marx, 2013; Kim, Kim, and Lee, 2025). At the same time, each firm's investors price local news immediately, so no stock shows momentum on its own. Yet good news at one firm predicts higher future returns for its correlated peers. Bundling peers into an industry

portfolio can generate industry momentum: one stock's gain today is followed by peer gains tomorrow (Moskowitz and Grinblatt, 1999). When peers also differ in information precision, this cross-firm predictability is asymmetric. Returns of firms with more precise local information lead the returns of their less-informed peers (Hou, 2007).

Third, value and growth are temporary states whose returns materialize as new information arrives. Because each stock's expected return equals the risk-free rate conditional on its own local information, no firm is a permanent high-return type. Instead, a stock enters a value portfolio when its local belief lies below its peers' beliefs. As new information arrives, local investors tend to revise their beliefs upward for value stocks relative to growth stocks. Value stocks therefore experience more favorable earnings surprises (La Porta, Lakonishok, Shleifer, and Vishny, 1997) and migrate toward growth portfolios as their valuations recover (Fama and French, 2007).

Fourth, anomaly returns concentrate on scheduled announcement days, such as earnings announcements (Engelberg, McLean, and Pontiff, 2018). To show this, I allow signal precision to vary over time and rise on announcement days. Because local investors know this schedule in advance, local prices already reflect it and local efficiency holds. Yet local information cannot predict the direction of the announcement surprise. On announcement days, local investors place greater weight on the more precise signal, generating larger belief revisions and closing more of the wedge. Anomaly returns are therefore higher on announcement days.

I test the model's mechanism on U.S. stocks from 1963 to 2024. Monthly cross-sectional regressions confirm that lower market-to-book and higher profitability jointly predict higher future returns (Novy-Marx, 2013). I then split each characteristic into its industry average and its within-industry deviation. Predictability comes from within-industry comparisons, while industry averages have no predictive power. A firm's own prior five-year characteristics have none either—as local efficiency implies. A stock therefore earns higher returns when it is cheap and profitable relative to its industry peers today, rather than relative to its own past. Next, I test the belief channel using consensus analyst forecasts of long-term earnings growth from I/B/E/S. Within an industry, firms with high profitability and low valuation have lower growth forecasts than their peers. I

show that analysts revise these forecasts upward relative to forecasts for industry peers over the following year. These revisions are consistent with the model's error-correction mechanism. Using quarterly earnings announcement dates from Compustat, I also show that return predictability nearly doubles in announcement months. Higher firm-level discount rates could account for cheap, profitable firms' higher returns, but do not by themselves explain why return predictability is peer-relative; why the same characteristics predict subsequent revisions in growth forecasts; and why predictability concentrates in earnings announcement months.

This paper proposes an alternative to risk-based and behavioral explanations of cross-sectional return predictability. Risk-based explanations attribute this predictability to differences in risk exposure (Fama and French, 1993; Zhang, 2005), whereas behavioral explanations attribute it to systematic biases in how investors form beliefs (De Bondt and Thaler, 1985; Lakonishok, Shleifer, and Vishny, 1994; Bordalo, Gennaioli, La Porta, and Shleifer, 2019). Here, investors are risk-neutral and rational Bayesians. Predictability comes instead from incomplete aggregation of information across firms.

Lewellen and Shanken (2002) show that rational Bayesian learning can generate return predictability under risk neutrality, but through a different mechanism. In their model, all investors observe the same information, and predictability arises from parameter uncertainty. That predictability appears in the time series and the cross-section, and no investor can exploit it. I shut down parameter uncertainty: each firm's investors know the true parameters governing its fundamentals. Risk-neutral pricing then makes each stock's return unpredictable from its own history. If all investors in my model shared the pooled information across firms—as in Lewellen and Shanken's model—returns would be unpredictable, although investors would still learn about unobservable drifts. By contrast, information held by one firm's investors helps predict another firm's future surprises. Comparisons across firms therefore supply information that local prices miss, and arbitrageurs have an incentive to aggregate information across firms.

A related literature studies economies in which investors specialize in subsets of firms, starting with Merton (1987).[3] In Merton's model, however, investors do not learn.

[3]See also Brennan and Cao (1997) and Van Nieuwerburgh and Veldkamp (2009) for models with different information sets across investors.

Specialization affects expected returns only because risk-averse investors bear idiosyncratic risk. I shut down that channel by assuming risk neutrality. Predictability nonetheless arises because peer information predicts surprises that local investors cannot anticipate.

The paper also contributes to the literature on information aggregation and limits to arbitrage. In noisy rational expectations models, private information is dispersed across investors who trade the same assets (Grossman and Stiglitz, 1980; Admati, 1985; Albagli, Hellwig, and Tsyvinski, 2024). Here, each firm's investors share its local information and price its stock efficiently, while information is dispersed across firms. Traditional limits to arbitrage rely on funding constraints or noise-trader risk (Shleifer and Vishny, 1997). Da, Nagel, and Xiu (2024) show that estimation error limits arbitrage when alphas are weak and rare. I focus instead on an informational limit: filtering histories across all firms is costly, and public characteristics capture only a projection of the wedge.

This setting also parallels Andrei, Cujean, and Wilson (2023) on the informational gap between investors and the econometrician (Hansen and Richard, 1987). When estimating betas from time series, the econometrician conditions down with less depth than investors and mismeasures risk. When sorting stocks on characteristics, the econometrician conditions across firms with more breadth than any single firm's investors and uncovers predictability. Finally, the model gives a counterpart to Samuelson's dictum that the stock market is "micro efficient" but "macro inefficient" (Jung and Shiller, 2005). Here, prices are efficient within each firm's local filtration yet inefficient across the pooled cross-section.

Section 2 presents the model and characterizes equilibrium information aggregation. Section 3 derives the empirical implications, Section 4 tests the mechanism on U.S. stock and analyst forecast data, and Section 5 concludes.

2 Model

Consider an economy with N firms, indexed by $n \in \{1, \ldots, N\}$. Firm n's operating profitability per unit of capital, $Z_{n,t}$, evolves as

$$dZ_{n,t} = (\mu_{n,t} - \kappa_{Z,n} Z_{n,t})dt + \sigma_{Z,n} dW^Z_{n,t},$$

where $\kappa_{Z,n} > 0$ is the speed of mean reversion and $\sigma_{Z,n} > 0$ is the volatility of firm-specific cash-flow shocks.

The latent profitability drift, $\mu_{n,t}$, follows

$$d\mu_{n,t} = \kappa_{\mu,n}(\bar{\mu}_n - \mu_{n,t})dt + \sigma_{\mu,n}dW^{\mu}_{n,t},$$

with mean-reversion speed $\kappa_{\mu,n} > 0$, long-run mean $\bar{\mu}_n$, and volatility $\sigma_{\mu,n} > 0$. The processes W^Z_n and W^{μ}_n are standard Brownian motions satisfying

$$d\langle W^Z_n, W^Z_m\rangle_t = 0 \quad (n \neq m), \qquad d\langle W^Z_n, W^{\mu}_m\rangle_t = 0, \qquad d\langle W^{\mu}_n, W^{\mu}_m\rangle_t = \rho_{nm}dt,$$

where $(\rho_{nm})^N_{n,m=1}$ is a symmetric, positive semidefinite correlation matrix with unit diagonal. Realized cash-flow shocks $dW^Z_{n,t}$ are purely idiosyncratic, whereas drift innovations $dW^{\mu}_{n,t}$ are correlated across firms through shared macroeconomic or industry links.

Firm n's local investors (hereafter "specialists") price its stock and know its structural parameters $(\kappa_{Z,n}, \sigma_{Z,n}, \kappa_{\mu,n}, \bar{\mu}_n, \sigma_{\mu,n})$. Alongside profitability $Z_{n,t}$, firm n's specialists observe a continuous flow of local information (e.g., management guidance, specialized analyst research) summarized by a signal $O_{n,t}$ with precision $\Lambda_n > 0$:

$$dO_{n,t} = \mu_{n,t}dt + \Lambda_n^{-1/2}dW^O_{n,t},$$

where W^O_n are standard Brownian motions, mutually independent across firms and independent of W^Z and W^{μ}. Firm n's local information set is the filtration generated by its own profitability and signal histories,

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

Conditional on $\mathcal{H}_{n,t}$, firm n's specialists form beliefs about $\mu_{n,t}$. Standard filtering theory (Liptser and Shiryaev, 2001) implies that the conditional distribution of $\mu_{n,t}$ is Gaussian with posterior mean $\mu^{\mathcal{H}}_{n,t} \equiv \mathbb{E}[\mu_{n,t} \mid \mathcal{H}_{n,t}]$ and posterior variance $\nu_{n,t} \equiv \mathrm{Var}(\mu_{n,t} \mid \mathcal{H}_{n,t})$. I assume the economy is in its jointly Gaussian stationary regime, where posterior uncertainty equals its steady state $\bar{\nu}_n$ (Dumas, Kurshev, and Uppal, 2009; Andrei, Mann, and Moyen, 2019):

$$\bar{\nu}_n = \frac{u_n - \kappa_{\mu,n}}{\Lambda_n + \sigma^{-2}_{Z,n}}, \qquad u_n \equiv \sqrt{\kappa^2_{\mu,n} + (\Lambda_n + \sigma^{-2}_{Z,n})\sigma^2_{\mu,n}}. \tag{1}$$

The filter $\mu_{n,t}^{\mathcal{H}}$ evolves as

$$d\mu_{n,t}^{\mathcal{H}} = \kappa_{\mu,n}(\bar{\mu}_n - \mu_{n,t}^{\mathcal{H}})dt + \bar{\nu}_n \left(\sigma_{Z,n}^{-2} dZ_{n,t}^{\mathcal{H}} + \Lambda_n dO_{n,t}^{\mathcal{H}}\right), \tag{2}$$

where $dZ_{n,t}^{\mathcal{H}} \equiv dZ_{n,t} - (\mu_{n,t}^{\mathcal{H}} - \kappa_{Z,n} Z_{n,t})dt$ and $dO_{n,t}^{\mathcal{H}} \equiv dO_{n,t} - \mu_{n,t}^{\mathcal{H}} dt$ denote the cash-flow and signal surprises relative to $\mathcal{H}_{n,t}$. The local forecast error $e_{n,t} \equiv \mu_{n,t} - \mu_{n,t}^{\mathcal{H}}$ satisfies $\mathbb{E}[e_{n,t} \mid \mathcal{H}_{n,t}] = 0$ and follows:

$$de_{n,t} = -u_n e_{n,t} dt + \sigma_{\mu,n} dW_{n,t}^{\mu} - \bar{\nu}_n \left(\sigma_{Z,n}^{-1} dW_{n,t}^{Z} + \Lambda_n^{1/2} dW_{n,t}^{O}\right). \tag{3}$$

The forecast error $e_{n,t}$ mean-reverts faster than the drift $\mu_{n,t}$ itself ($u_n > \kappa_{\mu,n}$) because Bayesian updating corrects errors at rate $u_n - \kappa_{\mu,n} > 0$ as new information arrives.

I assume that firm n's specialists neither know nor learn other firms' parameters or the correlation matrix (ρ_{nm}). Even if peer profitability $Z_{m,t}$ is public, they cannot use it to learn about $\mu_{n,t}$: doing so requires filtering $\mu_{m,t}$ using peer m's parameters and ρ_{nm}. Realized cash flows do not reveal ρ_{nm} through instantaneous covariation, as $d\langle Z_n, Z_m \rangle_t = 0$ and $\mu_{m,t}$ is latent. Peer signals $O_{m,t}$ are private, so firm n's specialists cannot recover them even by pooling all public profitability series. In Section 2.4, I relax this segmentation by introducing arbitrageurs who can process information across firms at a cost.

2.1 The Information Aggregation Wedge

I now compare local beliefs against the information from the entire cross-section. Let $\mathcal{G}_t \equiv \bigvee_{m=1}^{N} \mathcal{H}_{m,t}$ denote the pooled filtration and $\mu_{n,t}^{\mathcal{G}} \equiv \mathbb{E}[\mu_{n,t} \mid \mathcal{G}_t]$ denote the pooled belief about firm n's drift. The *information aggregation wedge* equals the pooled belief minus the local belief:

$$\mathcal{W}_{n,t} \equiv \mu_{n,t}^{\mathcal{G}} - \mu_{n,t}^{\mathcal{H}}.$$

Because $\mathcal{H}_{n,t} \subseteq \mathcal{G}_t$, the law of iterated expectations implies $\mathbb{E}[\mu_{n,t}^{\mathcal{G}} \mid \mathcal{H}_{n,t}] = \mu_{n,t}^{\mathcal{H}}$, and therefore $\mathbb{E}[\mathcal{W}_{n,t} \mid \mathcal{H}_{n,t}] = 0$. That is, conditional on firm n's local information, the wedge has mean zero and is unpredictable.

For any two firm-level variables $(x_{n,t})_{n=1}^{N}$ and $(y_{n,t})_{n=1}^{N}$, define the cross-sectional

covariance

$$\mathrm{Cov}_{CS}(x_{n,t}, y_{n,t}) \equiv \frac{1}{N}\sum_{n=1}^{N}(x_{n,t} - \bar{x}_t)(y_{n,t} - \bar{y}_t), \tag{4}$$

where $\bar{x}_t \equiv N^{-1}\sum_{n=1}^{N} x_{n,t}$ and $\bar{y}_t \equiv N^{-1}\sum_{n=1}^{N} y_{n,t}$. The following proposition characterizes the expected cross-sectional covariance between local beliefs and wedges.

Proposition 1 (Belief–Wedge Covariance)**.** *The expected cross-sectional covariance between local beliefs and wedges is*

$$\mathbb{E}\left[\mathrm{Cov}_{CS}(\mu^{\mathcal{H}}_{n,t}, \mathcal{W}_{n,t})\right] = -\frac{1}{N^2}\sum_{n=1}^{N}\sum_{m\neq n}\rho_{nm}\sigma_{\mu,n}\sigma_{\mu,m}\frac{u_n - \kappa_{\mu,n}}{(\kappa_{\mu,n} + u_m)(u_n + u_m)}. \tag{5}$$

Because $u_n > \kappa_{\mu,n}$, each ρ_{nm} has a strictly positive multiplier inside the sum in (5). Suppose drifts are nonnegatively correlated across firms ($\rho_{nm} \geq 0$ for all $n \neq m$) and $\rho_{nm} > 0$ for at least one pair. Then,

$$\mathbb{E}\left[\mathrm{Cov}_{CS}(\mu^{\mathcal{H}}_{n,t}, \mathcal{W}_{n,t})\right] < 0. \tag{6}$$

In the cross-section, firms with lower local beliefs than their peers' tend to have higher wedges. This negative covariance does not arise mechanically from $-\mu^{\mathcal{H}}_{n,t}$ appearing in $\mathcal{W}_{n,t}$: for a given firm, optimal filtering implies $\mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, \mathcal{W}_{n,t}) = 0$ in the time series. Equation (6) requires both Bayesian updating ($u_n > \kappa_{\mu,n}$) and cross-firm correlation ($\rho_{nm} > 0$). If either is missing, (5) gives zero. Moreover, (6) holds even though each firm's specialists estimate a firm-specific drift $\mu_{n,t}$ rather than a common factor.

2.2 Valuation and Return Predictability

Firm n has capital $K_{n,t}$, which accumulates according to

$$dK_{n,t} = (i_{n,t} - \delta_n)K_{n,t}dt,$$

where $i_{n,t}$ is the investment rate and $\delta_n > 0$ is the depreciation rate. Firm n's manager shares the local information set $\mathcal{H}_{n,t}$ with firm n's specialists. Both are risk-neutral and discount cash flows at the constant rate $r > 0$. The manager maximizes firm value over $i_{n,t}$:

$$V_n(K_{n,t}, Z_{n,t}, \mu^{\mathcal{H}}_{n,t}, \bar{\nu}_n) = \max_{\{i_{n,u}\}_{u\geq t}} \mathbb{E}\left[\int_t^\infty e^{-r(u-t)}K_{n,u}\Big(Z_{n,u} - \Psi_n(i_{n,u}, Z_{n,u})\Big)du \,\middle|\, \mathcal{H}_{n,t}\right],$$

where $\Psi_n(i_{n,t}, Z_{n,t})$ denotes investment and adjustment costs per unit of capital. Because cash flows and investment costs scale linearly with capital (Hayashi, 1982), the value function is homogeneous of degree one in $K_{n,t}$:

$$V_n(K_{n,t}, Z_{n,t}, \mu^{\mathcal{H}}_{n,t}, \bar{\nu}_n) = K_{n,t} v_n(Z_{n,t}, \mu^{\mathcal{H}}_{n,t}, \bar{\nu}_n). \tag{7}$$

The market-to-book ratio $v_{n,t} \equiv V_{n,t}/K_{n,t} = v_n(Z_{n,t}, \mu^{\mathcal{H}}_{n,t}, \bar{\nu}_n)$ therefore depends only on $Z_{n,t}$ and $\mu^{\mathcal{H}}_{n,t}$, given $\bar{\nu}_n$. I assume that $v_n(\cdot, \cdot, \bar{\nu}_n)$ is twice continuously differentiable and strictly positive, with $\partial v_n/\partial Z_{n,t} > 0$ and $\partial v_n/\partial \mu^{\mathcal{H}}_{n,t} > 0$.

