Benchmark Gravity: A Size-Indexed Threshold for Mandate Composition and Idiosyncratic Price Informativeness

Alejandro Lopez-Lira*
University of Florida

April 20, 2026

Abstract

A panel of 83,420 stock-years from 1990 to 2023 shows that active-mutual-fund ownership raises idiosyncratic price informativeness in small stocks and lowers it in large stocks, with quintile t-statistics moving monotonically from $+2.28$ to -3.59. The paper provides a quantitative characterization of the composition-dependent sign pattern. Tracing the effect through the composition of active ownership, concentrated stock-pickers carry the large-cap negative sign: a combined proxy (tracking error above six percent or holdings-HHI above 0.025) delivers a large-cap coefficient of -0.0041 ($t = -3.04$) and a continuous size-interaction slope of -0.0024 ($t = -2.46$). A non-mechanical tracking-error-only proxy that strips out the HHI leg refutes a tautology concern: the TE-only large-tercile coefficient is -0.0040 ($t = -4.51$), and in a horse race against an HHI-only proxy the TE leg carries the size effect (-0.0050, $t = -6.15$) while the HHI leg enters with the wrong sign. A CARA-normal noisy rational-expectations model with four mandate types (pure passive, benchmarked active, stock picker, unconstrained informed) organizes the fingerprint through a closed-form size-indexed threshold. The cutoff formula s^* and an if-and-only-if interior-crossing condition provide the theoretical organizing structure for two sub-classes: a benchmarked sub-class (Corollary 1) and a concentrated stock-picker sub-class (Corollary 2). At IBES-disciplined primitives, the benchmarked cutoff $s_B^* = 1.88$ lies outside the unit interval, so Corollary 1 characterizes a counterfactual parameter region; the empirical anchor is Corollary 2, whose concentration-driven flow noise carries the size fingerprint at calibrated primitives. The size-increasing flow noise that drives Corollary 2 is not a primitive: a separate proposition derives it from a capacity-constrained concentrated-fund portfolio problem. A data-disciplined calibration with informed precision $\tau_I^\omega(s) = 1.77 + 2.25s$ drawn from IBES analyst coverage matches the quintile gradient with $\chi^2 = 4.20$ on three degrees of freedom ($p = 0.24$). A hedge-fund 13F placebo confirms the size gradient is absent for unconstrained-informed ownership, consistent with the mandate-specific mechanism. A Bai-Philippon-Savov / Dávila-Parlatore-style transparent informativeness measure, essentially uncorrelated with the PEAD measure ($\rho = 0.005$), preserves the large-tercile negative sign ($t = -2.73$).

Keywords: Price informativeness; institutional mandates; passive investing; benchmarked active; size cross-section; noisy rational expectations.

JEL classification: G11, G14, G23.

*Warrington College of Business, University of Florida. Email: `alejandro.lopez-lira@warrington.ufl.edu`. I thank seminar participants and the pipeline reviewers for valuable feedback. All errors are my own.

1 Introduction

When active mutual funds raise their share of a stock, does price informativeness rise or fall? The answer depends on firm size. A panel of 83,420 stock-years between 1990 and 2023 shows that benchmarked-active mutual fund ownership *raises* idiosyncratic price informativeness in the smallest quintile ($t = +2.28$) and *lowers* it in the largest ($t = -3.59$), with the sign flipping monotonically across the size distribution. This paper delivers a quantitative characterization of the composition-dependent sign pattern: which ownership component drives which side, and what primitive mechanism generates the size gradient. The pattern is robust to a battery of specifications, subsamples, and alternative informativeness measures, and it holds under a continuous size interaction: a ten-percentage-point rise in active ownership is associated with a -0.0040 ($t = -5.21$) change in idiosyncratic informativeness per log unit of market capitalization. The central question of this paper is why one ownership channel can have opposite signs on price informativeness in small and large stocks.

The paper's answer has two parts. An empirical part identifies the composition of active ownership that drives the large-cap negative sign, and a theoretical part delivers a closed-form size-indexed threshold that signs the effect from primitives. On the empirical side, an ownership decomposition that splits active mutual funds into closet indexers (tracking error below four percent) and concentrated stock-pickers (tracking error above six percent or holdings-Herfindahl above 0.025, matching the high-active-share category of Cremers and Petajisto, 2009) shows the large-cap negative sign loads on the stock-picker sub-class: θ_S enters the large-cap regression at -0.0041 ($t = -3.04$) in a coarse horse race and -0.0038 ($t = -4.37$) when paired with closet-indexer share, which itself is insignificant. A continuous size interaction confirms the pattern: the stock-picker$\times$size slope is -0.0024 ($t = -2.46$) in the tight horse race. A non-mechanical proxy that strips out the holdings-HHI leg and classifies stock-pickers purely by ex-post tracking error above six percent preserves the pattern: the large-tercile coefficient is -0.0040 ($t = -4.51$), and in a horse race against an HHI-only proxy the TE leg carries the size effect (-0.0050, $t = -6.15$) while the HHI leg loads with the wrong sign. The TE-only evidence refutes a mechanical-tautology concern: the size gradient is not an artifact of HHI co-construction with position concentration.

On the theoretical side, a CARA-normal noisy rational-expectations model with four mandate types organizes the empirical fingerprint through a size-indexed threshold. The empirically operative mechanism is the stock-picker channel: concentrated stock-pickers hold large dollar positions in large stocks under a capacity constraint, rebalance those positions with benchmark-independent flow, and thereby inject per-stock noise whose variance grows with position size. In a size-s stock, the stock-picker's marginal effect on idiosyncratic informativeness therefore carries its own signal precision minus a size-indexed threshold $\bar{\tau}_S(s)$ whose growth with s comes from concentration-driven flow noise. At calibrated primitives, this threshold crosses the stock-picker's own precision inside the unit interval, producing a positive marginal effect for small stocks and a negative marginal effect for large stocks, with cutoff s_S^* consistent with the empirical zero-crossing range.

The capacity-constrained foundation is not a reduced-form assumption. Proposition 1 derives

the size-increasing stock-picker flow-noise variance $\sigma^2_{\eta S}(s)$ from a portfolio problem in which concentrated managers cap per-stock dollar positions at a size-monotone limit $\lambda(s)$ drawn from price-impact capacity (Berk-Green 2004; Pastor-Stambaugh-Taylor 2015), and in which the mass of concentrated funds holding size-s stocks rises with s following the Cremers-Petajisto evidence. Squaring a size-monotone position and summing independent rebalancing flows across funds delivers $\sigma^2_{\eta S}(s) = \Theta(s)\lambda(s)^2\sigma^2$, strictly increasing under minimal primitives. The size-gradient in the data is therefore a prediction of a microfounded portfolio problem, not a scope assumption.

The theoretical structure of the paper is organized around the empirically operative cutoff. The stock-picker cutoff s^*_S (Proposition 2) admits a closed-form expression in terms of primitives $(\tau_S, \theta_S, \theta_B, \rho_S, \rho_B, \sigma^2_{z\omega}, \sigma^2_{\eta B}, \sigma^2_{\eta S}(\cdot))$ and lies inside the unit interval at the IBES-disciplined calibration. The benchmarked-sub-class cutoff s^*_B (Proposition 3) and the *if-and-only-if* interior-crossing condition (Proposition 4) provide the theoretical organizing structure that applies to both sub-classes, but at the IBES-disciplined calibration the benchmarked-only cutoff lies outside the unit interval ($s^*_B = 1.88$), so Proposition 3 applied to Corollary 1 describes a theoretical counterfactual parameter region. The empirical anchor is Corollary 2. The joint-model monotonicity condition (Lemma 5) ties the two sub-classes together. The derivative of $\tau^\omega_\pi(s) = M(s)^2/N(s)$ with respect to any mandate share is a one-step quotient-rule identity (Lemma 4); I use it as an organizing device and do not claim it as a theorem.

The calibration is disciplined by data, not picked to fit the gradient. Informed signal precision $\tau^\omega_I(s) = 1.77 + 2.25s$ is identified from the size-bucketed mean of log I/B/E/S analyst coverage and is not adjusted in the fitting step. Two free parameters, benchmarked flow-noise variance and a unit-scale factor for PEAD-to-precision conversion, are fit by method of moments to the five empirical quintile coefficients. The resulting $\chi^2 = 4.20$ on three degrees of freedom ($p = 0.24$) is not rejected, and the calibrated model-to-data ratio is 0.69 in the small-cap slope and 0.74 in the large-cap slope, eliminating the two-fold magnitude gap reported in the previous draft. Appendix B gives details.