Let $dR^E_{n,t} \equiv \Big[dV_{n,t} + K_{n,t}\Big(Z_{n,t} - \Psi_n(i_{n,t}, Z_{n,t})\Big)dt\Big]/V_{n,t}$ denote firm n's total equity return. Because risk-neutral specialists price firm n's stock under $\mathcal{H}_{n,t}$, its expected return conditional on local information equals the risk-free rate:

$$\frac{\mathbb{E}[dR^E_{n,t} \mid \mathcal{H}_{n,t}]}{dt} = r. \tag{8}$$

Equation (8) is *local efficiency*: no variable in $\mathcal{H}_{n,t}$ predicts excess returns in firm n's time series. Under the pooled filtration $\mathcal{G}_t$, however, the surprises in (2) have conditional mean $\mathbb{E}[dZ^{\mathcal{H}}_{n,t} \mid \mathcal{G}_t] = \mathbb{E}[dO^{\mathcal{H}}_{n,t} \mid \mathcal{G}_t] = \mathcal{W}_{n,t}dt$. Applying Itô's lemma to (7) yields:

Proposition 2 (Return Predictability)**.** *The expected return conditional on the pooled filtration $\mathcal{G}_t$ is*

$$\frac{\mathbb{E}[dR^E_{n,t} \mid \mathcal{G}_t]}{dt} = r + \Gamma_{n,t}\mathcal{W}_{n,t}, \tag{9}$$

where

$$\Gamma_{n,t} \equiv \frac{1}{v_n}\left[\frac{\partial v_n}{\partial Z_{n,t}} + (u_n - \kappa_{\mu,n})\frac{\partial v_n}{\partial \mu^{\mathcal{H}}_{n,t}}\right] > 0. \tag{10}$$

Equations (8) and (9) describe the same stock return under two information sets. Given firm n's local information, excess returns are unpredictable. Given the pooled cross-section, expected excess returns are proportional to the wedge $\mathcal{W}_{n,t}$.

2.3 Beliefs, Characteristics, and the Cross-Section of Returns

Equations (6) and (9) suggest that local beliefs predict returns in the cross-section, even without observing the wedge $\mathcal{W}_{n,t}$.

Proposition 3 (Cross-Sectional Predictability)**.** *If $\rho_{nm} \geq 0$ for all $n \neq m$, with $\rho_{nm} > 0$ for at least one pair, and $\mathbb{E}[\Gamma_{n,t}^2] < \infty$ for all n, then*

$$\mathbb{E}\left[\mathrm{Cov}_{CS}\left(\mu_{n,t}^{\mathcal{H}}, \frac{dR_{n,t}^E}{dt} - r\right)\right] = -\frac{1}{N^2}\sum_{n\neq m}\frac{\rho_{nm}\sigma_{\mu,n}\sigma_{\mu,m}(u_n - \kappa_{\mu,n})}{(\kappa_{\mu,n} + u_m)(u_n + u_m)}\mathbb{E}[\Gamma_{m,t}] < 0. \quad (11)$$

Proposition 3 establishes a negative expected cross-sectional covariance between local beliefs and returns. In the time series, a firm's local belief does not predict its own return ($\mathrm{Cov}(\mu_{n,t}^{\mathcal{H}}, dR_{n,t}^E/dt - r) = 0$), but comparing firm n with its peers predicts returns in the cross-section. The sign holds even though $\Gamma_{n,t}$ is endogenous and firms differ in every model parameter.

I extend Proposition 3 to any differentiable characteristic $x_{n,t} = x_n(Z_{n,t}, \mu_{n,t}^{\mathcal{H}}, \bar{\nu}_n)$.

Corollary 3.1 (Characteristic Predictability)**.** *If $\mathbb{E}[\Gamma_{n,t}^2] < \infty$ for all n, then for any continuously differentiable characteristic $x_{n,t} = x_n(Z_{n,t}, \mu_{n,t}^{\mathcal{H}}, \bar{\nu}_n)$ with $\mathbb{E}[x_{n,t}^4] < \infty$ and square-integrable partial derivatives,*

$$\begin{aligned}&\mathbb{E}\left[\mathrm{Cov}_{CS}\left(x_{n,t}, \frac{dR_{n,t}^E}{dt} - r\right)\right]\\ &= -\frac{1}{N^2}\sum_{n\neq m}\frac{\rho_{nm}\sigma_{\mu,n}\sigma_{\mu,m}}{\kappa_{\mu,n} + u_m}\left\{\frac{u_n - \kappa_{\mu,n}}{u_n + u_m}\mathbb{E}\left[\Gamma_{m,t}\frac{\partial x_{n,t}}{\partial \mu_{n,t}^{\mathcal{H}}}\right] + \frac{1}{\kappa_{Z,n} + u_m}\mathbb{E}\left[\Gamma_{m,t}\frac{\partial x_{n,t}}{\partial Z_{n,t}}\right]\right\}. \quad (12)\end{aligned}$$

A characteristic $x_{n,t}$ depends on two local state variables: the belief $\mu_{n,t}^{\mathcal{H}}$ and profitability $Z_{n,t}$. The two terms in curly braces in (12) weight the sensitivities of $x_{n,t}$, $\partial x_{n,t}/\partial \mu_{n,t}^{\mathcal{H}}$ and $\partial x_{n,t}/\partial Z_{n,t}$, by peer return sensitivity $\Gamma_{m,t} > 0$. Because all weights are positive when $\rho_{nm} > 0$, any characteristic that increases in beliefs and profitability—such as the market-to-book ratio $v_n(Z_{n,t}, \mu_{n,t}^{\mathcal{H}}, \bar{\nu}_n)$—has a negative expected cross-sectional covariance with returns, generating a value premium. Finally, the pooled belief $\mu_{n,t}^{\mathcal{G}}$ does not appear in (12), so cross-sectional predictability does not require solving the pooled filtering problem.

Section 3.1 further decomposes equation (12) from Corollary 3.1, but a special case builds intuition first. Consider homogeneous firms sharing identical parameters (κ_Z, σ_Z, $\kappa_\mu, \bar{\mu}, \sigma_\mu, \delta, \Lambda$) and cost function Ψ, with equicorrelated drift innovations, $\rho_{nm} = \rho \in (0, 1)$ for all $n \neq m$. This case is only for interpretation, as the same intuition holds in the general model. The next proposition shows the implications for Fama–MacBeth cross-

sectional return regressions. It also verifies that predictability neither relies on parameter heterogeneity nor vanishes as $N \to \infty$.

Proposition 4 (Homogeneous Firms)**.** *Suppose $N \geq 2$ firms share identical parameters $(\kappa_Z, \sigma_Z, \kappa_\mu, \bar{\mu}, \sigma_\mu, \delta, \Lambda)$ and cost function Ψ, so that $u_n = u$, with equicorrelated drift innovations $\rho_{nm} = \rho \in (0,1)$ for all $n \neq m$ and square-integrable valuation sensitivities. Then:*

(a) The population cross-sectional regression slope of excess returns on local beliefs $\mu^{\mathcal{H}}_{n,t}$ is

$$b_\mu \equiv \frac{\mathbb{E}\left[\mathrm{Cov}_{CS}\left(\mu^{\mathcal{H}}_{n,t}, \frac{dR^E_{n,t}}{dt} - r\right)\right]}{\mathbb{E}\left[\mathrm{Var}_{CS}(\mu^{\mathcal{H}}_{n,t})\right]} = b_{\mathcal{W}}\mathbb{E}[\Gamma_{n,t}] < 0, \tag{13}$$

where the regression slope of the wedge $\mathcal{W}_{n,t}$ on $\mu^{\mathcal{H}}_{n,t}$ is

$$b_{\mathcal{W}} \equiv \frac{\mathbb{E}\left[\mathrm{Cov}_{CS}(\mu^{\mathcal{H}}_{n,t}, \mathcal{W}_{n,t})\right]}{\mathbb{E}\left[\mathrm{Var}_{CS}(\mu^{\mathcal{H}}_{n,t})\right]} = -\frac{\rho\kappa_\mu}{(1-\rho)u + \rho\kappa_\mu} \in (-1, 0). \tag{14}$$

The slope $b_{\mathcal{W}}$ vanishes when drifts are uncorrelated ($\rho = 0$) or when local learning is perfect ($\Lambda \to \infty$, so $u \to \infty$), and approaches -1 in the common-drift limit ($\rho \to 1$).

(b) Let $(b_{\mathcal{W},\mu}, b_{\mathcal{W},Z})$ denote the population cross-sectional regression slopes of $\mathcal{W}_{n,t}$ on $(\mu^{\mathcal{H}}_{n,t}, Z_{n,t})$. Conditional on $(\log v_{m,t}, Z_{m,t})_{m=1}^N$ and holding $(\bar{\mu}^{\mathcal{H}}_t, \bar{Z}_t)$ fixed, the expected marginal effects of $\log v_{n,t}$ and $Z_{n,t}$ on firm n's expected excess return are

$$\begin{aligned} b_v &= b_{\mathcal{W},\mu}\mathbb{E}\left[\Gamma_{n,t}\left(\frac{\partial \log v_{n,t}}{\partial \mu^{\mathcal{H}}_{n,t}}\right)^{-1}\right], \\ b_Z &= -b_{\mathcal{W},\mu}\mathbb{E}\left[\Gamma_{n,t}\frac{\partial v_{n,t}/\partial Z_{n,t}}{\partial v_{n,t}/\partial \mu^{\mathcal{H}}_{n,t}}\right] + b_{\mathcal{W},Z}\mathbb{E}[\Gamma_{n,t}]. \end{aligned} \tag{15}$$

If $u \geq \kappa_Z$, then $b_{\mathcal{W},\mu} \leq b_{\mathcal{W}} < 0$ and $b_{\mathcal{W},Z} \geq 0$, so $b_v < 0$ and $b_Z > 0$. When $\log v_{n,t}$ is affine in $(Z_{n,t}, \mu^{\mathcal{H}}_{n,t})$, $\Gamma_{n,t}$ is constant and (b_v, b_Z) equal the population cross-sectional regression slopes of excess returns on $(\log v_{n,t}, Z_{n,t})$.

(c) The slopes $b_{\mathcal{W}}$, b_μ, b_v, and b_Z are independent of N. As $N \to \infty$, the variance of

the wedge converges to

$$\lim_{N\to\infty} \mathrm{Var}(\mathcal{W}_{n,t}) = \frac{u-\tilde{u}}{\Lambda+\sigma_Z^{-2}} > 0, \qquad \tilde{u} \equiv \sqrt{\kappa_\mu^2 + (\Lambda+\sigma_Z^{-2})\sigma_\mu^2(1-\rho)} < u. \quad (16)$$

Parts (a) and (b) of Proposition 4 summarize cross-sectional predictability in beliefs and characteristics. Firms with higher local beliefs earn lower returns on average ($b_\mu < 0$). When local beliefs weight recent news at least as heavily as accounting profitability ($u \geq \kappa_Z$), conditioning jointly on log market-to-book and profitability (15) gives a negative valuation effect ($b_v < 0$) and a positive profitability effect ($b_Z > 0$): higher profitability at a given valuation implies a lower local belief $\mu_{n,t}^{\mathcal{H}}$, and $b_{\mathcal{W},Z} \geq 0$ reinforces this prediction. Every effect in (13) and (15) is driven by the wedge slopes $b_{\mathcal{W}}$ and $(b_{\mathcal{W},\mu}, b_{\mathcal{W},Z})$, so all predictability vanishes when $\rho = 0$ or under complete aggregation. Thus, return predictors that look like separate anomalies are projections of the same information aggregation wedge.

Proposition 4 also establishes two robustness results. First, predictability does not rely on parameter heterogeneity: firms are ex ante identical, yet comparing their local beliefs $\mu_{n,t}^{\mathcal{H}}$ predicts returns in the cross-section. Second, part (c) shows that predictability survives as $N \to \infty$. Pooling infinitely many peers reveals the common drift component, whereas local specialists cannot separate common from idiosyncratic drift shocks. The wedge therefore retains positive variance, as shown in (16).

The homogeneous case of Proposition 4 serves only to illustrate the mechanism. In the general economy of Corollary 3.1, the slopes are more involved (see Appendix A.4.1), but the conclusions are unchanged.

2.4 Equilibrium Information Aggregation

Why does this predictability survive arbitrage? To answer this question, I add a continuum of risk-neutral arbitrageurs. At date 0, a mass $M \in [0, \infty)$ of arbitrageurs enters the market. Each pays a fixed cost $c > 0$ to observe the pooled filtration $\mathcal{G}_t$ and trade across all N firms. Over dt, an arbitrageur who trades a fraction $\theta_{n,t}dt$ of firm n pays $V_{n,t}\theta_{n,t}dt = K_{n,t}v_{n,t}\theta_{n,t}dt$ at the market price and incurs a quadratic execution cost $\frac{\gamma_n}{2}K_{n,t}\theta_{n,t}^2 dt$, where $\gamma_n > 0$. Firm n's specialists absorb the aggregate order flow from arbitrageurs and independent noise

trading $W^Y_{n,t}$:

$$dY_{n,t} = M\theta_{n,t}dt + \sigma_{Y,n}dW^Y_{n,t}, \tag{17}$$

where $\sigma_{Y,n} > 0$. Observing order flow expands firm n's local filtration to $\mathcal{F}_{n,t} \equiv \mathcal{H}_{n,t} \vee \sigma\{Y_{n,u} : u \leq t\}$. The updated local belief is $\mu^{\mathcal{F}}_{n,t} \equiv \mathbb{E}[\mu_{n,t} \mid \mathcal{F}_{n,t}]$, and the residual wedge is $\mathcal{W}^{\mathcal{F}}_{n,t} \equiv \mu^{\mathcal{G}}_{n,t} - \mu^{\mathcal{F}}_{n,t}$. Firm n's manager also chooses investment adapted to $\mathcal{F}_{n,t}$, learning from the market as in Chen et al. (2007) and Bond, Edmans, and Goldstein (2012). Given this investment policy, let $U_{n,t}$ denote the arbitrageur's valuation of firm n per unit of capital given $\mathcal{G}_t \vee \mathcal{F}_{n,t}$. The local market-to-book ratio satisfies $v_{n,t} = \mathbb{E}[U_{n,t} \mid \mathcal{F}_{n,t}]$.

Proposition 5 (Equilibrium Information Aggregation)**.** *Under the order-flow process* (17) *and standard regularity conditions:*

(a) An entering arbitrageur's optimal trading rate in firm n is $\theta^*_{n,t} = (U_{n,t} - v_{n,t})/\gamma_n$, which yields expected discounted gross trading profits

$$\Pi(M) = \mathbb{E}\left[\sum_{n=1}^{N} \frac{1}{2\gamma_n} \int_0^\infty e^{-rt} K_{n,t}(U_{n,t} - v_{n,t})^2 dt\right]. \tag{18}$$

(b) As $M \to \infty$, $\mathbb{E}[(U_{n,t} - v_{n,t})^2] = O(M^{-1})$ for every firm n and $t > 0$, and gross trading profits per arbitrageur vanish:

$$\lim_{M\to\infty} \Pi(M) = 0. \tag{19}$$

(c) For any information cost $c > 0$, free entry determines a finite equilibrium mass of arbitrageurs $M^* < \infty$, satisfying $\Pi(M^*) = c$ if $0 < c < \Pi(0)$ and $M^* = 0$ if $c \geq \Pi(0)$, in which the residual wedge strictly survives: $\mathrm{Var}(\mathcal{W}^{\mathcal{F}}_{n,t}) > 0$ whenever $\mathrm{Var}(\mathcal{W}_{n,t}) > 0$.

Proposition 5 is a cross-sectional counterpart to Grossman and Stiglitz (1980). In (18), arbitrageurs profit from the gap $U_{n,t} - v_{n,t}$ between their valuation and the local price. As more arbitrageurs enter ($M \to \infty$), their trading makes order flow more informative. The gap narrows, and profits per arbitrageur fall to zero. Complete cross-sectional efficiency is therefore self-defeating: if trading eliminated the wedge, arbitrageurs could not recover the cost c.

Nor can arbitrageurs close the wedge for free by conditioning on public characteristics

alone. Public characteristics mix local beliefs $\mu^{\mathcal{H}}_{n,t}$ with current profitability $Z_{n,t}$. Even observing every firm's local belief at date t does not recover the pooled belief $\mu^{\mathcal{G}}_{n,t}$, which requires filtering full signal histories across all N firms (Timmermann, 2006).[4] Traders who condition only on public characteristics, without paying c, therefore have coarser information than $\mathcal{G}_t$ and cannot eliminate the wedge.

Because $M^* < \infty$, order flow is an additional signal of finite precision, and a residual wedge survives under the augmented filtration $\mathcal{F}_{n,t}$. In the remainder of the paper, I let local precision Λ_n absorb this signal and retain the baseline notation $\mathcal{H}_{n,t}$ and u_n. Within this reduced-form specification, the earlier predictability results apply.

3 Implications

This section develops the model's implications. First, I use the homogeneous-firms benchmark of Proposition 4 to contrast time-series and cross-sectional predictability. Next, in the general economy, I study which stocks earn higher returns. I then show how earnings surprises and belief revisions generate those returns and why stocks migrate across value and growth portfolios. Finally, I let signal precision vary over time to examine returns on scheduled announcement days.