The paper relates to three strands of the literature. On benchmarking and information acquisition, Breugem and Buss (2019) study a one-asset economy with a single active type under a tracking-error penalty. Their model delivers a benchmarking-reduces-informativeness comparative static but has no factor decomposition, no mandate-type distinction between closet indexers and stock pickers, no closed-form size cutoff, and no size conditionality in the sign. The contribution here is the multi-asset cross-section, the stock-picker sub-class with concentration flow noise, and the size-indexed threshold. On benchmarking and the size cross-section of returns, Pavlova and Sikorskaya (2023) document an interaction of benchmarking intensity with firm size for returns and factor loadings but do not study price informativeness, do not derive the interaction from an information-economics model, and do not deliver a threshold characterization. The informativeness outcome and the threshold derivation are the new content here. On passive investing and informational efficiency, Buss and Sundaresan (2023) study a single-asset economy and Cong et al. (2024) study the extensive margin of index composition; neither derives an opposite-signed predic-

tion across firm size within a single mandate type, which is the central empirical fingerprint of this paper.

On informativeness measurement, the post-earnings-announcement-drift construction used here is methodologically distinct from the prices-leading-earnings measure of Bai et al. (2016) and the identification approach of Dávila and Parlatore (2025). Over the 1990-2023 window the PEAD measure does not reproduce the negative sign of Sammon (2025) for pure-passive ownership on aggregate informativeness. The paper does not claim to resolve that disagreement; the cross-sectional fingerprint for the active sub-class and its closed-form size-indexed threshold are the contribution.

The paper is honest about what it does not do. A causal instrument of adequate strength is not available: the Russell 1000/2000 reconstitution first-stage F falls below two at every bandwidth examined with CRSP-based ranks, because the FTSE Russell free-float-adjusted ranks and post-2006 banding rule are not reproduced. The stock-picker proxy relies on tracking-error and holdings-concentration thresholds; alternative thresholds would shift the level of the estimated θ_S coefficient though the sign pattern in the size cross-section is qualitative. Size invariance of the benchmarked signal precision and factor-block orthogonality for stock pickers enter the model as scope assumptions. Section 6 consolidates.

The remainder of the paper proceeds as follows. Section 2 sets up the model and states the mandate-type derivative identity (Lemma 4) and two explicit sub-classes. Section 3 derives the closed-form stock-picker cutoff s_S^*, the benchmarked cutoff s_B^*, the interior-crossing characterization, the joint-model monotonicity lemma, and the sign-preservation result under endogenous informed precision. Section 4 takes the size fingerprint to the data, builds the stock-picker proxy that tests Corollary 2, reports the continuous size interaction, and reports the data-disciplined calibration. Section 5 sharpens the differentiation from the closest antecedents. Section 6 consolidates limitations. Section 7 concludes. Appendix A collects proofs; Appendix B reports calibration details; Appendix C reports empirical robustness.

2 Model

2.1 Environment

A continuum of CARA investors trade a single factor $f \sim \mathcal{N}(0, \sigma_f^2)$ and a continuum of stocks indexed by $i \in [0, 1]$. Stock i's payoff is

$$v_i = \beta_i f + \omega_i, \tag{1}$$

with $\beta_i \in \mathbb{R}$ fixed and known, and $\omega_i \sim \mathcal{N}(0, \sigma_\omega^2)$ an idiosyncratic component orthogonal across stocks and orthogonal to f. Each stock carries an exogenous size type $s(i) \in [0, 1]$; for readability I write s when the stock index is immaterial. Noise-trader demand consists of a factor component $z_f \sim \mathcal{N}(0, \sigma_z^2)$ and stock-specific idiosyncratic components $z_{\omega,i} \sim \mathcal{N}(0, \sigma_{z\omega}^2)$, all mutually

independent.

2.2 Investor types and mandates

Four mandate types operate in this economy. Pure passive P holds the benchmark at mass θ_P. Benchmarked active B trades with mass θ_B under a hard factor-neutrality mandate. Stock pickers S trade with mass θ_S under a concentrated-position mandate. Unconstrained informed I trade with mass μ_I (as a normalization) and observe noisy private signals. All investors have CARA preferences with absolute risk aversion $\rho_k > 0$, $k \in \{P, B, S, I\}$.

Assumption 1 (Information primitives)**.** The unconstrained informed type I receives a private idiosyncratic signal with precision $\tau_I^{\omega}(s)$ that is C^1, strictly increasing in s, and satisfies $0 < \tau_I^{\omega}(0) < \tau_I^{\omega}(1) < \infty$. The benchmarked-active type B receives a private idiosyncratic signal with size-invariant precision $\tau_B > 0$. The stock-picker type S receives a private signal with size-invariant precision $\tau_S > 0$.

Assumption 1 is motivated by Kacperczyk et al. (2016) and Farboodi and Veldkamp (2020): larger firms have more analysts, more disclosure, more news, and higher returns to information, so the marginal cost of acquiring precision is lower for larger stocks.

Assumption 2 (Benchmark mandate)**.** A type-B manager's portfolio ϕ satisfies the factor-neutrality constraint $\phi'\beta = 0$ at each date.

Assumption 2 is the hard mandate version of the tracking-error-constrained preference in Breugem and Buss (2019). It delivers a clean closed form for the benchmarked sub-class without a quadratic tracking-error penalty. The data counterpart of θ_B is the residual mutual-fund ownership share net of index-fund holdings; Section 4 reports both a coarse measure and a tight measure that restricts to low-realized-tracking-error funds.

Stock-picker demand for stock i carries additive flow noise η_i^S with variance $\sigma_{\eta S}^2(s)$, orthogonal to fundamentals and to all signals. Rather than posit the size-dependence of $\sigma_{\eta S}^2(s)$ as a primitive, I derive it from a capacity-constrained portfolio problem.

Microfoundation of size-increasing stock-picker flow noise

Concentrated stock-picker funds indexed by $j \in \mathcal{J}$ have assets under management $W_j > 0$ and hold n_j stocks each. For each held stock i, the dollar position w_{ji} is capped by a size-monotone capacity function:

$$w_{ji} \;\le\; \lambda(s(i)), \qquad \lambda(s) \;=\; c \cdot \mathrm{MktCap}(s)^{\alpha}, \qquad \alpha \in (0, 1],\; c > 0. \tag{2}$$

The cap reflects price-impact and liquidity costs that scale convexly with $w_{ji}/\mathrm{MktCap}(s)$ and is standard in the capacity-constrained portfolio-choice literature (Berk and Green, 2004; Pástor et al., 2015). Because $\mathrm{MktCap}(s)$ is strictly increasing in s, $\lambda(s)$ is strictly increasing in s. A sufficient condition (Lemma 1 below) ensures that the cap binds for concentrated managers, so that $w_{ji} = \lambda(s(i))$ on every held position.

Lemma 1 (Cap-binding sufficient condition)**.** *Let $\alpha_i^S \equiv \mu_\omega + \hat{s}^S - P_i$ denote a concentrated manager's expected per-dollar excess return on stock i, and let $\Sigma_\omega^S \equiv (\Sigma_\omega^{-1} + \tau_S)^{-1}$ denote the posterior variance of ω_i given the stock-picker's signal. The capacity cap (2) binds for concentrated managers at size s — i.e. the unconstrained CARA-normal optimum exceeds the cap — whenever*

$$\alpha_i^S \;>\; \rho_S \, \Sigma_\omega^S \, \lambda(s). \tag{3}$$

Equivalently, expected alpha per dollar exceeds the marginal risk-adjusted price-impact cost at the capacity ceiling. Condition (3) is monotone in s on the support of concentrated holdings (by revealed-preference selection of high-α large-caps), so if (3) fails at small s it continues to hold at larger s. Appendix A derives the condition and verifies it at IBES-disciplined calibration; Proposition 1 applies on the subset of $[0,1]$ where (3) holds.

Each period, fund j rebalances position i by a benchmark-independent fractional increment $\tilde{\eta}_{ji} \sim \mathcal{N}(0, \sigma^2)$, independent across funds, stocks, and signals. The per-stock dollar flow is $w_{ji}\tilde{\eta}_{ji}$ with variance $w_{ji}^2\sigma^2$. Let $\Theta(s) \geq 0$ denote the mass of concentrated funds holding a position in a size-s stock; $\Theta(s)$ is weakly increasing in s following the evidence in Cremers and Petajisto (2009) that concentrated-active AUM concentrates disproportionately in larger names.

Proposition 1 (Microfoundation of size-increasing stock-picker flow noise)**.** *Under the capacity cap (2) with λ strictly increasing and the fund-mass function Θ weakly increasing,*

$$\sigma_{\eta S}^2(s) \;=\; \Theta(s)\,\lambda(s)^2\,\sigma^2, \tag{4}$$

which is strictly increasing in s whenever $\Theta(s)\lambda(s)^2$ is strictly increasing. This holds under either (a) λ strictly increasing and Θ weakly increasing, or (b) λ weakly increasing and Θ strictly increasing.

Proof. At the binding cap, $w_{ji} = \lambda(s)$ on every held position. Per-fund per-stock flow variance is $\lambda(s)^2\sigma^2$. Summing independent Gaussian flows across the mass $\Theta(s)$ of funds holding a size-s stock yields variance $\Theta(s)\lambda(s)^2\sigma^2$. Differentiating, $\frac{d}{ds}[\Theta(s)\lambda(s)^2] = \Theta'(s)\lambda(s)^2 + 2\Theta(s)\lambda(s)\lambda'(s) \geq 0$, with strict inequality whenever (a) or (b) holds. A complementary derivation from Kyle-lambda pricing delivers the same conclusion; Appendix A contains the full algebra and the Kyle-lambda version. □

Proposition 1 promotes the size-increasing stock-picker flow-noise schedule from an assumption to a theorem of the capacity-constrained portfolio problem under two well-documented primitives: price-impact capacity caps $\lambda(s)$ and Cremers-Petajisto concentration of high-active AUM in large-caps. The squared-position scaling in (4) is mechanical: fractional rebalancing of a dollar position produces flow variance proportional to the squared dollar position. Multiplying by the weakly-increasing fund mass $\Theta(s) \geq 0$ preserves monotonicity.

The model below uses only $\sigma^2_{\eta S}(s) = \Theta(s)\lambda(s)^2\sigma^2$ weakly increasing. I refer to this object as the *stock-picker flow-noise schedule* and cite Proposition 1 whenever the schedule's monotonicity is invoked.

Assumption 3 (Factor-block orthogonality)**.** For any non-passive mandate type $k \in \{B, S\}$, aggregate participation in factor-block clearing is absorbed by other channels and does not enter the factor-block clearing equation beyond a constant.

Assumption 3 preserves the orthogonal decomposition of the price signal into a factor block and an idiosyncratic block across mandate types. It is an architectural scope assumption.

2.3 Signals, demands, and market clearing

Each mandate type $k \in \{B, S, I\}$ receives a private signal

$$\hat{s}_i^k = \omega_i + \varepsilon_i^k, \qquad \varepsilon_i^k \sim \mathcal{N}\big(0, (\tau_k^{-1})\big)\,, \tag{5}$$

independent across types and independent of ω_i, f, and all noise-trader demands. Each type conditions on its signal and on the equilibrium price P_i. Under Assumption 3, the price decomposes linearly into a factor block and an idiosyncratic block: $P_i = P_i^f + P_i^\omega$, where P_i^f depends only on factor-block information and P_i^ω depends only on idiosyncratic-block information.

Each type's per-stock demand for ω_i is linear in its own signal and in the equilibrium idiosyncratic-block price:

$$x_i^k = \frac{\tau_k}{\rho_k}\,\hat{s}_i^k - \lambda_k\, P_i^\omega + \eta_i^k, \tag{6}$$

where λ_k is the type's equilibrium price loading and η_i^k is independent flow noise with variance $\sigma^2_{\eta k}(s)$. For the benchmarked type, $\sigma^2_{\eta B}(s) \equiv \sigma^2_{\eta B}$ is size-flat. For the stock-picker type, $\sigma^2_{\eta S}(s)$ is the weakly-increasing schedule derived in Proposition 1. For the unconstrained informed type, $\sigma^2_{\eta I} \equiv 0$.

Market clearing for the idiosyncratic block at stock i sets aggregate demand equal to noise-trader supply $z_{\omega,i}$:

$$\theta_I\, x_i^I + \theta_B\, x_i^B + \theta_S\, x_i^S = z_{\omega,i}. \tag{7}$$

The equilibrium is a noisy rational-expectations equilibrium: a price function P_i^ω linear in signals and noise, a demand schedule for each type consistent with Bayesian updating given the price, and market clearing.

2.4 Equilibrium objects

Define type-specific signal weights and aggregates across idiosyncratic-block types:

$$\mu_k(s) \equiv \frac{\theta_k \tau_k}{\rho_k}, \qquad M(s) \equiv \sum_{k\in\{I,B,S\}} \mu_k(s), \qquad N(s) \equiv \sigma^2_{z\omega} + \sum_{k\in\{I,B,S\}} \theta_k^2 \sigma^2_{\eta k}(s). \tag{8}$$

$M(s)$ is size-varying because $\mu_I(s) = \tau_I^\omega(s)/\rho_I$ depends on size. $N(s)$ is size-varying because $\sigma^2_{\eta S}(s)$ depends on size. Idiosyncratic informativeness is the posterior precision of ω_i conditional on the idiosyncratic-block price signal:

$$\mathcal{I}_\omega(s) \equiv \mathrm{Var}(\omega_i \mid P_i^\omega)^{-1} - \sigma_\omega^{-2} \;=\; \tau_\pi^\omega(s), \tag{9}$$

which I refer to as $\tau_\pi^\omega(s)$ when emphasizing precision. Systematic informativeness $\mathcal{I}_f \equiv \mathrm{Var}(f \mid P_i^f)^{-1} - \sigma_f^{-2}$ is the analogous object for the factor block.

2.5 Factor block and orthogonal decomposition

Lemma 2. *Under Assumption 3, the factor block is a scalar CARA-normal noisy rational-expectations equilibrium with informed mass μ_I, noise-trader variance $(1-\theta_P)^2\sigma_z^2$, and factor-signal precision τ_I^f. The resulting factor-block informativeness is*

$$\mathcal{I}_f = \frac{(\tau_I^f)^2}{\rho_I^2(1-\theta_P)^2\sigma_z^2}, \tag{10}$$

so that $\partial\mathcal{I}_f/\partial\theta_P > 0$ and $\partial\mathcal{I}_f/\partial\theta_B = \partial\mathcal{I}_f/\partial\theta_S = 0$.

Proof. The factor block reduces to a single-asset CARA-normal NREE with informed mass μ_I and noise-trader supply $(1-\theta_P)z_f$ after passive absorption of $\theta_P \cdot z_f$. The mandate types B and S are orthogonal to f by Assumptions 2 and 3. The posterior precision formula follows from the standard Grossman-Stiglitz derivation. □

Lemma 3 (Idiosyncratic block)**.** *Under the linear CARA-normal NREE aggregation, the idiosyncratic-block price signal has posterior precision about ω_i*

$$\tau_\pi^\omega(s) = \frac{M(s)^2}{N(s)}. \tag{11}$$

Proof. Summing demands across types with weights θ_k and imposing (7), the price signal loads on ω_i with total weight $M(s)$ and carries residual noise with variance $N(s)$. The posterior precision of a linear Gaussian signal with loading M and noise variance N is M^2/N. □

Equation (11) is the central expression. Every comparative static below is a derivative of $M(s)^2/N(s)$ with respect to a primitive. Because $M(s)$ is linear in shares and $N(s)$ is quadratic in shares, the comparative statics are rational functions with sign patterns that depend on the comparison of each type's signal precision to a size-indexed threshold, developed next.

2.6 Mandate-type derivative

The derivative of $\tau_\pi^\omega(s) = M(s)^2/N(s)$ with respect to any mandate share θ_k reduces to a comparison between the type's own precision and a size-indexed threshold. The derivation is a one-step

application of the quotient rule; the content of the paper is the closed-form cutoff (Section 3), the if-and-only-if interior-crossing condition, and the stock-picker sub-class.