Table 1 summarizes excess-return predictability from the local belief $\mu^{\mathcal{H}}_{n,t}$ in Panel A, observable characteristics $(\log v_{n,t}, Z_{n,t})$ in Panel B (expected marginal effects), and the unobservable wedge $\mathcal{W}_{n,t}$ in Panel C. Each specification can be evaluated in the time series (one firm n over time) or the cross-section (the N firms at date t). Columns (1) and (2) show two cases with no return predictability. When drifts are uncorrelated across firms ($\rho = 0$, Column 1), peer histories contain no information about firm n. When information is fully aggregated ($\mathcal{F}_{n,t} = \mathcal{G}_t$, Column 2), local beliefs already reflect peer histories. In either case, the wedge vanishes ($\mathcal{W}_{n,t} = 0$). All time-series and cross-sectional entries are zero in Panels A and B, and are undefined in Panel C because the wedge does not vary.

Column (3) considers correlated drifts and incomplete aggregation ($\rho > 0$ and $M^* < \infty$).

[4] In the homogeneous-firms benchmark, Appendix A shows that $\mu^{\mathcal{H}}_{n,t}$ weights past signals at a single exponential decay rate u, whereas $\mu^{\mathcal{G}}_{n,t}$ weights cross-sectional averages and firm-specific deviations at two distinct rates (faster than u on the average and $\tilde{u} < u$ on deviations), so the date-t vector $(\mu^{\mathcal{H}}_{1,t}, \ldots, \mu^{\mathcal{H}}_{N,t})$ recovers only a linear projection of $\mu^{\mathcal{G}}_{n,t}$ rather than $\mu^{\mathcal{G}}_{n,t}$ itself.

Table 1: Return Predictability Across Information Structures. Population regression slopes of excess returns $dR^E_{n,t}/dt - r$ on the local belief $\mu^{\mathcal{H}}_{n,t}$ (Panel A), log market-to-book and profitability (Panel B, expected marginal effects, $u \geq \kappa_Z$), and the information aggregation wedge $\mathcal{W}_{n,t}$ (Panel C) in the homogeneous-firms benchmark of Proposition 4, with $b_{\mathcal{W}} = -\rho\kappa_\mu/[(1-\rho)u + \rho\kappa_\mu] \in (-1, 0)$ in Column (3b).

	(1)	(2)	(3) Incomplete Aggregation and $\rho > 0$	
			(3a)	(3b)
Independent Variable	$\rho = 0$	$\mathcal{F}_{n,t} = \mathcal{G}_t$	Time Series	Cross-Section
Panel A				
Local Belief ($\mu^{\mathcal{H}}_{n,t}$)	0	0	0	$b_\mu = b_{\mathcal{W}}\mathbb{E}[\Gamma_{n,t}] < 0$
Panel B				
Log Market-to-Book ($\log v_{n,t}$)	0	0	0	$b_v = b_{\mathcal{W},\mu}\mathbb{E}\left[\Gamma_{n,t}\left(\frac{\partial \log v_{n,t}}{\partial \mu^{\mathcal{H}}_{n,t}}\right)^{-1}\right] < 0$
Profitability ($Z_{n,t}$)	0	0	0	$b_Z = \mathbb{E}\left[\Gamma_{n,t}\left(b_{\mathcal{W},Z} - b_{\mathcal{W},\mu}\frac{\partial v_{n,t}/\partial Z_{n,t}}{\partial v_{n,t}/\partial \mu^{\mathcal{H}}_{n,t}}\right)\right] > 0$
Panel C				
Wedge ($\mathcal{W}_{n,t}$)	—	—	$\mathbb{E}[\Gamma_{n,t}] > 0$	$\mathbb{E}[\Gamma_{n,t}] > 0$

Here, the time series and the cross-section diverge. In a firm's own time series (Column 3a), local efficiency (8) implies that neither $\mu^{\mathcal{H}}_{n,t}$ nor $(\log v_{n,t}, Z_{n,t})$ predicts returns. Only the unobservable wedge $\mathcal{W}_{n,t}$ does, with slope $\mathbb{E}[\Gamma_{n,t}] > 0$. In the cross-section (Column 3b), comparing firms gives $b_{\mathcal{W}} < 0$ by Proposition 4, so higher beliefs predict lower returns ($b_\mu < 0$). Conditioning jointly on $(\log v_{n,t}, Z_{n,t})$ when $u \geq \kappa_Z$, log market-to-book has a negative marginal effect ($b_v < 0$) and profitability a positive marginal effect ($b_Z > 0$). Although Table 1 uses the homogeneous-firms benchmark to simplify the formulas, the same zeros and signs hold under moderate firm heterogeneity (see Appendix A.4.1).

Table 1 thus captures the paper's central mechanism. In Panel C, the wedge $\mathcal{W}_{n,t}$ predicts returns with the same slope $\mathbb{E}[\Gamma_{n,t}] > 0$ in Columns (3a) and (3b). Subsequent learning corrects forecast errors in both the time series and the cross-section. For a single firm, however, local beliefs and characteristics are orthogonal to the wedge, so they cannot predict this error correction. Across correlated firms, differences in beliefs and characteristics predict the wedge. Cross-sectional efficiency therefore requires more than stock-by-stock efficiency: the market must aggregate dispersed local information.

3.1 Cross-Sectional Predictability

Corollary 3.1 links cross-sectional predictability to two channels: the information parameters (ρ_{nm}, u_n) and the characteristic sensitivities $\partial x_{n,t}/\partial \mu^{\mathcal{H}}_{n,t}$ and $\partial x_{n,t}/\partial Z_{n,t}$. I examine each in turn.

3.1.1 Information Parameters

Equation (12) depends on two parameters that govern learning across and within firms: the drift correlations ρ_{nm} and the local filtering speeds u_n. Filtering speeds rise with signal precision Λ_n in (1). The correlation ρ_{nm} governs how much peer m's history reveals about firm n's drift, while u_n sets how fast firm n's specialists learn on their own. To isolate their effects, the next proposition holds the parameters $(\kappa_Z, \sigma_Z, \kappa_\mu, \bar{\mu}, \sigma_\mu, \delta)$ and cost function Ψ fixed across firms and varies only (ρ_{nm}, Λ_n). The same conclusions hold under moderate parameter heterogeneity.

Proposition 6 (Information Parameters)**.** *Suppose all firms share identical parameters* $(\kappa_Z, \sigma_Z, \kappa_\mu, \bar{\mu}, \sigma_\mu, \delta)$ *and cost function* Ψ*. Then:*

(a) With common precision $(\Lambda_n = \Lambda)$ *and* $\mathbb{E}[\Gamma^2_{n,t}] < \infty$*, partition firms into* $G \geq 2$ *equal-sized industries of* $N_G = N/G \geq 2$ *firms each, with drift correlation* $\rho_w \in (0,1)$ *within industries and* $\rho_a \in [0, \rho_w)$ *across industries. Let* $\bar{\mu}^{\mathcal{H}}_{g,t}$ *and* $d\bar{R}^E_{g,t}$ *denote the industry-g average belief and return, and define* $\psi \equiv \frac{\sigma^2_\mu(u-\kappa_\mu)}{2u(u+\kappa_\mu)} > 0$*. The expected within-industry and across-industry covariances are*

$$\mathbb{E}\left[\mathrm{Cov}^{(g)}_{CS}\left(\mu^{\mathcal{H}}_{n,t}, \frac{dR^E_{n,t}}{dt} - r\right)\right] = -\left(1 - \frac{1}{N_G}\right)\rho_w \psi\, \mathbb{E}[\Gamma_{n,t}] < 0, \tag{20}$$

$$\mathbb{E}\left[\mathrm{Cov}^{Ind}_{CS}\left(\bar{\mu}^{\mathcal{H}}_{g,t}, \frac{d\bar{R}^E_{g,t}}{dt} - r\right)\right] = \left(1 - \frac{1}{G}\right)\left[\left(1 - \frac{1}{N_G}\right)\rho_w - \rho_a\right]\psi\, \mathbb{E}[\Gamma_{n,t}]. \tag{21}$$

The across-industry covariance (21) is strictly positive whenever $(1 - 1/N_G)\rho_w > \rho_a$*.*

(b) Consider any pair of correlated firms $(\rho_{nm} > 0)$*. If firm n has higher signal precision than peer m* $(\Lambda_n > \Lambda_m$*, so* $u_n > u_m)$*, firm n's belief covaries more strongly with peer m's wedge than the reverse:*

$$\mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, \mathcal{W}_{m,t}) > \mathrm{Cov}(\mu^{\mathcal{H}}_{m,t}, \mathcal{W}_{n,t}) > 0. \tag{22}$$

More generally, $\mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, \mathcal{W}_{m,t})$ *strictly increases in* Λ_n *and strictly decreases in* Λ_m.

Part (a) of Proposition 6 shows how forming industry portfolios changes return predictability. Within an industry, (20) recovers the negative covariance between beliefs and returns from Proposition 3, scaled by ρ_w. Across industries, (21) can turn positive because drift correlation is stronger within than across them ($\rho_w > \rho_a$). Two empirical implications follow. First, value, profitability, and short-term reversal predict returns mainly within industries (Cohen and Polk, 1996; Asness, Porter, and Stevens, 2000; Novy-Marx, 2013; Kim et al., 2025). Second, when $(1-1/N_G)\rho_w > \rho_a$, the sign flip in (21) generates industry momentum (buying winning industries and selling losing ones, Moskowitz and Grinblatt 1999). Each firm's specialists price local news immediately, so no stock shows momentum on its own. Instead, because firm-specific drifts are correlated ($\rho_w > 0$), good news at one firm predicts higher future returns for its $N_G - 1$ industry peers. An industry portfolio bundles these peers: one stock's gain today is followed by peer gains tomorrow.

Part (b) of Proposition 6 considers two correlated firms ($\rho_{nm} > 0$). The second inequality in (22) shows that each firm's belief covaries positively with the other's wedge. Good news at firm n raises its belief and stock price today. Because their drifts are correlated, that news predicts positive future surprises at peer m. Thus, for economically linked firms (e.g., customers and suppliers), higher past returns at either firm predict higher future returns at the other (Cohen and Frazzini, 2008). The first inequality in (22) shows that unequal signal precision ($\Lambda_n > \Lambda_m$, so $u_n > u_m$) makes this predictability asymmetric. Because firm n's signal is more precise, its belief and stock price track its drift more closely, while peer m's forecast error decays more slowly. Returns of firms with more precise local information therefore lead those of less-informed peers (Hou, 2007).[5]

Local precision Λ_n also enters firm n's own return sensitivity $\Gamma_{n,t}$ in (10) through the error-correction rate $u_n - \kappa_{\mu,n}$. During scheduled disclosures such as earnings announcements, precision rises temporarily and specialists put more weight on the incoming signal.

[5]Equation (22) is stated for date-t beliefs, which map into date-t characteristics via Corollary 3.1. For cumulative excess returns $R^E_{n,t-h\to t} \equiv \int_{t-h}^{t}(dR^E_{n,s} - r\,ds)$ over $h > 0$, firm n's local shocks (dW^Z_n, dW^O_n) are independent of peer m, so a first-order approximation around steady state ($\Gamma_{k,t} \approx \bar{\Gamma}_k \equiv \mathbb{E}[\Gamma_{k,t}]$) and (3) yield $\mathrm{Cov}(R^E_{n,t-h\to t}, dR^E_{m,t}/dt - r) \approx \bar{\Gamma}_n\bar{\Gamma}_m \frac{\rho_{nm}\sigma^2_\mu}{u_n+u_m}\frac{1-e^{-u_m h}}{u_m} > 0$ $(n \neq m)$. Because $\bar{\Gamma}_n\bar{\Gamma}_m \frac{\rho_{nm}\sigma^2_\mu}{u_n+u_m}$ is symmetric in n and m while $(1-e^{-uh})/u$ strictly decreases in u, this covariance is larger from n to m than from m to n whenever $u_n > u_m$.

This higher weight raises $\Gamma_{n,t}$ and resolves more of the wedge. Anomaly returns therefore concentrate on announcement days (Engelberg et al., 2018). Because Λ_n is constant in the baseline model, Section 3.3 lets precision vary over time to formalize this effect.

3.1.2 Characteristic Sensitivities

The terms in curly braces in (12) show that a characteristic $x_{n,t}$ predicts returns through its sensitivities to the local belief $\mu^{\mathcal{H}}_{n,t}$ and current profitability $Z_{n,t}$. Market-to-book $v_{n,t}$ rises with both. A low valuation can thus reflect weak current profitability or a low local belief about the firm's drift. This explains the opposite signs in Panel B of Table 1. At a given $Z_{n,t}$, a lower valuation $v_{n,t}$ implies a lower belief $\mu^{\mathcal{H}}_{n,t}$ and predicts higher returns ($b_v < 0$). At a given $v_{n,t}$, higher profitability $Z_{n,t}$ also implies a lower belief $\mu^{\mathcal{H}}_{n,t}$ (otherwise valuation would be higher) and predicts higher returns ($b_Z > 0$). Value and profitability therefore predict returns better together than either alone (Novy-Marx, 2013).

This mechanism differs from the discount-rate interpretation in Novy-Marx (2013). There, a present-value identity links a firm's valuation and expected cash flows to its discount rate. Holding profitability fixed, a lower valuation implies a higher firm-level discount rate, as does higher profitability at a given valuation. Because that identity holds firm by firm, differences in discount rates would predict returns even when fundamentals are uncorrelated across firms ($\rho = 0$, Column 1 of Table 1). Time-varying discount rates would also predict returns in a firm's own time series (unlike Column 3a). In this learning model, the discount rate is constant (r), and predictability comes from comparing correlated firms. Cheap, profitable stocks earn higher returns when their local beliefs lie below those of correlated peers. Predictability is therefore purely cross-sectional and concentrates within industries, rather than in industry averages or a firm's own history. I test these predictions in Section 4.

Equation (12) also applies to other characteristics that increase in $\mu^{\mathcal{H}}_{n,t}$ and $Z_{n,t}$. If corporate investment $i_{n,t}$ rises with market-to-book $v_{n,t}$, as implied by the q-theory of investment (Tobin, 1969; Hayashi, 1982), higher investment predicts lower returns (Titman, Wei, and Xie, 2004; Cooper, Gulen, and Schill, 2008). Low valuation and low investment both identify firms whose local beliefs $\mu^{\mathcal{H}}_{n,t}$ lie below those of their peers. More generally, distinct anomalies in the "factor zoo" may proxy for the same information aggregation

wedge.

A characteristic can also proxy directly for the local belief $\mu^{\mathcal{H}}_{n,t}$, as do analysts' long-term growth forecasts. When $\partial x_{n,t}/\partial \mu^{\mathcal{H}}_{n,t} = 1$ and $\partial x_{n,t}/\partial Z_{n,t} = 0$, only the belief term in (12) remains. Stocks with high growth forecasts therefore earn lower returns (La Porta, 1996). Moreover, in the homogeneous benchmark (13)–(14), the expected cross-sectional covariance between beliefs and returns is $b_\mu \mathbb{E}[\mathrm{Var}_{CS}(\mu^{\mathcal{H}}_{n,t})]$. This suggests that for a given slope $b_\mu < 0$, greater dispersion in beliefs (a wider value spread) implies a larger value premium (Cohen, Polk, and Vuolteenaho, 2003).

3.2 Error Correction, Belief Revisions, and Migration

Section 3.1 shows which stocks earn higher expected returns in the cross-section. I now ask what drives those returns after portfolio formation and why stocks migrate across portfolios. Over dt, local efficiency (8) implies $\mathbb{E}[dR^E_{n,t} \mid \mathcal{H}_{n,t}] = r\,dt$. By (9), firm n's expected excess return under $\mathcal{G}_t$ therefore equals the gap between the pooled and local expectations, $\mathbb{E}[dR^E_{n,t} \mid \mathcal{G}_t] - \mathbb{E}[dR^E_{n,t} \mid \mathcal{H}_{n,t}] = \Gamma_{n,t}\mathcal{W}_{n,t}dt$. Substituting $\Gamma_{n,t}$ from (10) splits this gap into two terms:

$$\Gamma_{n,t}\mathcal{W}_{n,t}dt = \frac{1}{v_{n,t}}\left[\frac{\partial v_{n,t}}{\partial Z_{n,t}}\,\mathcal{W}_{n,t}dt + \frac{\partial v_{n,t}}{\partial \mu^{\mathcal{H}}_{n,t}}\,(u_n - \kappa_{\mu,n})\mathcal{W}_{n,t}dt\right]. \tag{23}$$

The first term of (23) affects valuation through $\partial v_{n,t}/\partial Z_{n,t}$, where $\mathcal{W}_{n,t}dt = \mathbb{E}[dZ_{n,t} \mid \mathcal{G}_t] - \mathbb{E}[dZ_{n,t} \mid \mathcal{H}_{n,t}]$. Consider a value stock whose local belief $\mu^{\mathcal{H}}_{n,t}$ is low relative to its peers' beliefs. On average, its wedge $\mathcal{W}_{n,t}$ is positive, so local expectations for the change in profitability lie below the pooled expectation. As realized profitability tends to beat those expectations, value stocks experience positive subsequent earnings surprises (La Porta et al., 1997).

The second term of (23) affects valuation through $\partial v_{n,t}/\partial \mu^{\mathcal{H}}_{n,t}$, where $(u_n - \kappa_{\mu,n})\mathcal{W}_{n,t}dt = \mathbb{E}[d\mu^{\mathcal{H}}_{n,t} \mid \mathcal{G}_t] - \mathbb{E}[d\mu^{\mathcal{H}}_{n,t} \mid \mathcal{H}_{n,t}]$. Under $\mathcal{H}_{n,t}$, specialists expect only mean reversion in the filter (2), which is already priced into $v_{n,t}$. On average, however, a value stock receives positive cash-flow and signal surprises, so $(u_n - \kappa_{\mu,n})\mathcal{W}_{n,t}dt$ is positive. The reverse holds for a growth stock. Specialists therefore tend to revise their beliefs upward for value stocks relative to growth stocks, a prediction I test with analyst forecast revisions in Section 4.

Together, these surprises and belief revisions raise $v_{n,t}$ for value stocks and lower it for growth stocks, so stocks migrate across portfolios. In fact, local efficiency (8) and the law of iterated expectations imply $\mathbb{E}[dR^E_{n,t} - r\,dt] = 0$ for every firm n. No firm is a permanent high-return type: buying and holding any individual stock earns zero unconditional expected excess return. A characteristic strategy can sustain a premium only through rebalancing: buying stocks when low local beliefs push $v_{n,t}$ down and selling them once $v_{n,t}$ recovers. The value premium thus *requires* portfolio migration. Consistent with this mechanism, Fama and French (2007) and Broussard, Mikkonen, and Puttonen (2016) find that the value premium comes mainly from migrating stocks rather than from stocks that remain in the same portfolio.

3.3 Time-Varying Precision and Announcement Returns

The model so far assumes constant signal precision Λ_n. I now let firm n's precision follow a deterministic schedule $\Lambda_{n,t} > 0$ that is bounded and piecewise continuous. At $t = 0$, the economy starts from the steady state of Section 2, with initial precision $\Lambda_{n,0}$. A natural example is a schedule that stays at a baseline $\underline{\Lambda}_n$ during routine trading and rises to $\overline{\Lambda}_n > \underline{\Lambda}_n$ during scheduled disclosures (e.g., earnings announcements).[6] The question I address here is: How does time-varying precision affect cross-sectional return predictability?

Conditional on $\mathcal{H}_{n,t}$, $\mu_{n,t}$ remains Gaussian with posterior mean $\mu^{\mathcal{H}}_{n,t}$ and deterministic posterior variance $\nu_{n,t} \equiv \mathrm{Var}(\mu_{n,t} \mid \mathcal{H}_{n,t})$, which evolves as

$$\frac{d\nu_{n,t}}{dt} = \sigma^2_{\mu,n} - 2\kappa_{\mu,n}\nu_{n,t} - (\Lambda_{n,t} + \sigma^{-2}_{Z,n})\nu^2_{n,t}. \tag{24}$$

Specialists know in advance when precision will rise, but they cannot predict the direction of future surprises. Local efficiency (8) therefore holds at every date t. Proposition 2 carries over with two changes: the market-to-book ratio $v_{n,t}$ now depends on calendar time, and the error-correction rate $u_n - \kappa_{\mu,n}$ in (10) is replaced by $(\Lambda_{n,t} + \sigma^{-2}_{Z,n})\nu_{n,t}$.

[6] Technically, $\Lambda_{n,t}$ is assumed to be right-continuous with left limits. One could also model a scheduled disclosure as a discrete signal $A_{n,T} = \mu_{n,T} + \varepsilon_{n,T}$ with independent Gaussian noise $\varepsilon_{n,T} \sim \mathcal{N}(0, J_n^{-1})$, which yields the same mechanism: the expected belief jump under $\mathcal{G}_{T^-}$ is $\mathbb{E}[\mu^{\mathcal{H}}_{n,T} - \mu^{\mathcal{H}}_{n,T^-} \mid \mathcal{G}_{T^-}] = g_{n,T}\mathcal{W}_{n,T^-}$, where $g_{n,T} \equiv J_n\nu_{n,T^-}/(1 + J_n\nu_{n,T^-}) \in (0,1)$ increases in both disclosure precision J_n and pre-announcement uncertainty ν_{n,T^-}.

Proposition 7 (Announcement Returns)**.** *At any date $t > 0$ where firm n's precision rises from $\Lambda_{n,t^-} = \underline{\Lambda}_n$ to $\Lambda_{n,t^+} = \overline{\Lambda}_n > \underline{\Lambda}_n$, the expected excess-return rate jumps by*

$$\left.\frac{\mathbb{E}[dR^E_{n,s} \mid \mathcal{G}_s]}{ds}\right|_{s=t^+} - \left.\frac{\mathbb{E}[dR^E_{n,s} \mid \mathcal{G}_s]}{ds}\right|_{s=t^-} = (\overline{\Lambda}_n - \underline{\Lambda}_n)\,\nu_{n,t}\,\frac{1}{v_{n,t}}\frac{\partial v_{n,t}}{\partial \mu^{\mathcal{H}}_{n,t}}\,\mathcal{W}_{n,t}. \tag{25}$$

If $\rho_{nm} \geq 0$ for all $m \neq n$, with $\rho_{nm} > 0$ for some $m \neq n$, and $\mathbb{E}[\Gamma^2_{n,t}] < \infty$, firm n's belief–return covariance $\mathrm{Cov}(\mu^{\mathcal{H}}_{n,s} - \bar{\mu}^{\mathcal{H}}_s, dR^E_{n,s}/ds - r) < 0$ becomes strictly more negative from $s = t^-$ to $s = t^+$.

Proposition 7 explains why anomaly returns concentrate around earnings announcements and other scheduled disclosures (Engelberg et al., 2018). When precision rises from $\underline{\Lambda}_n$ to $\overline{\Lambda}_n$ at date t, specialists put more weight on the incoming signal and resolve more of their forecast error. In (25), that extra weight raises the expected excess-return rate for positive-wedge firms and lowers it for negative-wedge firms, in proportion to pre-announcement uncertainty $\nu_{n,t}$. Although local prices already reflect the schedule, the return materializes only when the news arrives. Because peer-relative beliefs still predict the wedge, firm n's belief–return covariance becomes strictly more negative. Characteristic-sorted long-short returns are therefore larger during announcements than during routine trading.

During an announcement, high precision lowers uncertainty (24) and pulls the wedge toward zero in expectation. Between announcements, uncertainty builds back up. The sign predictions of Corollary 3.1 still hold under this cycle, and so do the arbitrage limits of Section 2.4: even when arbitrageurs trade ahead of a scheduled announcement, costly information, execution costs, and noisy order flow leave $\mathrm{Var}(\mathcal{W}^{\mathcal{F}}_{n,t}) > 0$ at date t. A natural step for future work is to replace this known schedule with stochastic precision, which would make uncertainty a state variable in valuation.

4 Empirical Illustration

This section tests the model's mechanism using market-to-book $v_{n,t}$ and profitability $Z_{n,t}$. In Panel B of Table 1, $\log v_{n,t}$ predicts returns negatively and $Z_{n,t}$ positively, even though each stock is locally efficient and discounted at the constant rate r. If incomplete

information aggregation drives this pattern, three predictions follow. First, predictability concentrates among correlated peers rather than across weakly correlated industries or in a firm's own history (Proposition 6). Second, because $v_{n,t}$ rises with both $Z_{n,t}$ and $\mu^{\mathcal{H}}_{n,t}$, a firm with high $Z_{n,t}$ and low $v_{n,t}$ must have a low local belief $\mu^{\mathcal{H}}_{n,t}$. When that belief lies below peers' beliefs, the wedge is positive on average ($\mathcal{W}_{n,t} > 0$), so local beliefs are later revised upward relative to peers (Proposition 3). Third, return predictability strengthens in earnings announcement months, as higher signal precision leads investors to put more weight on the incoming disclosure and resolves more of the wedge (Proposition 7).

I combine monthly stock returns from CRSP, annual accounting data from Compustat, and analyst forecasts from I/B/E/S. Market-to-book $v_{n,t}$ is market equity over book equity, and profitability $Z_{n,t}$ is operating income before depreciation over book equity. Consensus forecasts of long-term earnings growth ($\text{LTG}_{n,t}$) proxy for local beliefs $\mu^{\mathcal{H}}_{n,t}$. Following Fama and French (1992) and Novy-Marx (2013), each June t from 1963 to 2023, I match year $t-1$ characteristics to returns from July t through June $t+1$. Standard filters (Appendix B) yield 139,118 firm-years of U.S. non-financial common stocks (1,627,556 firm-months), 67,035 of which have initial I/B/E/S forecasts (1982–2023).

I start with returns. Each month τ from July t through June $t+1$, I estimate the cross-sectional regression

$$R_{n,\tau} = a_\tau + b_{v,\tau} \log v_{n,t} + b_{Z,\tau} Z_{n,t} + \eta_{n,\tau}, \tag{26}$$

where $R_{n,\tau}$ is firm n's return in percent per month. In (15), the model implies slopes $b_v < 0$ and $b_Z > 0$ (for $u \geq \kappa_Z$). As in Novy-Marx (2013), Column (1) of Table 2 confirms both signs, with estimates of -0.336 ($t = -4.17$) on $\log v_{n,t}$ and 0.694 ($t = 4.03$) on $Z_{n,t}$. A one-standard-deviation change in either characteristic (0.87 for $\log v_{n,t}$ and 0.35 for $Z_{n,t}$) moves expected returns by about 25 to 30 basis points per month.

Column (2) splits each characteristic into its two-digit SIC industry average and the firm's deviation from it. Predictability strengthens within industries (-0.369 and 0.767, both $|t| > 5.4$), while industry averages have no predictive power ($t = -1.25$ and $t = 0.18$). A firm earns higher returns only when it is cheap and profitable *relative to peers* in its industry. Consistent with part (a) of Proposition 6, because drifts covary more strongly within than across industries ($\rho_w > \rho_a$), peers' histories reveal when a firm's local belief

Table 2: Cross-Sectional Return Regressions. Monthly Fama–MacBeth regressions of stock returns (in percent per month, July t through June $t+1$) on log market-to-book ($\log v_{n,t}$) and operating profitability ($Z_{n,t}$). Column (2) decomposes each characteristic into its two-digit SIC deviation (Within-Industry) and industry average (Industry Mean). Column (3) adds each firm's prior five-year average (Past 5-Year) and leave-one-out two-digit SIC average (Industry-Peer). Twelve-lag Newey–West t-statistics are in parentheses; *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

Independent Variable	(1) Baseline	(2) Industry Decomp.	(3) History vs. Peers
Log Market-to-Book ($\log v_{n,t}$)	-0.336^{***}		-0.253^{***}
	(-4.17)		(-2.65)
Profitability ($Z_{n,t}$)	0.694^{***}		0.795^{***}
	(4.03)		(4.09)
Within-Industry $\log v_{n,t}$		-0.369^{***}	
		(-5.49)	
Within-Industry $Z_{n,t}$		0.767^{***}	
		(5.48)	
Industry Mean $\log v_{n,t}$		-0.198	
		(-1.25)	
Industry Mean $Z_{n,t}$		0.095	
		(0.18)	
Past 5-Year $\log v_{n,t}$			-0.015
			(-0.15)
Past 5-Year $Z_{n,t}$			-0.233
			(-1.02)
Industry-Peer $\log v_{n,t}$			0.235^{**}
			(2.17)
Industry-Peer $Z_{n,t}$			-1.367^{***}
			(-2.60)
Intercept	1.049^{***}	1.198^{***}	1.374^{***}
	(5.06)	(4.63)	(5.50)
Sample period	1963:07–2024:06	1963:07–2024:06	1966:07–2024:06
Months (T)	732	732	696
Avg. firms / month	2,223	2,223	1,664
Firm-months (N)	1,627,556	1,627,556	1,157,985
Avg. R^2 (%)	1.77	2.44	3.44

lies below the pooled belief ($\mathcal{W}_{n,t} > 0$).[7]

Column (3) adds two benchmarks alongside firm n's current characteristics: its prior

[7]Defining peers more finely strengthens the within-industry slopes: -0.414 ($t = -6.45$) on $\log v_{n,t}$ and 0.824 ($t = 6.25$) on $Z_{n,t}$ in three-digit SIC industries, and -0.406 ($t = -6.57$) and 0.857 ($t = 5.79$) in four-digit SIC industries, while industry averages remain insignificant ($|t| \leq 1.61$).

five-year average (Past 5-Year, inside $\mathcal{H}_{n,t}$) and the leave-one-out industry average (Industry-Peer, outside $\mathcal{H}_{n,t}$). The firm's five-year history has no predictive power ($t = -0.15$ and $t = -1.02$). This is consistent with local efficiency (8): specialists already price that history, so comparing a firm with its past reveals nothing about its wedge.

By contrast, Industry-Peer $\log v_{n,t}$ enters positively (0.235, $t = 2.17$) and Industry-Peer $Z_{n,t}$ negatively (-1.367, $t = -2.60$), opposite in sign to firm n's own slopes (-0.253 and 0.795). These opposite signs follow from the wedge $\mathcal{W}_{n,t} = \mu^{\mathcal{G}}_{n,t} - \mu^{\mathcal{H}}_{n,t}$: firm n's belief enters negatively, whereas peers' beliefs raise the pooled belief $\mu^{\mathcal{G}}_{n,t}$. Holding firm n's characteristics fixed, higher peer valuation (given profitability) implies higher peer beliefs and raises $\mathcal{W}_{n,t}$. Conversely, higher peer profitability (given valuation) implies lower peer beliefs and lowers $\mathcal{W}_{n,t}$. Thus, a firm's standing relative to peers today, rather than its own past, identifies the wedge $\mathcal{W}_{n,t}$.

I now turn from returns to beliefs. Each year t, I estimate the cross-sectional regression

$$y_{n,t} = a_t + c_{v,t} \log v_{n,t} + c_{Z,t} Z_{n,t} + \varepsilon_{n,t}, \tag{27}$$

where $y_{n,t}$ is either the initial growth forecast $\log \mathrm{LTG}_{n,t}$ (Columns 1–2 of Table 3) or the subsequent one-year revision $\Delta \log \mathrm{LTG}_{n,t\to t+1}$ (Columns 3–4). Columns (2) and (4) replace $\log v_{n,t}$ and $Z_{n,t}$ with their within-industry deviations. For initial forecasts $y_{n,t} = \log \mathrm{LTG}_{n,t}$, inverting the valuation function (7) predicts $c_{v,t} > 0$ and $c_{Z,t} < 0$: at a given $\log v_{n,t}$, a more profitable firm must have a lower local belief. Columns (1) and (2) confirm both signs unconditionally (0.374 and -0.511, both $|t| > 13$) and within industries (0.284 and -0.356, both $|t| > 8$). Within an industry, low valuation and high profitability identify firms with lower growth forecasts than peers, and thus a positive average wedge.

Columns (3) and (4) track how analysts revise their forecasts from June t to June $t+1$, the same twelve-month window as the returns in Table 2. Because firms with lower initial beliefs carry a positive wedge ($\mathcal{W}_{n,t} > 0$), incoming news pulls their forecasts upward, as in (23). The signs in (27) should therefore flip to $c_{v,t} < 0$ and $c_{Z,t} > 0$. Both columns confirm this sign reversal unconditionally (-0.037, $t = -6.75$ on $\log v_{n,t}$, and 0.027, $t = 1.92$ on $Z_{n,t}$) and within industries (-0.034, $t = -6.14$, and 0.019, $t = 1.54$). Analysts start with lower growth forecasts for cheap, profitable firms than for industry peers, and revise them upward relative to peers over the following year.

Table 3: Analyst Growth Beliefs and Subsequent Revisions. Annual Fama–MacBeth regressions ($t = 1982, \ldots, 2023$) of December $t-1$ log consensus I/B/E/S long-term earnings growth forecasts ($\log \mathrm{LTG}_{n,t}$, Columns 1–2) and subsequent June t to June $t+1$ forecast revisions ($\Delta \log \mathrm{LTG}_{n,t\to t+1}$, Columns 3–4) on log market-to-book ($\log v_{n,t}$) and operating profitability ($Z_{n,t}$). Columns (2) and (4) use deviations from two-digit SIC industry averages (Within-Ind.). Three-lag Newey–West t-statistics are in parentheses; *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

	Initial Belief ($\log \mathrm{LTG}_{n,t}$)		1-Year Revision ($\Delta \log \mathrm{LTG}_{n,t\to t+1}$)	
	(1)	(2)	(3)	(4)
Independent Variable	Baseline	Within-Ind.	Baseline	Within-Ind.
Log Market-to-Book ($\log v_{n,t}$)	0.374^{***}		-0.037^{***}	
	(13.51)		(-6.75)	
Profitability ($Z_{n,t}$)	-0.511^{***}		0.027^{*}	
	(-14.56)		(1.92)	
Within-Industry $\log v_{n,t}$		0.284^{***}		-0.034^{***}
		(19.75)		(-6.14)
Within-Industry $Z_{n,t}$		-0.356^{***}		0.019
		(-8.51)		(1.54)
Intercept	2.518^{***}	2.649^{***}	-0.021^{*}	-0.043^{***}
	(57.02)	(69.79)	(-1.89)	(-5.19)
Sample period	1982–2023	1982–2023	1982–2023	1982–2023
Years (T)	42	42	42	42
Avg. firms / year	1,596	1,596	1,393	1,393
Firm-years (N)	67,035	67,035	58,522	58,522
Avg. R^2 (%)	22.72	10.28	1.14	0.79

Finally, I test Proposition 7 by comparing return predictability in months when a firm announces quarterly earnings ($\mathrm{Ann}_{n,\tau} = 1$, from Compustat `rdq` over 1972–2024) with months when it does not ($\mathrm{Ann}_{n,\tau} = 0$). Higher precision $\Lambda_{n,t}$ during announcements resolves a larger share of the wedge $\mathcal{W}_{n,t}$ (Engelberg et al., 2018). Columns (1) and (2) of Table 4 estimate (26) separately on the 1,561 non-announcing and 755 announcing firms each month, and Column (3) reports their difference from a pooled interaction regression. The slope on $\log v_{n,t}$ nearly doubles from -0.251 ($t = -2.84$) in non-announcement months to -0.493 ($t = -4.97$) in announcement months, a difference of -0.242 ($t = -4.46$). The slope on $Z_{n,t}$ rises from 0.420 ($t = 2.27$) to 0.529 ($t = 2.23$). Although announcement months represent a third of the sample (32.6% of firm-months), they generate nearly half of the annual valuation spread ($4 \times 0.493\% = 1.97\%$ out of $4 \times 0.493\% + 8 \times 0.251\% = 3.98\%$).