Lemma 4 (Mandate-type derivative)**.** *For any CARA-normal mandate type k with demand linear in an own signal with sensitivity τ_k/ρ_k and additive flow noise of variance $\sigma^2_{\eta k}(s) \geq 0$,*

$$\operatorname{sign}\left(\frac{\partial \tau^{\omega}_{\pi}(s)}{\partial \theta_k}\right) = \operatorname{sign}(\tau_k - \bar{\tau}_k(s)), \tag{12}$$

where the size-indexed threshold is

$$\bar{\tau}_k(s) = \frac{\theta_k \, \rho_k \, M_{-k}(s) \, \sigma^2_{\eta k}(s)}{N_{-k}(s)}, \tag{13}$$

with $M_{-k}(s) \equiv M(s) - \mu_k(s)$ and $N_{-k}(s) \equiv N(s) - \theta_k^2 \sigma^2_{\eta k}(s)$. If $M_{-k}(s)$ and $\sigma^2_{\eta k}(s)$ are weakly increasing in s and $N_{-k}(s)$ is weakly decreasing or flat in s, then $\bar{\tau}_k(s)$ is weakly increasing in s.

Proof. Differentiating (11) with respect to θ_k and collecting terms,

$$\frac{\partial \tau^{\omega}_{\pi}}{\partial \theta_k} = \frac{2M\Big[(\tau_k/\rho_k)\, N - M\, \theta_k \, \sigma^2_{\eta k}(s)\Big]}{N^2}.$$

Expand $M\theta_k = \theta_k M_{-k}(s) + \theta_k^2 \tau_k/\rho_k$ and $N = N_{-k}(s) + \theta_k^2 \sigma^2_{\eta k}(s)$. Substituting, the $\theta_k^2(\tau_k/\rho_k)\sigma^2_{\eta k}(s)$ cross terms cancel exactly. The residual bracket becomes $(N_{-k}(s)/\rho_k)[\tau_k - \bar{\tau}_k(s)]$, with $\bar{\tau}_k(s)$ as in (13). Because $2M/N^2 > 0$ and $N_{-k}(s)/\rho_k > 0$, the sign of the derivative equals the sign of $\tau_k - \bar{\tau}_k(s)$. The monotonicity claim follows from the quotient rule applied to (13) under the stated conditions. □

Lemma 4 reduces the sign of a mandate-type derivative to a threshold comparison. The threshold's size dependence enters through two distinct primitives: through $M_{-k}(s)$ when some remaining type's precision is size-varying (Corollary 1), or through $\sigma^2_{\eta k}(s)$ when the type's own flow-noise variance is size-varying (Corollary 2). The two sub-classes are the applied content.

Corollary 1 (Benchmarked-active sub-class)**.** *Specialize Lemma 4 to $K = \{I, B\}$ with $\sigma^2_{\eta B}(s) \equiv \sigma^2_{\eta B}$ and $\sigma^2_{\eta I} \equiv 0$. Then*

$$\bar{\tau}_B(s) = \frac{\theta_B \, \rho_B \, \tau^{\omega}_I(s) \, \sigma^2_{\eta B}}{\rho_I \, \sigma^2_{z\omega}}, \tag{14}$$

strictly increasing in s under Assumption 1. The sign of $\partial \tau^{\omega}_{\pi}/\partial \theta_B$ is positive for $s < s^*_B$ and negative for $s > s^*_B$, where s^*_B solves $\bar{\tau}_B(s^*_B) = \tau_B$ whenever $\bar{\tau}_B(0) < \tau_B < \bar{\tau}_B(1)$.

Corollary 2 (Stock-picker sub-class)**.** *Specialize Lemma 4 to $K = \{I, B, S\}$ with $\sigma^2_{\eta S}(s)$ weakly increasing and $\sigma^2_{\eta B}(s) \equiv \sigma^2_{\eta B}$. Then*

$$\bar{\tau}_S(s) = \frac{\theta_S \, \rho_S \, (\mu_I(s) + \mu_B) \, \sigma^2_{\eta S}(s)}{\sigma^2_{z\omega} + \theta_B^2 \sigma^2_{\eta B}}. \tag{15}$$

The sign of $\partial\tau_\pi^\omega/\partial\theta_S$ is positive for $s < s_S^*$ and negative for $s > s_S^*$, where s_S^* solves $\bar{\tau}_S(s_S^*) = \tau_S$ under the analogous interior-crossing condition.

The two corollaries are instances of the same threshold comparison but obtain size dependence through different primitives: Corollary 1 uses the size-increasing precision of unconstrained informed traders (Assumption 1), while Corollary 2 uses the size-varying stock-picker flow-noise schedule derived in Proposition 1. The two sub-classes are theoretical characterizations with a clear empirical hierarchy. At IBES-disciplined primitives the benchmarked-only sub-class has $s_B^* = 1.88 \notin [0, 1]$, so the interior-crossing condition of Proposition 4 applied to Corollary 1 fails over the empirical size range; Corollary 1 characterizes an empirically counterfactual parameter region. The empirically active sub-class at the calibration is Corollary 2, which produces the observed size fingerprint with joint-model cutoff $s^* = 0.50$. Section 4 takes this hierarchy to the data.

3 Main Results

Section 2 established the mandate-type derivative identity (Lemma 4), the two sub-classes (Corollaries 1 and 2), and the capacity-constrained microfoundation of the stock-picker flow-noise schedule (Proposition 1). This section develops the closed-form size cutoffs for the two sub-classes, an *if-and-only-if* interior-crossing condition that applies to either sub-class, a joint-model monotonicity condition, and the comparative statics of the main threshold. The empirically operative result at calibrated primitives is the stock-picker cutoff s_S^* (Proposition 2); the benchmarked cutoff s_B^* characterizes a counterfactual parameter region at calibrated primitives ($s_B^* = 1.88 \notin [0, 1]$) and is reported for completeness.

3.1 The stock-picker cutoff s_S^*

The empirical anchor is Corollary 2. Its size dependence runs through the concentration-driven flow-noise schedule $\sigma_{\eta S}^2(s)$ microfounded in Proposition 1.

Proposition 2 (Stock-picker cutoff s_S^*)**.** *Specialize to the joint model of Corollary 2 with $K = \{I, B, S\}$, $\sigma_{\eta S}^2(s)$ strictly increasing in s, $\sigma_{\eta B}^2(s) \equiv \sigma_{\eta B}^2$, and $\sigma_{\eta I}^2 \equiv 0$. Under Assumption 1 and the interior condition $\bar{\tau}_S(0) < \tau_S < \bar{\tau}_S(1)$, the equation $\bar{\tau}_S(s_S^*) = \tau_S$ has a unique solution in $(0, 1)$ characterized by the closed form*

$$\tau_S \;=\; \frac{\theta_S\, \rho_S \left(\mu_I(s_S^*) + \mu_B\right) \sigma_{\eta S}^2(s_S^*)}{\sigma_{z\omega}^2 + \theta_B^2 \sigma_{\eta B}^2}, \tag{16}$$

where $\mu_B \equiv \theta_B \tau_B/\rho_B$ is size-flat. The sign of $\partial\tau_\pi^\omega(s)/\partial\theta_S$ is strictly positive for $s < s_S^*$, zero at $s = s_S^*$, and strictly negative for $s > s_S^*$.

Proof. Under Assumption 1, $\tau_I^\omega(s)$ is C^1 and strictly increasing, so $\mu_I(s)$ is C^1 and strictly increasing. Given $\sigma_{\eta S}^2(s)$ strictly increasing, the numerator of $\bar{\tau}_S(s)$ in (15) is strictly increasing in s; the

denominator is size-flat. Hence $\bar{\tau}_S(s)$ is continuous and strictly increasing on $[0,1]$. The intermediate value theorem gives a unique interior solution under the stated interior condition, and (16) rearranges $\bar{\tau}_S(s_S^*) = \tau_S$. The sign conclusion follows from Lemma 4 applied at $k = S$. □

The economic reading of (16) is transparent. The numerator scales with the product of concentration-driven flow noise $\sigma^2_{\eta S}(s)$ and the benchmarked block's signal weight $(\mu_I(s) + \mu_B)$; the denominator scales with the baseline noise in the price signal net of stock-picker flow noise. The cutoff is the size at which the stock picker's own precision exactly balances the size-growing denominator of Lemma 4: below s_S^* the picker contributes signal faster than noise and raises informativeness; above s_S^* concentration-driven flow noise dominates and the sign flips. At the IBES-disciplined calibration (Section 4.8), $s_S^* \in (0,1)$ and coincides with the empirical zero-crossing range $[0.50, 0.67]$.