Columns (4) through (6) combine Propositions 6 and 7 by repeating this comparison

Table 4: Return Predictability in Earnings Announcement Months. Monthly Fama–MacBeth regressions (July 1972 through June 2024, $T = 624$) of stock returns (in percent per month) on $\log v_{n,t}$ and $Z_{n,t}$ (Columns 1–3) and on their within-industry and industry-mean components (Columns 4–6), estimated separately for firms without ($\text{Ann}_{n,\tau} = 0$, Columns 1 and 4) and with ($\text{Ann}_{n,\tau} = 1$, Columns 2 and 5) a Compustat quarterly earnings announcement (`rdq`) in month τ. Columns (3) and (6) report the announcement-month interaction slopes ($\times\text{Ann}_{n,\tau}$) from pooled monthly regressions. Twelve-lag Newey–West t-statistics are in parentheses; *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

	Baseline Characteristics			Industry Decomposition		
	(1)	(2)	(3)	(4)	(5)	(6)
Independent Variable	Non-Ann.	Ann. Month	Diff. ($\times$Ann)	Non-Ann.	Ann. Month	Diff. ($\times$Ann)
Log Market-to-Book ($\log v_{n,t}$)	-0.251^{***}	-0.493^{***}	-0.242^{***}			
	(-2.84)	(-4.97)	(-4.46)			
Profitability ($Z_{n,t}$)	0.420^{**}	0.529^{**}	0.109			
	(2.27)	(2.23)	(0.56)			
Within-Industry $\log v_{n,t}$				-0.302^{***}	-0.545^{***}	-0.243^{***}
				(-3.96)	(-5.82)	(-4.06)
Within-Industry $Z_{n,t}$				0.529^{***}	0.628^{***}	0.098
				(3.40)	(3.02)	(0.49)
Industry Mean $\log v_{n,t}$				-0.093	-0.233	-0.140
				(-0.55)	(-1.25)	(-1.16)
Industry Mean $Z_{n,t}$				-0.291	-0.035	0.257
				(-0.52)	(-0.04)	(0.51)
Intercept / $\text{Ann}_{n,\tau}$	0.883^{***}	1.475^{***}	0.592^{***}	1.064^{***}	1.526^{***}	0.463^{**}
	(3.98)	(5.97)	(7.23)	(4.06)	(4.38)	(2.43)
Sample period	1972–2024	1972–2024	1972–2024	1972–2024	1972–2024	1972–2024
Months (T)	624	624	624	624	624	624
Avg. firms / month	1,561	755	2,316	1,561	755	2,316
Firm-months (N)	973,966	471,155	1,445,121	973,966	471,155	1,445,121
Avg. R^2 (%)	1.99	2.23	2.19	2.72	3.16	2.94

under the industry decomposition. The announcement effect comes entirely from within industries. The slope on Within-Industry $\log v_{n,t}$ jumps from -0.302 ($t = -3.96$) in non-announcement months to -0.545 ($t = -5.82$) in announcement months, a difference of -0.243 ($t = -4.06$), and the slope on Within-Industry $Z_{n,t}$ rises from 0.529 ($t = 3.40$) to 0.628 ($t = 3.02$). By contrast, the slopes on Industry Mean $\log v_{n,t}$ and $Z_{n,t}$ are insignificant across all three columns ($|t| \leq 1.25$). Return predictability concentrates where peer comparisons identify the wedge (within industries) and when precise disclosures arrive to resolve it.

Together, Tables 2 through 4 support the model's predictions. Market-to-book and profitability predict returns relative to current industry peers, not through industry

averages or a firm's own history. The same characteristics identify firms whose initial growth forecasts lie below peers' and are later revised upward. Return predictability also nearly doubles in announcement months. This evidence does not rule out variation in discount rates, but a firm-level discount rate alone cannot explain why return predictability is peer-relative; why the same characteristics predict subsequent revisions in growth forecasts; and why predictability concentrates in earnings announcement months.

5 Conclusion

Hayek (1945, p. 520) observed that society's economic problem is how to use "knowledge not given to anyone in its totality." This paper applies Hayek's insight to the cross-section of returns. Each firm's risk-neutral specialists price its stock efficiently given their local information. That price reflects other firms' information only through arbitrageurs' trades, and only partially, because processing information across firms is costly. The resulting information aggregation wedge—the gap between the pooled and local beliefs about a firm's drift—drives cross-sectional predictability. The wedge itself is unobservable. Yet to detect this predictability, no one needs to know the true drifts, their correlations, or even the pooled beliefs. Comparing firms is enough.

Incomplete aggregation also helps explain why anomaly returns decline after publication (McLean and Pontiff, 2016; Marrow and Nagel, 2024). If these returns compensated for risk, publishing an academic paper would not erode them. If they reflected data mining alone, they would vanish out of sample even before publication. In my model, predictability is real before publication because aggregating information across firms is costly. Once research documents a predictor, arbitrageurs trade on it, local prices incorporate more peer information, and the predictor's expected return shrinks. Yet because public characteristics capture only a projection of the wedge, arbitraging published predictors does not eliminate cross-sectional predictability.

Several questions remain for future research. In this paper, signal precision follows a known schedule, so posterior uncertainty is deterministic. If precision instead varied stochastically, uncertainty would become a state variable in valuation. The model also takes specialization as given. Letting specialists learn about peer firms at a cost would tie

the size of the wedge to investor attention. Finally, the empirical illustration defines peers coarsely through two-digit SIC industries. Measuring peer characteristics across closer economic links, such as supply chains or product-market networks, should identify the wedge more precisely.

References

Admati, A. R. (1985). A noisy rational expectations equilibrium for multi-asset securities markets. *Econometrica 53*(3), 629–657.

Albagli, E., C. Hellwig, and A. Tsyvinski (2024). Information aggregation with asymmetric asset payoffs. *Journal of Finance 79*(4), 2715–2758.

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.

Asness, C., R. B. Porter, and R. Stevens (2000). Predicting stock returns using industry-relative firm characteristics. Working paper, University of Florida.

Bai, J., T. Philippon, and A. Savov (2016). Have financial markets become more informative? *Journal of Financial Economics 122*(3), 625–654.

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

Bond, P., A. Edmans, and I. Goldstein (2012). The real effects of financial markets. *Annual Review of Financial Economics 4*(1), 339–360.

Bordalo, P., N. Gennaioli, R. La Porta, and A. Shleifer (2019). Diagnostic expectations and stock returns. *Journal of Finance 74*(6), 2839–2874.

Brennan, M. J. and H. H. Cao (1997). International portfolio investment flows. *Journal of Finance 52*(5), 1851–1880.

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. (2011). Presidential address: Discount rates. *Journal of Finance 66*(4), 1047–1108.

Cohen, L. and A. Frazzini (2008). Economic links and predictable returns. *Journal of Finance 63*(4), 1977–2011.

Cohen, R. B. and C. Polk (1996). An investigation of the impact of industry factors in asset-pricing tests. Working paper WP96-002, Harvard Business School.

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

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

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

De Bondt, W. F. M. and R. Thaler (1985). Does the stock market overreact? *Journal of Finance 40*(3), 793–805.

Dumas, B., A. Kurshev, and R. Uppal (2009). Equilibrium portfolio strategies in the presence of sentiment risk and excess volatility. *Journal of Finance 64*(2), 579–629.

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

Engelberg, J., R. D. McLean, J. Pontiff, and M. C. Ringgenberg (2023). Do cross-sectional predictors contain systematic information? *Journal of Financial and Quantitative Analysis 58*(3), 1172–1201.

Fama, E. F. (1991). Efficient capital markets: II. *Journal of Finance 46*(5), 1575–1617.

Fama, E. F. and K. R. French (1992). The cross-section of expected stock returns. *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.

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

Goyal, A., I. Welch, and A. Zafirov (2024). A comprehensive 2022 look at the empirical performance of equity premium prediction. *Review of Financial Studies 37*(11), 3490–3557.

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

Gu, S., B. Kelly, and D. Xiu (2020). Empirical asset pricing via machine learning. *Review of Financial Studies 33*(5), 2223–2273.

Hansen, L. P. and S. F. Richard (1987). The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. *Econometrica 55*(3), 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 50*(1), 213–224.

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

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

Hou, K. (2007). Industry information diffusion and the lead-lag effect in stock returns. *Review of Financial Studies 20*(4), 1113–1138.

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 86*, 108347.

La Porta, R. (1996). Expectations and the cross-section of stock returns. *Journal of Finance 51*(5), 1715–1742.

La Porta, R., J. Lakonishok, A. Shleifer, and R. W. Vishny (1997). Good news for value stocks: Further evidence on market efficiency. *Journal of Finance 52*(2), 859–874.

Lakonishok, J., A. Shleifer, and R. W. Vishny (1994). Contrarian investment, extrapolation, and risk. *Journal of Finance 49*(5), 1541–1578.

Lewellen, J. and J. Shanken (2002). Learning, asset-pricing tests, and market efficiency. *Journal of Finance 57*(3), 1113–1145.

Liptser, R. S. and A. N. Shiryaev (2001). *Statistics of Random Processes II: Applications* (2nd ed.). Springer-Verlag, New York.

Marrow, B. and S. Nagel (2024). Real-time discovery and tracking of return-based anomalies. Working paper, University of Chicago.

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

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

Moskowitz, T. J. and M. Grinblatt (1999). Do industries explain momentum? *Journal of Finance 54*(4), 1249–1290.

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

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

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

Timmermann, A. (2006). Forecast combinations. In G. Elliott, C. W. J. Granger, and A. Timmermann (Eds.), *Handbook of Economic Forecasting*, Volume 1, Chapter 4, pp. 135–196. Elsevier.

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

Tobin, J. (1969). A general equilibrium approach to monetary theory. *Journal of Money, Credit and Banking 1*(1), 15–29.

Van Nieuwerburgh, S. and L. Veldkamp (2009). Information immobility and the home bias puzzle. *Journal of Finance 64*(3), 1187–1215.

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

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

A Appendix: Proofs

A.1 Proof of Proposition 1

Expanding (4) and taking expectations gives

$$\mathbb{E}\left[\mathrm{Cov}_{CS}(\mu^{\mathcal{H}}_{n,t}, \mathcal{W}_{n,t})\right] = \frac{1}{N}\sum_{n=1}^{N}\mathbb{E}[\mu^{\mathcal{H}}_{n,t}\mathcal{W}_{n,t}] - \frac{1}{N^2}\sum_{n=1}^{N}\sum_{m=1}^{N}\mathbb{E}[\mu^{\mathcal{H}}_{n,t}\mathcal{W}_{m,t}].$$

For $m = n$, $\mu^{\mathcal{H}}_{n,t} \in \mathcal{H}_{n,t}$ and $\mathbb{E}[\mathcal{W}_{n,t} \mid \mathcal{H}_{n,t}] = 0$ imply $\mathbb{E}[\mu^{\mathcal{H}}_{n,t}\mathcal{W}_{n,t}] = 0$. For $n \neq m$, $\mu^{\mathcal{H}}_{n,t} \in \mathcal{G}_t$, $\mathcal{W}_{m,t} = \mathbb{E}[e_{m,t} \mid \mathcal{G}_t]$, and $\mathbb{E}[e_{m,t}] = 0$ imply

$$\mathbb{E}[\mu^{\mathcal{H}}_{n,t}\mathcal{W}_{m,t}] = \mathbb{E}[\mu^{\mathcal{H}}_{n,t}e_{m,t}] = \mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, e_{m,t}).$$

From $d\mu_{n,t} = \kappa_{\mu,n}(\bar{\mu}_n - \mu_{n,t})dt + \sigma_{\mu,n}dW^{\mu}_{n,t}$ and (3), the stationary Ornstein–Uhlenbeck covariances for $n \neq m$ satisfy

$$\mathrm{Cov}(\mu_{n,t}, e_{m,t}) = \frac{\rho_{nm}\sigma_{\mu,n}\sigma_{\mu,m}}{\kappa_{\mu,n} + u_m}, \qquad \mathrm{Cov}(e_{n,t}, e_{m,t}) = \frac{\rho_{nm}\sigma_{\mu,n}\sigma_{\mu,m}}{u_n + u_m}.$$

Because $\mu^{\mathcal{H}}_{n,t} = \mu_{n,t} - e_{n,t}$, subtracting the second covariance from the first gives

$$\mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, e_{m,t}) = \rho_{nm}\sigma_{\mu,n}\sigma_{\mu,m}\frac{u_n - \kappa_{\mu,n}}{(\kappa_{\mu,n} + u_m)(u_n + u_m)}.$$

Substituting this expression into the double sum yields (5). □

A.2 Proof of Proposition 2

Under $\mathcal{G}_t$, $\mathbb{E}[dZ^{\mathcal{H}}_{n,t} \mid \mathcal{G}_t] = \mathbb{E}[dO^{\mathcal{H}}_{n,t} \mid \mathcal{G}_t] = \mathcal{W}_{n,t}dt$. Using (1) and (2), the conditional drift differences between $\mathcal{G}_t$ and $\mathcal{H}_{n,t}$ are

$$\mathbb{E}[dZ_{n,t} \mid \mathcal{G}_t] - \mathbb{E}[dZ_{n,t} \mid \mathcal{H}_{n,t}] = \mathcal{W}_{n,t}dt, \quad \mathbb{E}[d\mu^{\mathcal{H}}_{n,t} \mid \mathcal{G}_t] - \mathbb{E}[d\mu^{\mathcal{H}}_{n,t} \mid \mathcal{H}_{n,t}] = (u_n - \kappa_{\mu,n})\mathcal{W}_{n,t}dt.$$

Applying Itô's lemma to $V_{n,t} = K_{n,t}v_n(Z_{n,t}, \mu^{\mathcal{H}}_{n,t}, \bar{\nu}_n)$ in (7) gives

$$\begin{aligned} dV_{n,t} = {} & v_n dK_{n,t} + K_{n,t}\left(\frac{\partial v_n}{\partial Z_{n,t}}dZ_{n,t} + \frac{\partial v_n}{\partial \mu^{\mathcal{H}}_{n,t}}d\mu^{\mathcal{H}}_{n,t}\right) \\ & + \frac{1}{2}K_{n,t}\left(\frac{\partial^2 v_n}{\partial Z^2_{n,t}}d\langle Z_n\rangle_t + 2\frac{\partial^2 v_n}{\partial Z_{n,t}\partial \mu^{\mathcal{H}}_{n,t}}d\langle Z_n, \mu^{\mathcal{H}}_n\rangle_t + \frac{\partial^2 v_n}{\partial (\mu^{\mathcal{H}}_{n,t})^2}d\langle \mu^{\mathcal{H}}_n\rangle_t\right). \end{aligned}$$

Because payouts $K_{n,t}(Z_{n,t} - \Psi_n(i_{n,t}, Z_{n,t}))dt$, capital growth $dK_{n,t}$, and the quadratic-variation terms are identical under $\mathcal{G}_t$ and $\mathcal{H}_{n,t}$, dividing by $V_{n,t} = K_{n,t}v_n$ and taking conditional expecta-

tions yields

$$\mathbb{E}[dR^E_{n,t} \mid \mathcal{G}_t] - \mathbb{E}[dR^E_{n,t} \mid \mathcal{H}_{n,t}] = \frac{1}{v_n}\left[\frac{\partial v_n}{\partial Z_{n,t}} + (u_n - \kappa_{\mu,n})\frac{\partial v_n}{\partial \mu^{\mathcal{H}}_{n,t}}\right]\mathcal{W}_{n,t}dt.$$

Combining this equation with (8) yields (9) and (10), where $\Gamma_{n,t} > 0$ follows from $v_n > 0$, $\partial v_n/\partial Z_{n,t} > 0$, $\partial v_n/\partial \mu^{\mathcal{H}}_{n,t} > 0$, and $u_n > \kappa_{\mu,n}$. □

A.3 Proof of Proposition 3

Because $\mu^{\mathcal{H}}_{n,t} \in \mathcal{G}_t$, the law of iterated expectations and (9) give

$$\mathbb{E}\left[\mathrm{Cov}_{CS}\left(\mu^{\mathcal{H}}_{n,t}, \frac{dR^E_{n,t}}{dt} - r\right)\right] = \frac{N-1}{N^2}\sum_{n=1}^{N}\mathbb{E}[\mu^{\mathcal{H}}_{n,t}\Gamma_{n,t}\mathcal{W}_{n,t}] - \frac{1}{N^2}\sum_{n=1}^{N}\sum_{m\neq n}\mathbb{E}[\mu^{\mathcal{H}}_{n,t}\Gamma_{m,t}\mathcal{W}_{m,t}].$$