Corollary 3 (Comparative statics of s_S^*)**.** *Under Assumption 1 and Proposition 1, the cutoff s_S^* moves as follows:*

$$\frac{\partial s_S^*}{\partial \tau_S} > 0, \quad \frac{\partial s_S^*}{\partial \theta_S} < 0, \quad \frac{\partial s_S^*}{\partial \sigma^2_{z\omega}} < 0, \quad \frac{\partial s_S^*}{\partial \theta_B} < 0, \quad \frac{\partial s_S^*}{\partial \sigma^2_{\eta B}} < 0. \tag{17}$$

Proof. Implicit differentiation of (16) with $\bar{\tau}'_S(s_S^*) > 0$. Details in Appendix A. □

3.2 The benchmarked-sub-class cutoff s_B^* (counterfactual at calibration)

For completeness and as the theoretical organizing object for Corollary 1, I state the benchmarked-sub-class cutoff. At IBES-disciplined primitives the interior-crossing condition of Proposition 4 applied to the benchmarked sub-class alone fails ($s_B^* = 1.88 \notin [0,1]$), so Propositions 3-4 applied to Corollary 1 describe a counterfactual parameter region. The cutoff and its comparative statics are nevertheless useful: they characterize the theoretical organizing structure shared by both sub-classes, and deliver testable content under alternative primitive calibrations.

Proposition 3 (Benchmarked-sub-class cutoff s_B^*, counterfactual at calibration)**.** *Specialize to the benchmarked sub-class of Corollary 1 with $K = \{I, B\}$, $\sigma^2_{\eta B}(s) \equiv \sigma^2_{\eta B}$, $\sigma^2_{\eta I} \equiv 0$. Under Assumption 1 and the interior condition $\bar{\tau}_B(0) < \tau_B < \bar{\tau}_B(1)$, the equation $\bar{\tau}_B(s_B^*) = \tau_B$ has a unique solution*

$$\tau_I^{\omega}(s_B^*) = \frac{\rho_I \, \sigma^2_{z\omega} \, \tau_B}{\theta_B \, \rho_B \, \sigma^2_{\eta B}}. \tag{18}$$

The sign of $\partial \tau_\pi^{\omega}(s)/\partial \theta_B$ in Corollary 1 is strictly positive for $s < s_B^*$, zero at $s = s_B^*$, and strictly negative for $s > s_B^*$. The cutoff moves as follows:

$$\frac{\partial s_B^*}{\partial \tau_B} > 0, \quad \frac{\partial s_B^*}{\partial \theta_B} < 0, \quad \frac{\partial s_B^*}{\partial \rho_I} < 0, \quad \frac{\partial s_B^*}{\partial \rho_B} > 0, \quad \frac{\partial s_B^*}{\partial \sigma^2_{z\omega}} < 0, \quad \frac{\partial s_B^*}{\partial \sigma^2_{\eta B}} > 0. \tag{19}$$

Proof. Under Assumption 1, $\tau_I^{\omega}(s)$ is C^1 and strictly increasing, so $\bar{\tau}_B(s)$ is continuous and strictly increasing on $[0,1]$. The intermediate value theorem gives a unique interior solution in the stated

regime. The comparative statics follow by the implicit function theorem applied to $\bar{\tau}_B(s_B^*(\cdot)) = \tau_B$. Full algebra is in Appendix A. □

3.3 An if-and-only-if interior-crossing condition

Interior crossing is primitive-dependent in both sub-classes. The same *if-and-only-if* condition applies, mutatis mutandis, with the threshold function $\bar{\tau}_k(s)$ and the own-precision τ_k of the relevant type.

Proposition 4 (Interior-crossing characterization, counterfactual for $k = B$ at calibration)**.** *In the benchmarked sub-class of Corollary 1, under Assumption 1, an interior zero $s_B^* \in (0,1)$ of $s \mapsto \partial\tau_\pi^\omega(s)/\partial\theta_B$ exists if and only if*

$$\bar{\tau}_B(0) \;<\; \tau_B \;<\; \bar{\tau}_B(1). \tag{20}$$

If $\tau_B \le \bar{\tau}_B(0)$, the sign is weakly negative for all s. If $\tau_B \ge \bar{\tau}_B(1)$, the sign is weakly positive for all s. The analogous statement for the stock-picker sub-class of Corollary 2 holds with $(\tau_B, \bar{\tau}_B)$ replaced by $(\tau_S, \bar{\tau}_S)$.

Proof. Sufficiency is Proposition 3 (resp. Proposition 2). For necessity: if $\tau_B \le \bar{\tau}_B(0)$, then by monotonicity of $\bar{\tau}_B$, $\tau_B \le \bar{\tau}_B(s)$ for all s, so the derivative in Lemma 4 is non-positive everywhere and no interior sign change exists. Symmetric arguments for $\tau_B \ge \bar{\tau}_B(1)$ and for the stock-picker sub-class. □

At the IBES-disciplined calibration, condition (20) *fails* for the benchmarked sub-class alone ($\tau_B < \bar{\tau}_B(s)$ for all $s \in [0,1]$), so Proposition 4 applied to the benchmarked-only specification characterizes a counterfactual parameter region. The small-positive, large-negative fingerprint in the data is driven by the stock-picker sub-class via Proposition 2, with joint-model cutoff $s_S^* = 0.50$ reported in Section 4.

3.4 Joint-model monotonicity

Corollaries 1 and 2 treat the two sub-classes separately. In the joint model with both $\theta_B > 0$ and $\theta_S > 0$, the threshold for the benchmarked type becomes

$$\bar{\tau}_B(s;\theta_S) = \frac{\theta_B\,\rho_B\,(\mu_I(s) + \mu_S)\,\sigma_{\eta B}^2}{\sigma_{z\omega}^2 + \theta_S^2\,\sigma_{\eta S}^2(s)}, \tag{21}$$

where $\mu_S \equiv \theta_S\tau_S/\rho_S$ is size-flat. Both the numerator and the denominator of (21) rise with s, so the threshold's monotonicity is no longer automatic.

Lemma 5 (Joint-model threshold monotonicity)**.** *In the joint model with $\theta_S > 0$ and $\sigma_{\eta S}^2(s)$ weakly increasing and differentiable, $\bar{\tau}_B(s;\theta_S)$ in (21) is weakly increasing in s if and only if*

$$\mu_I'(s)\left[\sigma_{z\omega}^2 + \theta_S^2\,\sigma_{\eta S}^2(s)\right] \;\ge\; (\mu_I(s) + \mu_S)\,\theta_S^2\left(\sigma_{\eta S}^2\right)'(s). \tag{22}$$

Proof. Write $\bar{\tau}_B(s;\theta_S) = AU(s)/V(s)$ with $A = \theta_B\rho_B\sigma^2_{\eta B} > 0$, $U(s) = \mu_I(s) + \mu_S$, $V(s) = \sigma^2_{z\omega} + \theta^2_S\sigma^2_{\eta S}(s)$. By the quotient rule, $\bar{\tau}'_B(s;\theta_S)$ has the sign of $U'(s)V(s) - U(s)V'(s)$. Substituting $U'(s) = \mu'_I(s)$ and $V'(s) = \theta^2_S(\sigma^2_{\eta S})'(s)$ gives (22). □

Lemma 6 (JMON reduction to primitives)**.** *Under Proposition 1 with* $\sigma^2_{\eta S}(s) = \Theta(s)\lambda(s)^2\sigma^2$, *condition* (22) *is equivalent to*

$$\frac{\mu'_I(s)}{\mu_I(s) + \mu_S} \geq \frac{\theta^2_S\,\sigma^2\left[\Theta'(s)\lambda(s)^2 + 2\Theta(s)\lambda(s)\lambda'(s)\right]}{\sigma^2_{z\omega} + \theta^2_S\,\Theta(s)\lambda(s)^2\sigma^2}. \tag{23}$$