For $m = n$, $\mu^{\mathcal{H}}_{n,t}, \Gamma_{n,t} \in \mathcal{H}_{n,t}$ and $\mathbb{E}[\mathcal{W}_{n,t} \mid \mathcal{H}_{n,t}] = 0$ imply $\mathbb{E}[\mu^{\mathcal{H}}_{n,t}\Gamma_{n,t}\mathcal{W}_{n,t}] = 0$. For $n \neq m$, joint Gaussianity and $\mathbb{E}[\mathcal{W}_{m,t} \mid \mathcal{H}_{m,t}] = 0$ imply that $\mathbb{E}[\mu^{\mathcal{H}}_{n,t}\mathcal{W}_{m,t} \mid \mathcal{H}_{m,t}] = \mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, \mathcal{W}_{m,t} \mid \mathcal{H}_{m,t})$ is deterministic and equals $\mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, \mathcal{W}_{m,t}) = \mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, e_{m,t})$ by the law of total covariance. Conditioning on $\mathcal{H}_{m,t}$ and using $\Gamma_{m,t} \in \mathcal{H}_{m,t}$ gives

$$\mathbb{E}[\mu^{\mathcal{H}}_{n,t}\Gamma_{m,t}\mathcal{W}_{m,t}] = \mathbb{E}\left[\Gamma_{m,t}\mathbb{E}[\mu^{\mathcal{H}}_{n,t}\mathcal{W}_{m,t} \mid \mathcal{H}_{m,t}]\right] = \mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, e_{m,t})\mathbb{E}[\Gamma_{m,t}].$$

Substituting $\mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, e_{m,t})$ from the proof of Proposition 1 yields (11), where $\mathbb{E}[\Gamma_{m,t}] > 0$ and $u_n > \kappa_{\mu,n}$ establish the strict inequality. □

A.3.1 Proof of Corollary 3.1

Because $x_{n,t} \in \mathcal{H}_{n,t} \subseteq \mathcal{G}_t$ and $\Gamma_{m,t} \in \mathcal{H}_{m,t} \subseteq \mathcal{G}_t$, the own-firm terms ($m = n$) satisfy $\mathbb{E}[x_{n,t}\Gamma_{n,t}\mathcal{W}_{n,t}] = 0$, while for $n \neq m$ the law of iterated expectations gives $\mathbb{E}[x_{n,t}\Gamma_{m,t}\mathcal{W}_{m,t}] = \mathbb{E}[x_{n,t}\Gamma_{m,t}e_{m,t}]$. Because $(Z_{m,t}, \mu^{\mathcal{H}}_{m,t})$ and $e_{m,t}$ are jointly Gaussian and uncorrelated, $e_{m,t}$ is independent of $(Z_{m,t}, \mu^{\mathcal{H}}_{m,t})$, so $(Z_{n,t}, \mu^{\mathcal{H}}_{n,t}, e_{m,t})$ conditional on $(Z_{m,t}, \mu^{\mathcal{H}}_{m,t})$ is Gaussian with $\mathbb{E}[e_{m,t} \mid Z_{m,t}, \mu^{\mathcal{H}}_{m,t}] = 0$ and conditional covariances $\mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, e_{m,t})$ and $\mathrm{Cov}(Z_{n,t}, e_{m,t})$. Applying bivariate Stein's lemma conditional on $(Z_{m,t}, \mu^{\mathcal{H}}_{m,t})$, multiplying by $\Gamma_{m,t}$, and taking expectations gives

$$\mathbb{E}[x_{n,t}\Gamma_{m,t}e_{m,t}] = \mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, e_{m,t})\mathbb{E}\left[\Gamma_{m,t}\frac{\partial x_{n,t}}{\partial \mu^{\mathcal{H}}_{n,t}}\right] + \mathrm{Cov}(Z_{n,t}, e_{m,t})\mathbb{E}\left[\Gamma_{m,t}\frac{\partial x_{n,t}}{\partial Z_{n,t}}\right].$$

For $n \neq m$, $d\langle Z_n, e_m\rangle_t = 0$, so the stationary covariance equation for $(Z_{n,t}, e_{m,t})$ yields

$$(\kappa_{Z,n} + u_m)\,\mathrm{Cov}(Z_{n,t}, e_{m,t}) = \mathrm{Cov}(\mu_{n,t}, e_{m,t}) = \frac{\rho_{nm}\sigma_{\mu,n}\sigma_{\mu,m}}{\kappa_{\mu,n} + u_m}.$$

Substituting $\mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, e_{m,t})$ and $\mathrm{Cov}(Z_{n,t}, e_{m,t})$ into the expansion of (4) yields (12). □

A.4 Proof of Proposition 4

Under common parameters $(\kappa_Z, \sigma_Z, \kappa_\mu, \bar{\mu}, \sigma_\mu, \delta, \Lambda)$ and cost function Ψ, with $\rho_{nm} = \rho \in (0,1)$ for $n \neq m$, (1) implies $\bar{\nu} = \sigma_\mu^2/(u + \kappa_\mu)$. Because $\mu_{n,t} = \mu^{\mathcal{H}}_{n,t} + e_{n,t}$ with $\mu^{\mathcal{H}}_{n,t} \perp e_{n,t}$, the stationary moments of local beliefs are

$$\mathrm{Var}(\mu^{\mathcal{H}}_{n,t}) = \frac{\sigma_\mu^2}{2\kappa_\mu} - \frac{\sigma_\mu^2}{u + \kappa_\mu} = \frac{\sigma_\mu^2(u - \kappa_\mu)}{2\kappa_\mu(u + \kappa_\mu)}, \quad \mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, \mu^{\mathcal{H}}_{m,t}) = \rho\frac{u - \kappa_\mu}{u}\,\mathrm{Var}(\mu^{\mathcal{H}}_{n,t}) \quad (n \neq m).$$

Specializing (5) and (11) and using $\mathbb{E}[\mathrm{Var}_{CS}(\mu^{\mathcal{H}}_{n,t})] = (1 - N^{-1})[\mathrm{Var}(\mu^{\mathcal{H}}_{n,t}) - \mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, \mu^{\mathcal{H}}_{m,t})]$ gives

$$\mathbb{E}\left[\mathrm{Cov}_{CS}(\mu^{\mathcal{H}}_{n,t}, \mathcal{W}_{n,t})\right] = -\left(1 - \frac{1}{N}\right)\rho\frac{\kappa_\mu}{u}\,\mathrm{Var}(\mu^{\mathcal{H}}_{n,t}),$$

$$\mathbb{E}\left[\mathrm{Var}_{CS}(\mu^{\mathcal{H}}_{n,t})\right] = \left(1 - \frac{1}{N}\right)\frac{(1-\rho)u + \rho\kappa_\mu}{u}\,\mathrm{Var}(\mu^{\mathcal{H}}_{n,t}).$$

Dividing these expressions yields (14) and (13) in part (a). By joint Gaussianity and exchangeability across firms, conditional on $\mu^{\mathcal{H}}_t \equiv (\mu^{\mathcal{H}}_{1,t}, \ldots, \mu^{\mathcal{H}}_{N,t})$,

$$\mathbb{E}[\mathcal{W}_{n,t} - \bar{\mathcal{W}}_t \mid \mu^{\mathcal{H}}_t] = b_{\mathcal{W}}(\mu^{\mathcal{H}}_{n,t} - \bar{\mu}^{\mathcal{H}}_t).$$

For part (b), $v_{n,t} > 0$ and $\partial v_{n,t}/\partial \mu^{\mathcal{H}}_{n,t} > 0$ imply $\partial \log v_{n,t}/\partial \mu^{\mathcal{H}}_{n,t} > 0$, so $(Z_{n,t}, \mu^{\mathcal{H}}_{n,t}) \mapsto (Z_{n,t}, \log v_{n,t})$ is invertible with $\mu^{\mathcal{H}}_{n,t} = \mu^{\mathcal{H}}(\log v_{n,t}, Z_{n,t})$ and partial derivatives

$$\left.\frac{\partial \mu^{\mathcal{H}}_{n,t}}{\partial \log v_{n,t}}\right|_{Z_{n,t}} = \left(\frac{\partial \log v_{n,t}}{\partial \mu^{\mathcal{H}}_{n,t}}\right)^{-1} > 0, \qquad \left.\frac{\partial \mu^{\mathcal{H}}_{n,t}}{\partial Z_{n,t}}\right|_{\log v_{n,t}} = -\frac{\partial v_{n,t}/\partial Z_{n,t}}{\partial v_{n,t}/\partial \mu^{\mathcal{H}}_{n,t}} < 0.$$

Thus $\sigma\{(\log v_{m,t}, Z_{m,t})_{m=1}^N\} = \sigma\{(\mu^{\mathcal{H}}_{m,t}, Z_{m,t})_{m=1}^N\} \subset \mathcal{G}_t$. Under equicorrelation, the cross-sectional averages $(\bar{\mu}^{\mathcal{H}}_t, \bar{Z}_t)$ are independent of the deviations $(\mu^{\mathcal{H}}_{m,t} - \bar{\mu}^{\mathcal{H}}_t, Z_{m,t} - \bar{Z}_t)_{m=1}^N$, so (9) gives

$$\begin{aligned}\mu^{CS}_{R,n,t} &\equiv \mathbb{E}[dR^E_{n,t} \mid (\log v_{m,t}, Z_{m,t})_{m=1}^N]/dt - r \\ &= \Gamma_{n,t}\Big[\mathbb{E}[\bar{\mathcal{W}}_t \mid \bar{\mu}^{\mathcal{H}}_t, \bar{Z}_t] + b_{\mathcal{W},\mu}(\mu^{\mathcal{H}}(\log v_{n,t}, Z_{n,t}) - \bar{\mu}^{\mathcal{H}}_t) + b_{\mathcal{W},Z}(Z_{n,t} - \bar{Z}_t)\Big].\end{aligned}$$

For $m \neq n$, define $V_\mu \equiv \mathrm{Var}(\mu^{\mathcal{H}}_{n,t}) - \mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, \mu^{\mathcal{H}}_{m,t})$, $V_Z \equiv \mathrm{Var}(Z_{n,t}) - \mathrm{Cov}(Z_{n,t}, Z_{m,t})$, and $C_{\mu Z} \equiv \mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, Z_{n,t}) - \mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, Z_{m,t})$. Using $\mathrm{Cov}(\mu_{n,t}, Z_{n,t}) = \frac{\sigma^2_\mu}{2\kappa_\mu(\kappa_\mu+\kappa_Z)}$, $\mathrm{Cov}(e_{n,t}, Z_{n,t}) = 0$, and $\mathrm{Cov}(Z_{n,t}, e_{m,t}) = \frac{\rho\sigma^2_\mu}{(u+\kappa_\mu)(u+\kappa_Z)}$ gives

$$V_Z = \frac{\sigma^2_Z}{2\kappa_Z} + \frac{(1-\rho)\sigma^2_\mu}{2\kappa_\mu\kappa_Z(\kappa_\mu+\kappa_Z)}, \qquad C_{\mu Z} = \sigma^2_\mu\left[\frac{1-\rho}{2\kappa_\mu(\kappa_\mu+\kappa_Z)} + \frac{\rho}{(u+\kappa_\mu)(u+\kappa_Z)}\right] > 0.$$

Because $1 - N^{-1}$ cancels in the bivariate normal equations for $\mathcal{W}_{n,t}$ on $(\mu^{\mathcal{H}}_{n,t}, Z_{n,t})$, solving them yields

$$b_{\mathcal{W},Z} = \frac{\rho(1-\rho)\sigma^4_\mu(u-\kappa_\mu)^2(u-\kappa_Z)}{4\kappa_\mu u(u+\kappa_\mu)^2(u+\kappa_Z)(\kappa_\mu+\kappa_Z)(V_\mu V_Z - C^2_{\mu Z})}, \qquad b_{\mathcal{W},\mu} = b_{\mathcal{W}} - \frac{C_{\mu Z}}{V_\mu} b_{\mathcal{W},Z},$$

with $V_\mu V_Z - C^2_{\mu Z} > 0$ by Cauchy–Schwarz. Differentiating $\mu^{CS}_{R,n,t}$ holding $(\bar{\mu}^{\mathcal{H}}_t, \bar{Z}_t)$ fixed and taking expectations yields $b_v \equiv \mathbb{E}[\partial\mu^{CS}_{R,n,t}/\partial \log v_{n,t}|_{Z_{n,t},\bar{\mu}^{\mathcal{H}}_t,\bar{Z}_t}]$ and $b_Z \equiv \mathbb{E}[\partial\mu^{CS}_{R,n,t}/\partial Z_{n,t}|_{\log v_{n,t},\bar{\mu}^{\mathcal{H}}_t,\bar{Z}_t}]$ in (15), because for $g_{n,t} \in \{\partial\Gamma_{n,t}/\partial \log v_{n,t}|_{Z_{n,t}}, \partial\Gamma_{n,t}/\partial Z_{n,t}|_{\log v_{n,t}}\}$, $\sigma(\log v_{n,t}, Z_{n,t})$-measurability of $g_{n,t}$ and $\mathbb{E}[\mathcal{W}_{n,t} \mid \mathcal{H}_{n,t}] = 0$ give $\mathbb{E}[g_{n,t}\mathbb{E}[\mathcal{W}_{n,t} \mid (\log v_{m,t}, Z_{m,t})^N_{m=1}]] = \mathbb{E}[g_{n,t}\mathcal{W}_{n,t}] = 0$. When $u \geq \kappa_Z$, $b_{\mathcal{W},Z} \geq 0$ and $C_{\mu Z}/V_\mu > 0$ imply $b_{\mathcal{W},\mu} \leq b_{\mathcal{W}} < 0$, so $b_v < 0$ and $b_Z > 0$. When $\log v_{n,t} = \alpha_0 + \alpha_Z Z_{n,t} + \alpha_\mu \mu^{\mathcal{H}}_{n,t}$ is affine ($\alpha_Z, \alpha_\mu > 0$), (10) gives constant $\Gamma_{n,t} = \alpha_Z + (u-\kappa_\mu)\alpha_\mu > 0$ and Corollary 3.1 implies $\mathbb{E}[\mathrm{Cov}_{CS}(x_{n,t}, dR^E_{n,t}/dt - r)] = \Gamma_{n,t}\mathbb{E}[\mathrm{Cov}_{CS}(x_{n,t}, \mathcal{W}_{n,t})]$ for $x_{n,t} \in \{\log v_{n,t}, Z_{n,t}\}$, so the population OLS slopes on $(\log v_{n,t}, Z_{n,t})$ equal (b_v, b_Z). In Table 1, $\mathcal{W}_{n,t} = 0$ gives zeros in Columns (1) and (2), (8) gives zeros in Column (3a) of Panels A and B, and conditioning on $\mathcal{H}_{m,t}$ gives

$$\mathbb{E}[\Gamma_{m,t}\mathcal{W}_{n,t}\mathcal{W}_{m,t}] = \mathbb{E}[\Gamma_{n,t}]\mathbb{E}[\mathcal{W}_{n,t}\mathcal{W}_{m,t}],$$

which yields the slope $\mathbb{E}[\Gamma_{n,t}] > 0$ on $\mathcal{W}_{n,t}$ in Panel C. For part (c), define the sufficient local signal $dS_{n,t} \equiv (\Lambda + \sigma_Z^{-2})^{-1}[\sigma_Z^{-2}(dZ_{n,t} + \kappa_Z Z_{n,t}dt) + \Lambda dO_{n,t}] = \mu_{n,t}dt + (\Lambda + \sigma_Z^{-2})^{-1/2}dW^S_{n,t}$ (so $d\mu^{\mathcal{H}}_{n,t} = \kappa_\mu(\bar{\mu} - \mu^{\mathcal{H}}_{n,t})dt + (u-\kappa_\mu)(dS_{n,t} - \mu^{\mathcal{H}}_{n,t}dt)$ decays at rate u). Under equicorrelation, the cross-sectional average shocks and deviation shocks have zero quadratic covariation, so the N-firm pooled filter $\mu^{\mathcal{G}}_{n,t} = \bar{\mu}^{\mathcal{G}}_t + (\mu^{\mathcal{G}}_{n,t} - \bar{\mu}^{\mathcal{G}}_t)$ decouples into two orthogonal scalar filters with decay rates $u_0 \equiv \sqrt{\kappa^2_\mu + (\Lambda + \sigma_Z^{-2})\sigma^2_\mu(1 + (N-1)\rho)} > u$ and $\tilde{u} \equiv \sqrt{\kappa^2_\mu + (\Lambda + \sigma_Z^{-2})\sigma^2_\mu(1-\rho)} < u$:

$$d\bar{\mu}^{\mathcal{G}}_t = \kappa_\mu(\bar{\mu} - \bar{\mu}^{\mathcal{G}}_t)dt + (u_0 - \kappa_\mu)(d\bar{S}_t - \bar{\mu}^{\mathcal{G}}_t dt),$$

$$d(\mu^{\mathcal{G}}_{n,t} - \bar{\mu}^{\mathcal{G}}_t) = -\kappa_\mu(\mu^{\mathcal{G}}_{n,t} - \bar{\mu}^{\mathcal{G}}_t)dt + (\tilde{u} - \kappa_\mu)(dS_{n,t} - d\bar{S}_t - (\mu^{\mathcal{G}}_{n,t} - \bar{\mu}^{\mathcal{G}}_t)dt).$$