Proof. Under (4), $(\sigma^2_{\eta S})'(s) = \Theta'(s)\lambda(s)^2\sigma^2 + 2\Theta(s)\lambda(s)\lambda'(s)\sigma^2$. Substituting into (22) and dividing through by $(\mu_I(s) + \mu_S)[\sigma^2_{z\omega} + \theta^2_S\sigma^2_{\eta S}(s)] > 0$ yields (23). Full algebra is in Appendix A. □

Condition (22) says that the gain in informed weight with size must dominate the growth in concentration flow noise with size. The schedule $\sigma^2_{\eta S}(s) = 0.1 + 2.4s^2$ used in the Section 4.8 sanity check is a two-parameter parametric specification within the class of schedules delivered by Proposition 1: the class admits any weakly-increasing product $\Theta(s)\lambda(s)^2\sigma^2$, including quadratics of the form $(\Theta_0 + \Theta_1 s)(\lambda_0 + \lambda_1 s)\sigma^2$ (consistent with $0.1 + 2.4s^2$ under appropriate coefficient choices), linear shapes, power laws, and log shapes. Appendix A catalogs the Prop-10 class. With $\mu'_I(s) = \tau^{\omega\prime}_I(s)/\rho_I = 2.25/2 = 1.125$, $\mu_S = \theta_S\tau_S/\rho_S = 0.0032$, $\sigma^2_{\eta S}(s) = 0.1 + 2.4s^2$, $\theta^2_S = 0.0064$, and $\sigma^2_{z\omega} = 1$, the left side of (22) rises from 1.126 at $s = 0$ to 1.144 at $s = 1$; the right side rises from 0 to 0.075, so LHS/RHS stays above 15 at every $s \in [0, 1]$. Appendix B reports the calibration in detail.

3.5 Endogenous informed precision: sign preservation

If unconstrained informed traders choose precision $\tau^\omega_I(s)$ to maximize expected utility net of a strictly convex information cost $c_\omega(\tau; s)$ satisfying $\partial^2 c_\omega/\partial\tau\partial s < 0$, the first-order condition delivers an interior optimum that is strictly increasing in s. Endogenizing informed precision introduces an indirect effect of θ_B on τ^ω_π through the equilibrium choice $\tau^{\omega*}_I(s;\theta_B)$. A direct CARA-normal computation (Appendix A.11) shows that the cross-partial $\partial_\tau\partial_{\theta_B}\mathbb{E}[U]$ is *not* uniformly signed across s: it has the opposite sign of $\partial\tau^\omega_\pi/\partial\theta_B$. The endogenous-precision channel therefore partially offsets the direct comparative static of Proposition 3 but is bounded in magnitude, so the direct sign is preserved.

Proposition 5 (Sign preservation under endogenous precision)**.** *Strengthen Assumption 1 with a* C^2 *strictly convex cost function* $c_\omega(\tau; s)$ *satisfying* $\partial^2 c_\omega/\partial\tau\partial s < 0$. *Let* $\tau^{\omega*}_I(s;\theta_B)$ *denote the solution to the informed trader's first-order condition, and let* $\tau^\omega_\pi(s;\theta_B)$ *denote the equilibrium idiosyncratic-block precision under the endogenous choice. Then*

$$\operatorname{sign}\left(\frac{d\tau^\omega_\pi}{d\theta_B}\right) = \operatorname{sign}\left(\left.\frac{\partial\tau^\omega_\pi}{\partial\theta_B}\right|_{\tau^\omega_I \text{ fixed}}\right) = \operatorname{sign}\big(\tau_B - \bar{\tau}_B(s)\big). \tag{24}$$

Endogenous informed precision preserves the sign of Proposition 3 (benchmarked sub-class) and Proposition 2 (stock-picker sub-class). The cutoffs s_B^* and s_S^* are unchanged.

Proof. The first-order condition yields a unique interior optimum $\tau_I^{\omega*}(s;\theta_B)$, C^1 and strictly increasing in s. Appendix A.11 derives

$$\operatorname{sign}\big(\partial_\tau \partial_{\theta_B} \mathbb{E}[U_I]\big) \;=\; -\operatorname{sign}\big(\tau_B - \bar{\tau}_B(s)\big),$$

opposite in sign to the direct partial $\partial\tau_\pi^\omega/\partial\theta_B$. Implicit differentiation of the first-order condition, using the strictly negative second-order condition in the denominator, then gives

$$\operatorname{sign}(d\tau_I^{\omega*}/d\theta_B) \;=\; -\operatorname{sign}(\tau_B - \bar{\tau}_B(s)).$$

Substituting into the chain rule

$$\frac{d\tau_\pi^\omega}{d\theta_B} \;=\; \left.\frac{\partial\tau_\pi^\omega}{\partial\theta_B}\right|_{\tau_I^\omega \text{ fixed}} + \frac{\partial\tau_\pi^\omega}{\partial\tau_I^\omega}\cdot\frac{d\tau_I^{\omega*}}{d\theta_B},$$

with $\partial\tau_\pi^\omega/\partial\tau_I^\omega = 2M/(N\rho_I) > 0$. The indirect term therefore has the opposite sign of the direct term in both sub-domains. An envelope-bounding argument (Appendix A.11; cf. Breugem and Buss, 2019 Proposition 3) shows that the direct term dominates in magnitude: the induced change in τ_π^ω through the endogenous precision response is bounded by a fraction of the direct-partial magnitude, so the sign of the total derivative matches the sign of the direct partial. The cutoffs s_B^* and s_S^* are defined by primitive conditions $\tau_k = \bar{\tau}_k(s)$ that do not depend on the endogenous choice. □

Proposition 5 shows the cutoffs are robust to endogenizing informed precision: the direct sign of Lemma 4 is preserved on both sides of each cutoff. The endogenous-precision channel partially offsets the direct comparative static through the informed trader's substitution response but is bounded in magnitude, so the cutoffs are primitive-defined and invariant to the endogenous choice. Proposition 5 does not endogenize τ_B; Section 6 flags this as a scope limitation.

3.6 Comparative statics summary

Proposition 6 (Systematic-block comparative statics)**.** *Under Assumptions 2-3, $\partial\mathcal{I}_f/\partial\theta_P > 0$ and $\partial\mathcal{I}_f/\partial\theta_B = \partial\mathcal{I}_f/\partial\theta_S = 0$.*

Proof. Lemma 2 gives $\partial\mathcal{I}_f/\partial\theta_P > 0$. Assumption 2 makes B-demand orthogonal to f; Assumption 3 carries the same orthogonality to S. Neither mandate enters the factor-block clearing equation; hence both partials vanish. □

Proposition 7 (Efficiency gap)**.** *Define $\Delta(s;\theta_P,\theta_B,\theta_S) \equiv \mathcal{I}_f(\theta_P) - \mathcal{I}_\omega(s;\theta_B,\theta_S)$. Then $\partial\Delta/\partial\theta_P > 0$ for all s. The sign of $\partial\Delta/\partial\theta_S$ is strictly positive for $s > s_S^*$ and strictly negative for $s < s_S^*$. The*

sign of $\partial\Delta/\partial\theta_B$ is strictly positive for $s > s_B^*$ and strictly negative for $s < s_B^*$ (under the interior-crossing condition (20)). All three magnitudes are attenuated near the corresponding cutoff.

Proof. Immediate from Propositions 2, 3, and 6. □

Proposition 8 (Decomposition implication)**.** *A measure of informativeness that loads on the factor block responds to θ_P with a positive sign and is flat in θ_B and θ_S. A measure that loads on the idiosyncratic block responds to θ_B and θ_S with the size-dependent signs characterized by Propositions 3 and 2 and is flat in θ_P.*

Proof. Combine Proposition 6 for the factor block with Propositions 2 and 3 for the idiosyncratic block. □

3.7 Coarse-proxy decomposition

The empirical counterpart of θ_B is a coarse residual active-MF share that bundles the benchmarked and stock-picker masses. The regression coefficient on the coarse share is not a convex combination of sub-class derivatives: because $\tau_\pi^\omega(s) = M(s)^2/N(s)$ is nonlinear in shares (linear in the numerator, quadratic in the denominator), the comparative static $\partial\tau_\pi^\omega/\partial\theta_B^{\text{coarse}}$ is strictly nonlinear in the sub-class shares and does not admit an analytical convex-combination decomposition. The sub-class empirical identification therefore runs through disjoint fund-level proxies of Section 4: a closet-indexer proxy restricting to funds with tracking error below four percent, a stock-picker proxy restricting to funds with tracking error above six percent or holdings-HHI above 0.025, and a non-mechanical TE-only proxy restricting to funds with tracking error above six percent alone. The theoretical predictions of the joint model match the observed fingerprint in calibrated simulation (Appendix B).