Because $\mu_{n,t} - \mu^{\mathcal{G}}_{n,t} \perp \mathcal{W}_{n,t}$, the steady-state pooled posterior variance and wedge variance are

$$\mathrm{Var}(\mu_{n,t} \mid \mathcal{G}_t) = \frac{1}{\Lambda + \sigma_Z^{-2}} \left[\frac{u_0 - \kappa_\mu}{N} + \left(1 - \frac{1}{N}\right)(\tilde{u} - \kappa_\mu) \right],$$

$$\mathrm{Var}(\mathcal{W}_{n,t}) = \frac{1}{\Lambda + \sigma_Z^{-2}} \left[u - \frac{u_0}{N} - \left(1 - \frac{1}{N}\right)\tilde{u} \right].$$

Because $u_0/N = O(N^{-1/2}) \to 0$ as $N \to \infty$, taking the limit yields (16). □

A.4.1 Cross-Sectional Regression Slopes in the General Case

In the general economy of Propositions 1–3 and Corollary 3.1, each firm's local filter loads on independent signal shocks $(W^Z_{n,t}, W^O_{n,t})$ with $\bar{\nu}_n > 0$, so $\mathbb{E}[\mathrm{Var}_{CS}(\mu^{\mathcal{H}}_{n,t})] > 0$ for every $N \geq 2$. Specifically, using $\mu^{\mathcal{H}}_{n,t} = \mu_{n,t} - e_{n,t}$ and $\mathbb{E}[\mu^{\mathcal{H}}_{n,t}] = \bar{\mu}_n$,

$$\mathbb{E}\left[\mathrm{Var}_{CS}(\mu^{\mathcal{H}}_{n,t})\right] = \mathrm{Var}_{CS}(\bar{\mu}_n) + \frac{N-1}{N^2}\sum_{n=1}^{N}\mathrm{Var}(\mu^{\mathcal{H}}_{n,t}) - \frac{1}{N^2}\sum_{n\neq m}\mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, \mu^{\mathcal{H}}_{m,t}) > 0,$$

where $\mathrm{Var}(\mu^{\mathcal{H}}_{n,t}) = \frac{\sigma^2_{\mu,n}(u_n - \kappa_{\mu,n})}{2\kappa_{\mu,n}(u_n + \kappa_{\mu,n})}$ and, for $n \neq m$,

$$\mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, \mu^{\mathcal{H}}_{m,t}) = \rho_{nm}\sigma_{\mu,n}\sigma_{\mu,m}\frac{(u_n - \kappa_{\mu,n})(u_m - \kappa_{\mu,m})(\kappa_{\mu,n} + \kappa_{\mu,m} + u_n + u_m)}{(\kappa_{\mu,n} + \kappa_{\mu,m})(\kappa_{\mu,n} + u_m)(u_n + \kappa_{\mu,m})(u_n + u_m)}.$$

Dividing (5), (11), and (12) by the regressor dispersion yields the population univariate slopes

$$b_{\mathcal{W}} = -\frac{1}{N^2\mathbb{E}[\mathrm{Var}_{CS}(\mu^{\mathcal{H}}_{n,t})]}\sum_{n\neq m}\mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, e_{m,t}),$$

$$b_\mu = -\frac{1}{N^2\mathbb{E}[\mathrm{Var}_{CS}(\mu^{\mathcal{H}}_{n,t})]}\sum_{n\neq m}\mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, e_{m,t})\mathbb{E}[\Gamma_{m,t}],$$

$$b_x = -\frac{1}{N^2\mathbb{E}[\mathrm{Var}_{CS}(x_{n,t})]}\sum_{n\neq m}\left\{\mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, e_{m,t})\mathbb{E}\left[\Gamma_{m,t}\frac{\partial x_{n,t}}{\partial \mu^{\mathcal{H}}_{n,t}}\right] + \mathrm{Cov}(Z_{n,t}, e_{m,t})\mathbb{E}\left[\Gamma_{m,t}\frac{\partial x_{n,t}}{\partial Z_{n,t}}\right]\right\},$$

where $\mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, e_{m,t})$ and $\mathrm{Cov}(Z_{n,t}, e_{m,t})$ are given in the proofs of Proposition 1 and Corollary 3.1. Whenever $\rho_{nm} \geq 0$ for all $n \neq m$ (with $\rho_{nm} > 0$ for at least one pair), $b_{\mathcal{W}} < 0$ and $b_\mu < 0$, and $b_x < 0$ for any differentiable characteristic with $\mathbb{E}[\mathrm{Var}_{CS}(x_{n,t})] > 0$, $\partial x_{n,t}/\partial \mu^{\mathcal{H}}_{n,t} > 0$, and $\partial x_{n,t}/\partial Z_{n,t} \geq 0$ (such as $x_{n,t} = \log v_{n,t}$).

Next, suppose firms share common parameters $(\kappa_Z, \sigma_Z, \kappa_\mu, \bar{\mu}, \sigma_\mu, \delta, \Lambda)$ and cost function Ψ, while drift shocks have an arbitrary correlation matrix (ρ_{nm}) with average pairwise correlation $\bar{\rho} \equiv \frac{1}{N(N-1)}\sum_{n\neq m}\rho_{nm} \in (0,1)$. Because each local filter conditions only on $\mathcal{H}_{n,t}$, all own-firm

second moments and $\mathbb{E}[\Gamma_{n,t}]$ are identical across firms and independent of (ρ_{nm}), whereas every cross-firm covariance $(n \neq m)$ among $(\mu^{\mathcal{H}}_{n,t}, Z_{n,t})$ and with $\mathcal{W}_{m,t}$ is proportional to ρ_{nm}. Averaging across pairs in $\mathbb{E}[\mathrm{Cov}_{CS}(\cdot,\cdot)]$ therefore replaces ρ by $\bar{\rho}$ in (14), (13), and the bivariate wedge slopes $(b_{\mathcal{W},\mu}, b_{\mathcal{W},Z})$ from the proof of part (b) above, preserving $b_{\mathcal{W},\mu} \leq b_{\mathcal{W}} < 0$ and $b_{\mathcal{W},Z} \geq 0$ when $u \geq \kappa_Z$. In particular, when $\log v_{n,t}$ is affine in $(Z_{n,t}, \mu^{\mathcal{H}}_{n,t})$, the population cross-sectional OLS slopes (b_v, b_Z) on $(\log v_{n,t}, Z_{n,t})$ equal (15) with ρ replaced by $\bar{\rho}$, giving $b_v < 0$ and $b_Z > 0$ for $u \geq \kappa_Z$.

Finally, in Table 1, all zeros hold under arbitrary firm heterogeneity: when $\rho_{nm} = 0$ for all $n \neq m$ (Column 1) or $\mathcal{F}_{n,t} = \mathcal{G}_t$ (Column 2), the wedge vanishes identically for every firm, and in firm n's own time series (Column 3a), local efficiency (8) sets the slopes and marginal effects of all $\mathcal{H}_{n,t}$-measurable variables to zero, while conditioning on $\mathcal{H}_{n,t}$ gives $\mathrm{Cov}(\mathcal{W}_{n,t}, dR^E_{n,t}/dt - r)/\mathrm{Var}(\mathcal{W}_{n,t}) = \mathbb{E}[\Gamma_{n,t}] > 0$ whenever $\mathrm{Var}(\mathcal{W}_{n,t}) > 0$. In the cross-section (Column 3b), $b_\mu < 0$ holds globally under arbitrary heterogeneity by Proposition 3. In Panel C, conditioning on $\mathcal{H}_{m,t}$ gives

$$\frac{\mathbb{E}\left[\mathrm{Cov}_{CS}(\mathcal{W}_{n,t}, dR^E_{n,t}/dt - r)\right]}{\mathbb{E}\left[\mathrm{Var}_{CS}(\mathcal{W}_{n,t})\right]} = \frac{N^{-1}\sum_{m=1}^{N}\mathbb{E}[\Gamma_{m,t}]\,\mathbb{E}[\mathcal{W}_{m,t}(\mathcal{W}_{m,t} - \bar{\mathcal{W}}_t)]}{\mathbb{E}\left[\mathrm{Var}_{CS}(\mathcal{W}_{n,t})\right]},$$

which equals $\mathbb{E}[\Gamma_{n,t}] > 0$ under common firm parameters (for any correlation matrix with $\mathbb{E}[\mathrm{Var}_{CS}(\mathcal{W}_{n,t})] > 0$) and remains strictly positive under moderate dispersion in $\mathbb{E}[\Gamma_{m,t}]$. In Panel B, $(Z_{n,t}, \mu^{\mathcal{H}}_{n,t}) \mapsto (Z_{n,t}, \log v_{n,t})$ remains invertible for each firm n, and at the homogeneous equicorrelated benchmark with $u \geq \kappa_Z$ both inequalities $b_v < 0$ and $b_Z > 0$ in (15) are strict. Under continuity of $(v_n, \partial v_n/\partial Z_{n,t}, \partial v_n/\partial \mu^{\mathcal{H}}_{n,t})$ and the stationary Gaussian projection coefficients in firm parameters and (ρ_{nm}), the expected cross-sectional marginal effects (holding $(\bar{\mu}^{\mathcal{H}}_t, \bar{Z}_t)$ fixed via uniform peer adjustments) remain strictly $b_{v,n} < 0$ and $b_{Z,n} > 0$ under moderate firm heterogeneity.

A.5 Proof of Proposition 5

Assume admissible investment rates are bounded ($i_{n,s} \in [\underline{i}_n, \bar{i}_n]$, so $0 < K_{n,0}e^{(\underline{i}_n - \delta_n)t} \leq K_{n,t} \leq K_{n,0}e^{(\bar{i}_n - \delta_n)t}$), an equilibrium exists for each $M \geq 0$ with $\Pi(M) < \infty$ continuous in M and uniformly integrable discounted integrand in (18), and $\mathbb{E}[(U_{n,t} - U_{n,s})^2] \leq L_{n,T}|t-s|$ on each $[0,T]$ with $L_{n,T} > 0$ independent of M. Define $X_{n,t} \equiv K^{-1}_{n,t}\int_t^\infty e^{-r(s-t)}K_{n,s}(Z_{n,s} - \Psi_n(i_{n,s}, Z_{n,s}))ds$, so $U_{n,t} = \mathbb{E}[X_{n,t} \mid \mathcal{G}_t \vee \mathcal{F}_{n,t}]$ and $v_{n,t} = \mathbb{E}[X_{n,t} \mid \mathcal{F}_{n,t}] = \mathbb{E}[U_{n,t} \mid \mathcal{F}_{n,t}]$. An entrant with zero initial holdings ($H_{n,0} = 0$) chooses $\mathcal{G}_t \vee \mathcal{F}_{n,t}$-adapted trading rates $\{\theta_{n,t}\}_{t\geq 0}$, with cumulative holdings

$H_{n,s} \equiv \int_0^s \theta_{n,t} dt$ satisfying $\mathbb{E}[\int_0^\infty \int_0^s e^{-rs} |\theta_{n,t} K_{n,s}(Z_{n,s} - \Psi_n(i_{n,s}, Z_{n,s}))| dt ds] < \infty$, to maximize

$$\mathbb{E}\left[\sum_{n=1}^{N} \int_0^\infty e^{-rs} \left(H_{n,s} K_{n,s}(Z_{n,s} - \Psi_n(i_{n,s}, Z_{n,s})) - K_{n,s} v_{n,s} \theta_{n,s} - \frac{\gamma_n}{2} K_{n,s} \theta_{n,s}^2 \right) ds \right].$$

Applying Fubini's theorem to the payout integral, conditioning on $\mathcal{G}_t \vee \mathcal{F}_{n,t}$ at each date t, and completing the square rewrites this problem as

$$\max_{\{\theta_{n,t}\}_{t \geq 0}} \mathbb{E}\left[\sum_{n=1}^{N} \int_0^\infty e^{-rt} K_{n,t} \left(\frac{(U_{n,t} - v_{n,t})^2}{2\gamma_n} - \frac{\gamma_n}{2} \left(\theta_{n,t} - \frac{U_{n,t} - v_{n,t}}{\gamma_n} \right)^2 \right) dt \right].$$

Because each arbitrageur is atomistic, $K_{n,t} > 0$, and $(U_{n,t} - v_{n,t})/\gamma_n$ is $\mathcal{G}_t \vee \mathcal{F}_{n,t}$-measurable, setting the squared term to zero pointwise gives $\theta^*_{n,t} = (U_{n,t} - v_{n,t})/\gamma_n$ and (18) in part (a). For part (b), define the $\mathcal{F}_{n,t}$-adapted process $Q_{n,t} \equiv Y_{n,t} + \frac{M}{\gamma_n} \int_0^t v_{n,s} ds$, which satisfies $dQ_{n,s} = \frac{M}{\gamma_n} U_{n,s} ds + \sigma_{Y,n} dW^Y_{n,s}$ under $\theta_{n,s} = \theta^*_{n,s}$ by (17), and for $M > 0$ and $h \in (0, t)$ define the $\mathcal{F}_{n,t}$-measurable estimator

$$\widehat{U}^h_{n,t} \equiv \frac{\gamma_n}{Mh}(Q_{n,t} - Q_{n,t-h}) = \frac{1}{h} \int_{t-h}^{t} U_{n,s} ds + \frac{\gamma_n \sigma_{Y,n}}{Mh}(W^Y_{n,t} - W^Y_{n,t-h}).$$

Because $v_{n,t} = \mathbb{E}[U_{n,t} \mid \mathcal{F}_{n,t}]$ minimizes mean squared error over $\mathcal{F}_{n,t}$-measurable variables, applying $(a-b)^2 \leq 2a^2 + 2b^2$, Jensen's inequality, and $\mathbb{E}[(U_{n,t} - U_{n,s})^2] \leq L_{n,T}|t-s|$ on $[0,T]$ $(T \geq t)$ gives

$$\mathbb{E}\left[(U_{n,t} - v_{n,t})^2\right] \leq \mathbb{E}\left[(U_{n,t} - \widehat{U}^h_{n,t})^2\right] \leq L_{n,T} h + \frac{2\gamma_n^2 \sigma^2_{Y,n}}{M^2 h}.$$

Setting $h = M^{-1}$ for $M > t^{-1}$ yields $\mathbb{E}[(U_{n,t} - v_{n,t})^2] \leq (L_{n,T} + 2\gamma_n^2 \sigma^2_{Y,n}) M^{-1} = O(M^{-1})$. Because $K_{n,t} \leq K_{n,0} e^{(\bar{\imath}_n - \delta_n)t}$, $\mathbb{E}[K_{n,t}(U_{n,t} - v_{n,t})^2] \to 0$ for every $t > 0$, and uniform integrability of the discounted integrand in (18) gives (19). For part (c), continuity of $\Pi(M)$ and (19) imply a finite free-entry mass $M^* < \infty$ ($M^* = 0$ if $c \geq \Pi(0)$, and $\Pi(M^*) = c$ by the intermediate value theorem if $0 < c < \Pi(0)$). Because W^Y is independent of fundamentals, $\mathbb{E}[\mu_{n,t} \mid \mathcal{G}_t \vee \mathcal{F}_{n,t}] = \mu^{\mathcal{G}}_{n,t}$, so $\mu^{\mathcal{F}}_{n,t} = \mathbb{E}[\mu^{\mathcal{G}}_{n,t} \mid \mathcal{F}_{n,t}]$ by iterated expectations since $\mathcal{F}_{n,t} \subseteq \mathcal{G}_t \vee \mathcal{F}_{n,t}$, and $\mathcal{W}^{\mathcal{F}}_{n,t} = \mathcal{W}_{n,t} - \mathbb{E}[\mathcal{W}_{n,t} \mid \mathcal{F}_{n,t}]$. At any $M^* < \infty$, $K_{n,s} \geq K_{n,0} e^{(\underline{\imath}_n - \delta_n)s}$ and $\Pi(M^*) < \infty$ imply $\int_0^t \mathbb{E}[(U_{n,s} - v_{n,s})^2] ds < \infty$, so conditional mutual information transmitted by the additive-noise channel (17) over $[0,t]$ given $\mathcal{H}_{n,t}$ is at most $\frac{(M^*)^2}{2\gamma_n^2 \sigma^2_{Y,n}} \int_0^t \mathbb{E}[(U_{n,s} - v_{n,s})^2] ds < \infty$ (Liptser and Shiryaev, 2001). Because $\mathcal{W}_{n,t}$ is Gaussian and independent of $\mathcal{H}_{n,t}$, the Gaussian rate-distortion lower bound (Liptser and

Shiryaev, 2001) yields

$$\mathrm{Var}(\mathcal{W}^{\mathcal{F}}_{n,t}) \geq \mathrm{Var}(\mathcal{W}_{n,t}) \exp\left(-\frac{(M^*)^2}{\gamma_n^2 \sigma^2_{Y,n}} \int_0^t \mathbb{E}\left[(U_{n,s} - v_{n,s})^2\right] ds\right) > 0$$

for every $t \geq 0$ whenever $\mathrm{Var}(\mathcal{W}_{n,t}) > 0$ (and conditioning further on $\mathcal{G}_{t-\Delta} \vee \mathcal{F}_{n,t-\Delta}$ over a fixed window $[t-\Delta, t]$ replaces $\mathrm{Var}(\mathcal{W}_{n,t})$ by $\mathrm{Var}(\mu^{\mathcal{G}}_{n,t} \mid \mathcal{G}_{t-\Delta} \vee \mathcal{H}_{n,t}) > 0$, keeping $\liminf_{t\to\infty} \mathrm{Var}(\mathcal{W}^{\mathcal{F}}_{n,t}) > 0$ whenever $\sup_{s\geq 0} \mathbb{E}[(U_{n,s} - v_{n,s})^2] < \infty$). □