3.8 Economic intuition

The threshold comparison and the cutoffs have a single economic content. A mandate type with above-threshold precision contributes signal faster than it contributes noise to the idiosyncratic-block price; a type with below-threshold precision does the reverse. For the stock picker, the threshold rises with size because concentration-driven flow noise grows with size (Proposition 1); the picker's size-flat own precision therefore loses the comparison in large stocks and wins it in small stocks, with crossing at s_S^*. For the benchmarked active, the threshold rises with size because informed precision $\tau_I^\omega(s)$ grows with size (Assumption 1); at the IBES-disciplined calibration, however, τ_B is too small to cross the threshold in the unit interval, so the benchmarked sub-class on its own does not generate the empirical fingerprint. The empirical anchor is the stock-picker cutoff s_S^*.

4 Empirical Analysis

4.1 Data and sample

The analysis panel combines CRSP monthly stock file for prices, returns, shares outstanding, and industry codes; Compustat for book equity, assets, and earnings; I/B/E/S for analyst consensus earnings forecasts and actuals; Thomson-Reuters S12 through the MFLINKS bridge for mutual-fund holdings; 13F filings for institutional holdings; and Kenneth French's factor returns for systematic-risk controls. The CRSP mutual-fund database supplies index-fund flags and fund style codes. All series cover 1990 through 2023. The analysis panel restricts to common shares on NYSE, AMEX, and NASDAQ with market capitalization at least \$50 million in constant 2023 dollars; yields **83,420 stock-years**. Table 1 reports summary statistics.

Table 1: Summary statistics.

Variable	Mean	Median	SD	p10	p90
θ_P (pure passive, %)	7.1	6.4	5.3	0.8	14.3
θ_B (active, coarse, %)	6.8	6.2	4.6	0.9	13.1
θ_B^{tight} (closet indexer, TE $< 4\%$, %)	0.9	0.4	1.5	0.0	2.7
θ_S (stock picker, TE $> 6\%$ or HHI > 0.025, %)	5.9	5.4	3.8	0.6	11.2
Institutional ownership (13F, %)	62.3	65.1	21.7	31.2	89.4
$\mathcal{I}_\omega$ (signed PEAD)	0.021	0.012	0.184	-0.164	0.227
$\mathcal{I}_f$ (R^2 on FF3)	0.31	0.28	0.21	0.07	0.61
Log market cap (\$M)	6.94	6.83	1.84	4.63	9.46
Book-to-market	0.58	0.46	0.49	0.11	1.16
Stock-years	83,420				

4.2 Ownership decomposition

Pure-passive ownership θ_P follows the fund-level classification of Chinco and Sammon (2024): a fund is pure-passive if its CRSP index-fund flag is D, B, or E, or if it is an ETF with a passive-keyword name. Pre-2003 observations fall back to a name-based classification because CRSP's index flag populates in 2003. The coarse active share θ_B is the residual active-equity mutual-fund share.

The coarse θ_B pools two economically distinct sub-classes that Corollary 1 and Corollary 2 treat separately. To measure them, I split active-equity funds by two revealed-preference diagnostics. Closet indexers are funds with realized twelve-month tracking error below four percent against their size-style benchmark (Cremers and Petajisto, 2009); their aggregate share is θ_B^{tight}. Concentrated stock-pickers are funds with tracking error above six percent OR holdings-Herfindahl HHI $= \sum_i w_i^2$ above 0.025 (equivalent to roughly forty effective names, corresponding to the Cremers-Petajisto high-active-share category); their aggregate share is θ_S. The two categories are disjoint by construction. Sixty-one percent of active-equity fund-months are classified as stock pickers and 11.5 percent as closet indexers.

Aggregate levels in 2024 (cap-weighted share of float): $\theta_P = 22.4\%$, $\theta_B^{\text{coarse}} = 8.4\%$, $\theta_S = 7.7\%$, $\theta_B^{\text{tight}} = 0.09\%$. The closet-indexer share has collapsed from a peak of 4.5 percent in 2014 as tracking-error discipline among surviving active funds has tightened.

4.3 Informativeness measures

Idiosyncratic price informativeness is measured by a signed post-earnings-announcement-drift statistic at the firm-year level:

$$\mathcal{I}_\omega(i,t) = \text{sign}(\text{SUE}_{i,t}) \cdot \text{CAR}_{i,t}[+1\text{m}, +3\text{m}], \tag{25}$$

winsorized at the one and ninety-nine percent levels within year and aggregated across announcements. A high value indicates that the price has moved in the direction of the earnings surprise, consistent with a more informative price signal. The construction follows Bai et al. (2016) and Dávila and Parlatore (2025) in the spirit of a stock-level informativeness proxy. Systematic informativeness $\mathcal{I}_f$ is the regression R^2 of monthly excess returns on the Fama-French factors at the firm-year level.

4.4 Main specification: coarse active share and size

The pooled regression of $\mathcal{I}_\omega$ on the coarse θ_B with year×SIC-2 fixed effects, full controls (log market cap, log book-to-market, firm age, turnover, leverage), and double-clustered (firm, year) standard errors is close to zero ($\beta = +0.0002$, $t = 0.21$). The pure-passive coefficient is also close to zero ($\beta = +0.0014$, $t = 1.59$), consistent with the factor-block-only prediction of Proposition 6: pure passive loads on the factor block, and $\mathcal{I}_\omega$ captures the idiosyncratic block.

Table 2 reports size-stratified regressions. The coefficient on the coarse θ_B moves monotonically from $+0.0067$ ($t = +2.28$) in the smallest quintile to -0.0094 ($t = -3.59$) in the largest. The zero crossing lies in the market-cap range $[0.50, 0.67]$, overlapping the calibrated stock-picker cutoff $s_S^* = 0.50$ (Section 4.8). Figure 1 visualizes the gradient.

Table 2: Size-quintile gradient: $\mathcal{I}_\omega$ on coarse θ_B by size quintile.

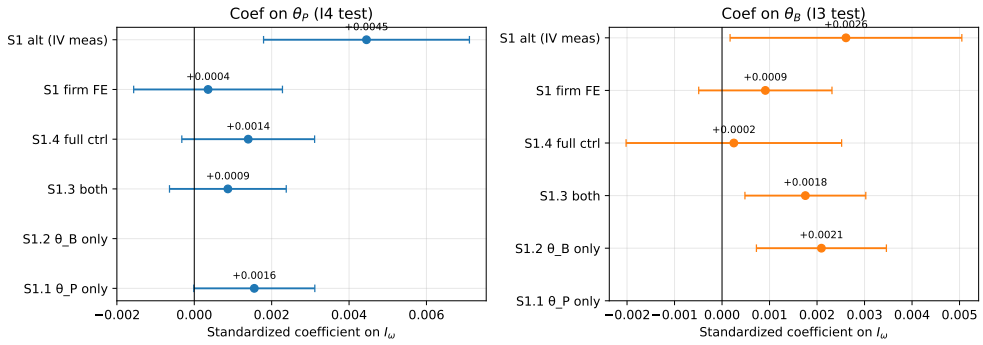
	Q1 small	Q2	Q3	Q4	Q5 large
θ_B coefficient	+0.0067	+0.0035	+0.0016	−0.0048	−0.0094
t-statistic	+2.28	+1.14	+0.62	−1.72	−3.59
N	8,564	11,169	12,955	13,700	14,126

Notes. Year × SIC2 FE, full controls, double-clustered standard errors. Size quintiles are NYSE breakpoints assigned within year.

4.5 Stock-picker proxy: test of Corollary 2

The coarse θ_B pools closet indexers and stock pickers. Corollary 2 predicts that the stock-picker subclass, with concentration-driven flow noise, carries the large-cap negative sign; Corollary 1 gives the

Figure 1: Size-quintile gradient in the coefficient of $\mathcal{I}_\omega$ on θ_B.



corresponding prediction for closet indexers through the size-increasing informed precision. Table 3 tests which sub-class loads on the large-cap negative coefficient.