A.6 Proof of Proposition 6

Under common parameters $(\kappa_Z, \sigma_Z, \kappa_\mu, \bar{\mu}, \sigma_\mu, \delta, \Lambda)$ and cost function Ψ, each firm has the same valuation function $v_n = v$ and marginal stationary law for $(Z_{n,t}, \mu^{\mathcal{H}}_{n,t})$, so $\mathbb{E}[\Gamma_{m,t}] = \mathbb{E}[\Gamma_{n,t}] > 0$. Because the $N \times N$ block correlation matrix has strictly positive eigenvalues $1-\rho_w$, $1-\rho_w + N_G(\rho_w - \rho_a)$, and $1 - \rho_w + N_G(\rho_w - \rho_a) + N\rho_a$, it is positive definite. From the proof of Proposition 3, $\mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, dR^E_{m,t}/dt - r)$ equals 0 for $m = n$ and $\rho_{nm}\psi\mathbb{E}[\Gamma_{n,t}]$ for $m \neq n$, where $\psi \equiv \frac{\sigma_\mu^2(u-\kappa_\mu)}{2u(u+\kappa_\mu)} > 0$. Expanding $\mathrm{Cov}^{(g)}_{CS}(x_{n,t}, y_{n,t}) \equiv N_G^{-1} \sum_{n\in g}(x_{n,t} - \bar{x}_{g,t})(y_{n,t} - \bar{y}_{g,t})$ within industry g (N_G firms, $\rho_{nm} = \rho_w$ for $m \neq n$) yields (20), while averaging within each industry $g \in \{1, \ldots, G\}$ gives

$$\mathrm{Cov}\left(\bar{\mu}^{\mathcal{H}}_{g,t}, \frac{d\bar{R}^E_{g,t}}{dt} - r\right) = \left(1 - \frac{1}{N_G}\right)\rho_w \psi\, \mathbb{E}[\Gamma_{n,t}],$$

$$\mathrm{Cov}\left(\bar{\mu}^{\mathcal{H}}_{g,t}, \frac{d\bar{R}^E_{g',t}}{dt} - r\right) = \rho_a \psi\, \mathbb{E}[\Gamma_{n,t}] \quad (g' \neq g).$$

Substituting these industry covariances into $\mathbb{E}[\mathrm{Cov}^{\mathrm{Ind}}_{CS}(\bar{\mu}^{\mathcal{H}}_{g,t}, d\bar{R}^E_{g,t}/dt - r)]$, where $\mathrm{Cov}^{\mathrm{Ind}}_{CS}(\bar{x}_{g,t}, \bar{y}_{g,t}) \equiv G^{-1}\sum_{g=1}^G (\bar{x}_{g,t} - \bar{x}_t)(\bar{y}_{g,t} - \bar{y}_t)$, gives (21) in part (a). For part (b), with common (κ_μ, σ_μ) and $\rho_{nm} > 0$, the proof of Proposition 1 gives $\mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, \mathcal{W}_{m,t}) = \rho_{nm}\sigma_\mu^2 \frac{u_n - \kappa_\mu}{(\kappa_\mu + u_m)(u_n + u_m)} > 0$. Therefore

$$\mathrm{Cov}(\mu^{\mathcal{H}}_{n,t}, \mathcal{W}_{m,t}) - \mathrm{Cov}(\mu^{\mathcal{H}}_{m,t}, \mathcal{W}_{n,t}) = \rho_{nm}\sigma_\mu^2 \frac{u_n^2 - u_m^2}{(\kappa_\mu + u_m)(\kappa_\mu + u_n)(u_n + u_m)} > 0$$

whenever $\Lambda_n > \Lambda_m$ (so $u_n > u_m$), which proves (22). Differentiating with respect to u_n and u_m,

$$\frac{\partial}{\partial u_n}\left[\frac{u_n - \kappa_\mu}{(\kappa_\mu + u_m)(u_n + u_m)}\right] = \frac{1}{(u_n + u_m)^2} > 0,$$

$$\frac{\partial}{\partial u_m}\left[\frac{u_n - \kappa_\mu}{(\kappa_\mu + u_m)(u_n + u_m)}\right] = -(u_n - \kappa_\mu)\frac{\kappa_\mu + u_n + 2u_m}{(\kappa_\mu + u_m)^2(u_n + u_m)^2} < 0.$$

Because $\partial u_k/\partial \Lambda_k = \sigma_\mu^2/(2u_k) > 0$, $\mathrm{Cov}(\mu_{n,t}^{\mathcal{H}}, \mathcal{W}_{m,t})$ strictly increases in Λ_n and strictly decreases in Λ_m. □

A.7 Proof of Proposition 7

Under a deterministic, piecewise-continuous schedule $\Lambda_{n,s}$ bounded away from zero and infinity (right-continuous with left limits) and initialized at $\nu_{n,0} = \bar{\nu}_n(\Lambda_{n,0}) > 0$, the solution $\nu_{n,s}$ of (24) is continuous, bounded, and strictly positive for all $s \geq 0$, and the local filter and forecast error follow

$$d\mu_{n,s}^{\mathcal{H}} = \kappa_{\mu,n}(\bar{\mu}_n - \mu_{n,s}^{\mathcal{H}})ds + \nu_{n,s}\left(\sigma_{Z,n}^{-2} dZ_{n,s}^{\mathcal{H}} + \Lambda_{n,s} dO_{n,s}^{\mathcal{H}}\right),$$

$$de_{n,s} = -u_{n,s}e_{n,s}ds + \sigma_{\mu,n}dW_{n,s}^{\mu} - \nu_{n,s}\left(\sigma_{Z,n}^{-1} dW_{n,s}^{Z} + \Lambda_{n,s}^{1/2} dW_{n,s}^{O}\right),$$

where $u_{n,s} \equiv \kappa_{\mu,n} + (\Lambda_{n,s} + \sigma_{Z,n}^{-2})\nu_{n,s} > \kappa_{\mu,n}$. Applying Itô's lemma to $V_{n,s} = K_{n,s}v_n(s, Z_{n,s}, \mu_{n,s}^{\mathcal{H}})$ on either side of a jump date $t > 0$ and subtracting (8) cancels all $\mathcal{H}_{n,s}$-measurable terms (including $\partial v_n/\partial s$ and the quadratic-variation terms) and yields

$$\frac{\mathbb{E}[dR_{n,s}^E \mid \mathcal{G}_s]}{ds} - r = \Gamma_{n,s}\mathcal{W}_{n,s}, \qquad \Gamma_{n,s} = \frac{1}{v_{n,s}}\left[\frac{\partial v_{n,s}}{\partial Z_{n,s}} + (\Lambda_{n,s} + \sigma_{Z,n}^{-2})\nu_{n,s}\frac{\partial v_{n,s}}{\partial \mu_{n,s}^{\mathcal{H}}}\right] > 0.$$

Because no discrete signal arrives at $s = t$, local efficiency (8) imposes value matching $v_n(t^-, Z, \mu^{\mathcal{H}}) = v_n(t^+, Z, \mu^{\mathcal{H}})$ across $(Z, \mu^{\mathcal{H}})$, so $v_{n,s}$, $\partial v_{n,s}/\partial Z_{n,s}$, and $\partial v_{n,s}/\partial \mu_{n,s}^{\mathcal{H}}$ are continuous across $s = t$ (only $\partial v_n/\partial s$ jumps). Because $\nu_{n,s}$ and $\mathcal{W}_{n,s}$ are also continuous across $s = t$ while $\Lambda_{n,t^+} - \Lambda_{n,t^-} = \overline{\Lambda}_n - \underline{\Lambda}_n > 0$, subtracting $s = t^-$ from $s = t^+$ yields (25). Furthermore, because $\{\Lambda_{k,s}\}_{s\geq 0}$ is deterministic and initialized at the Gaussian steady state, the joint system remains Gaussian with $\mathbb{E}[e_{n,s} \mid \mathcal{H}_{n,s}] = 0$ and $\mathbb{E}[dR_{n,s}^E/ds - r] = 0$. For $m \neq n$, joint Gaussianity and the law of total covariance imply that $\mathbb{E}[\mu_{m,s}^{\mathcal{H}} e_{n,s} \mid \mathcal{H}_{n,s}] = \mathrm{Cov}(\mu_{m,s}^{\mathcal{H}}, e_{n,s} \mid \mathcal{H}_{n,s})$ is deterministic and equals $C_{mn}(s) \equiv \mathrm{Cov}(\mu_{m,s}^{\mathcal{H}}, e_{n,s})$, so conditioning on $\mathcal{H}_{n,s}$ gives

$$\mathrm{Cov}\left(\mu_{n,s}^{\mathcal{H}} - \bar{\mu}_s^{\mathcal{H}}, \frac{dR_{n,s}^E}{ds} - r\right) = -\frac{1}{N}\sum_{m\neq n} C_{mn}(s)\mathbb{E}[\Gamma_{n,s}],$$

where $R_{mn}(s) \equiv \mathrm{Cov}(e_{m,s}, e_{n,s})$ and $C_{mn}(s)$ solve, for $m \neq n$,

$$\frac{dR_{mn}(s)}{ds} = -(u_{m,s} + u_{n,s})R_{mn}(s) + \rho_{nm}\sigma_{\mu,m}\sigma_{\mu,n},$$

$$\frac{dC_{mn}(s)}{ds} = -(\kappa_{\mu,m} + u_{n,s})C_{mn}(s) + (u_{m,s} - \kappa_{\mu,m})R_{mn}(s).$$

Because $u_{m,s} > \kappa_{\mu,m} > 0$ and $u_{n,s} > \kappa_{\mu,n} > 0$ are bounded and piecewise continuous, integrating factors from $R_{mn}(0), C_{mn}(0) \geq 0$ (strictly positive when $\rho_{nm} > 0$) imply that $C_{mn}(s)$ is continuous at $s = t$, nonnegative whenever $\rho_{nm} \geq 0$, and strictly positive whenever $\rho_{nm} > 0$. Consequently, $-\frac{1}{N}\sum_{m\neq n} C_{mn}(t) < 0$ is continuous across $s = t$, whereas $\Gamma_{n,t^+} - \Gamma_{n,t^-} = (\overline{\Lambda}_n - \underline{\Lambda}_n)\nu_{n,t} v_{n,t}^{-1} \partial v_{n,t}/\partial \mu_{n,t}^{\mathcal{H}} > 0$ (with $0 < \Gamma_{n,t^-} < \Gamma_{n,t^+} = \Gamma_{n,t}$ and $\mathbb{E}[\Gamma_{n,t}^2] < \infty$) implies $\mathbb{E}[\Gamma_{n,t^+}] > \mathbb{E}[\Gamma_{n,t^-}] > 0$, so $\mathrm{Cov}(\mu_{n,s}^{\mathcal{H}} - \bar{\mu}_s^{\mathcal{H}}, dR_{n,s}^E/ds - r) < 0$ becomes strictly more negative from $s = t^-$ to $s = t^+$. □

B Appendix: Data

Compustat and CRSP sample. I combine the Compustat North America Fundamentals Annual file (`comp.funda`) with the CRSP Monthly Stock File (`crsp.msf`) and Monthly Delisting File (`crsp.msedelist`) using the CRSP–Compustat historical link table (`crsp.ccmxpf_lnkhist`), requiring the link to be active on the fiscal year-end date (`datadate`). Following Fama and French (1992, 1993) and Novy-Marx (2013), each June $t \in \{1963, \ldots, 2023\}$ I match accounting data from fiscal years ending in calendar year $t-1$ (keeping the latest fiscal year-end if multiple records appear) to monthly stock returns from July t through June $t+1$. Book equity (`be`) equals common equity (`ceq`) plus balance-sheet deferred taxes (`txdb`, set to zero when missing), and operating profitability is $Z_{n,t} = \texttt{oibdp}/\texttt{be}$, where `oibdp` is operating income before depreciation. Market equity (in millions of dollars) at the end of any month is $|\texttt{PRC}| \times \texttt{SHROUT}/1000$, and market-to-book $v_{n,t}$ divides December $t-1$ CRSP market equity by fiscal year $t-1$ book equity. Monthly delisting-adjusted returns are $R_{n,\tau} = \left[(1 + \texttt{RET})(1 + \texttt{DLRET}) - 1\right] \times 100$ (treating a missing component as zero when the other is observed, and missing when both are missing). As of June t, I retain ordinary U.S. common shares (CRSP `SHRCD` 10 or 11) listed on NYSE, AMEX, or Nasdaq (`EXCHCD` 1, 2, or 3) from the CRSP historical security header file (`crsp.msenames`), exclude financial firms (`SICCD` 6000–6999) and observations without a valid positive SIC code, and require positive book equity, positive sales (`sale`), non-missing `oibdp`, positive market equity in December $t-1$ and June t, and a June t share price of at least \$5. Within each formation year t, I winsorize $\log v_{n,t}$ and $Z_{n,t}$ at the 1st and 99th percentiles, and then drop two-digit SIC industry-years ($\lfloor \texttt{SICCD}/100 \rfloor$) with fewer than five firms. The resulting panel contains 139,118 firm-years across 13,989 firms (averaging 2,281 firms across 50 two-digit SIC industries per year) and 1,627,556 firm-months across 732 months from July 1963 through June 2024 (2,223 firms per month). Across firm-years, $\log v_{n,t}$ has mean 0.639 and standard deviation 0.868 (median 0.559), $Z_{n,t}$ has mean 0.280 and standard deviation 0.346 (median 0.275), and monthly returns

$R_{n,\tau}$ have mean 1.052% and standard deviation 14.016%.

Characteristic decompositions, earnings announcements, and Fama–MacBeth inference. For each characteristic $x_{n,t} \in \{\log v_{n,t}, Z_{n,t}\}$ in two-digit SIC industry g and formation year t (with $N_{g,t} \geq 5$ firms), Industry Mean is $\bar{x}_{g,t} = N_{g,t}^{-1} \sum_{m \in g} x_{m,t}$, Within-Industry is $x_{n,t} - \bar{x}_{g,t}$, and Industry-Peer is the leave-one-out average $(N_{g,t} - 1)^{-1} \sum_{m \in g,\, m \neq n} x_{m,t}$. Past 5-Year is firm n's average of $x_{n,s}$ over prior formation years $s \in \{t-5, \ldots, t-1\}$, requiring at least three observed prior years; Column (3) of Table 2 therefore spans July 1966 through June 2024 (696 months, 1,157,985 firm-months). In Table 4, I obtain quarterly earnings announcement dates (`rdq`) from the Compustat North America Fundamentals Quarterly file (`comp.fundq`), link `gvkey` to CRSP `permno` via `crsp.ccmxpf_lnkhist` requiring the link to be active on `rdq`, and set $\text{Ann}_{n,\tau} = 1$ if firm n has a quarterly earnings announcement in month τ (and 0 otherwise). Retaining firm-years from July 1972 through June 2024 ($T = 624$ months) with at least one reported `rdq` between July t and June $t+1$ yields 1,445,121 firm-months, of which 471,155 (32.6%) are announcement months and 973,966 are non-announcement months. In Tables 2 through 4, each specification is estimated by cross-sectional ordinary least squares in every period ($T = 732$, 696, or 624 months for returns; $T = 42$ years for growth forecasts), and I report the time-series mean of the slope estimates. Standard errors use the Bartlett-kernel Newey–West estimator with lag truncation $K = 12$ for monthly regressions and $K = 3$ for annual regressions, and two-sided statistical significance (*, **, ***) is evaluated from the Student-t distribution with $T - 1$ degrees of freedom.

I/B/E/S analyst growth forecasts. I obtain consensus mean long-term annual earnings-per-share growth forecasts (`ltg`, in percent per year, requiring `ltg` > 0) from the I/B/E/S Unadjusted U.S. EPS Summary file (`ibes.statsum_epsus`, available from December 1981). I link eight-digit I/B/E/S `cusip` identifiers to CRSP `permno` via the CRSP historical security header file (`crsp.msenames`), requiring `cusip` = `ncusip` and that the I/B/E/S statistical period date (`statpers`) fall within the active name range [`namedt`, `nameendt`] (keeping the latest `statpers` if multiple snapshots appear). For Columns (1) and (2) of Table 3, the initial growth belief $\log \text{LTG}_{n,t} = \log(\texttt{ltg})$ uses the December $t-1$ consensus forecast to align with December $t-1$ market equity in $v_{n,t}$; merging with the June t formation panel over $t \in \{1982, \ldots, 2023\}$ yields 67,035 firm-years across 8,453 firms (mean forecast 17.53%, median 15.00%, covered by 4.4 analysts on average). For Columns (3) and (4), I compute the subsequent one-year forecast revision from the monthly I/B/E/S summary snapshots in June t and June $t+1$ as $\Delta \log \text{LTG}_{n,t \to t+1} = \log \text{LTG}_{n,\text{June } t+1} - \log \text{LTG}_{n,\text{June } t}$, winsorize $\Delta \log \text{LTG}_{n,t \to t+1}$ at the 1st

and 99th percentiles within each formation year across the I/B/E/S–CRSP June panel, and merge with the June t formation panel. This yields 58,522 firm-years across 7,552 firms (mean revision -0.046, standard deviation 0.354).