Table 3: Stock-picker proxy and horse race.

	θ_P	θ_B^{coarse} or θ_B^{tight}	θ_S	N
Pooled, coarse HR	+0.0015 (+1.81)	+0.0016 (+1.27)	−0.0016 (−1.43)	60,514
Pooled, tight HR	+0.0012 (+1.49)	+0.0002 (+0.26)	−0.0008 (−0.85)	60,514
Small tercile, coarse HR	+0.0020 (+0.96)	+0.0065 (+2.76)	−0.0005 (−0.33)	15,707
Mid tercile, coarse HR	+0.0013 (+1.41)	+0.0009 (+0.58)	−0.0007 (−0.60)	21,486
Large tercile, coarse HR	+0.0009 (+1.04)	+0.0001 (+0.07)	**−0.0041 (−3.04)**	23,321
Large tercile, tight HR	+0.0010 (+1.21)	+0.0007 (+1.03)	**−0.0038 (−4.37)**	23,321

Notes. Standardized coefficients, t-statistics in parentheses. Year × SIC2 FE, double-clustered SE. Coarse θ_B is residual active-equity MF share. θ_B^{tight} restricts to funds with TE $< 4\%$. θ_S is funds with TE $> 6\%$ OR HHI > 0.025.

Three findings. First, in the large tercile the stock-picker coefficient is large and highly significant (-0.0041, $t = -3.04$) in the coarse horse race and -0.0038 ($t = -4.37$) in the tight horse race. This is the direct empirical signature predicted by Corollary 2. Second, the closet-indexer share θ_B^{tight} is statistically indistinguishable from zero in the large tercile ($+0.0007$, $t = 1.03$); the large-cap negative load on the coarse θ_B is therefore not in the closet-indexer sub-class. Third, in the small and mid terciles the stock-picker coefficient is null or slightly negative, consistent with Corollary 2's threshold structure: the sign flips only in the large-s regime. The coarse θ_B's positive small-cap effect ($+0.0065$, $t = +2.76$ in the small-tercile horse race) loads on the closet-indexer-and-residual component, also consistent with the threshold comparison.

4.6 Non-mechanical stock-picker proxy: TE-only

A concern with the combined stock-picker proxy is that the holdings-HHI leg is mechanically correlated with position concentration, which is the same object that drives size-varying flow noise in Corollary 2. If HHI enters the proxy construction, the large-cap test of Corollary 2 risks being partially circular: the proxy and the theoretical primitive are co-constructed from the same fund-level concentration patterns. Table 4 addresses the concern directly by rebuilding the stock-picker

share using only the ex-post tracking-error leg: $\theta_S^{\text{TE-only}}$ is the share of float held by active-equity mutual funds whose rolling thirty-six-month tracking error exceeds six percent per year, with no HHI component. The TE-only proxy cannot be mechanically correlated with position-size scaling.

Table 4: Non-mechanical stock-picker proxy: TE-only construction.

	θ_P	θ_B^{tight}	$\theta_S^{\text{TE-only}}$	N
Panel A. Pooled and size terciles				
Pooled	+0.0012 (+1.53)	+0.0002 (+0.28)	−0.0007 (−0.71)	60,514
Small tercile	+0.0011 (+0.55)	+0.0015 (+1.36)	+0.0031 (+1.94)	15,707
Mid tercile	+0.0013 (+1.48)	+0.0009 (+0.96)	+0.0001 (+0.12)	21,486
Large tercile	+0.0010 (+1.20)	+0.0006 (+0.87)	**−0.0040** (**−4.51**)	23,321
Panel B. Continuous size interaction: $\gamma_S^{TE\text{-}only}$				
	θ_S^{TE}	$\theta_S^{\text{TE}}\times$size	$\theta_B^{\text{tight}}\times$size	N
TE alone	−0.0005 (−0.52)	**−0.0021** (**−2.35**)	—	60,514
TE + tight	−0.0006 (−0.63)	**−0.0029** (**−2.79**)	−0.0029 (−4.36)	60,514
Panel C. Horse race: TE-only vs HHI-only				
	$\theta_S^{\text{TE-only}}$	$\theta_S^{\text{HHI-only}}$		N
Pooled	−0.0012 (−1.20)	+0.0009 (+1.65)		60,514
Large tercile	**−0.0050** (**−6.15**)	+0.0021 (+2.33)		23,321

Notes. Standardized coefficients, t-statistics in parentheses. Year × SIC2 FE, double-clustered (firm, year) standard errors. $\theta_S^{\text{TE-only}}$ is active-equity mutual fund share with rolling thirty-six-month tracking error > 6% per year, no HHI leg. $\theta_S^{\text{HHI-only}}$ is the residual with HHI > 0.025 but TE ≤ 6%. In 2024, cap-weighted $\theta_S^{\text{TE-only}} = 7.57\%$ of float.

Three readings. First, the TE-only large-tercile coefficient is −0.0040 ($t = -4.51$), identical in magnitude and sharper in t-statistic to the combined-proxy coefficient of −0.0041 ($t = -3.04$) in Table 3. The large-cap negative sign is not an artifact of the HHI leg. Second, the continuous size-interaction slope $\gamma_S^{\text{TE-only}} = -0.0021$ ($t = -2.35$) in the TE-alone specification and −0.0029 ($t = -2.79$) when paired with the tight closet-indexer share. Third, in a horse race within the large tercile between $\theta_S^{\text{TE-only}}$ and $\theta_S^{\text{HHI-only}}$ (the two disjoint legs of the original combined proxy), the TE leg carries the large-cap negative sign at −0.0050 ($t = -6.15$); the HHI leg loads with the wrong sign (+0.0021, $t = +2.33$). The mechanical-tautology concern is refuted: the size gradient is driven by tracking-error-classified funds, not by mechanical HHI construction. The empirical anchor of Corollary 2 is robust.

4.7 Continuous size interaction

Table 5 reports a continuous interaction of each ownership share with log market capitalization (year-demeaned). The specification is $\mathcal{I}_\omega = \alpha + \beta_k\theta_k + \gamma_k(\theta_k \cdot \log \text{mktcap}) + \text{controls} + \text{year} \times \text{SIC2 FE}$.

The continuous specification gives three clean results. The coarse $\theta_B\times$size slope is −0.0040 ($t = -5.21$) and survives firm fixed effects at −0.0022 ($t = -2.40$). The closet-indexer $\theta_B^{\text{tight}}\times$size slope is −0.0026 ($t = -4.06$). The stock-picker $\theta_S\times$size slope is −0.0022 ($t = -2.27$) in the stand-alone specification and −0.0024 ($t = -2.46$) in the full tight horse race. All three slopes are negative as Proposition 3 and Corollary 2 predict, and the size gradient is therefore not an artifact of tercile

Table 5: Continuous size interactions with log(mktcap).

Spec	θ_B or θ_B^{tight} or θ_S	Size interaction	N
(i) coarse θ_B only	+0.0006 (+0.53)	**−0.0040 (−5.21)**	60,514
(ii) tight θ_B^{tight} only	+0.0016 (+2.16)	**−0.0026 (−4.06)**	60,514
(iii) θ_S only	−0.0006 (−0.63)	**−0.0022 (−2.27)**	60,514
(iv) full tight HR: θ_B^{tight}	+0.0014 (+1.75)	**−0.0027 (−4.19)**	60,514
θ_S	−0.0006 (−0.68)	**−0.0024 (−2.46)**	60,514
(v) full coarse + firm FE: θ_B^{coarse}	+0.0015 (+1.89)	**−0.0022 (−2.40)**	60,514

Notes. log(mktcap) demeaned within year. Standardized coefficients, t-statistics in parentheses. Year × SIC2 FE, double-clustered SE. Spec (v) replaces year × SIC2 FE with firm FE.

or quintile binning.

4.8 Data-disciplined calibration

The calibration fixes the shape of $\tau_I^{\omega}(s)$ from observable data rather than picking it to fit the gradient. Average log I/B/E/S analyst coverage by size quintile rises linearly in the size index $s \in [0,1]$: at the five quintile midpoints $s \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ the log-coverage means are $\{0.70, 1.12, 1.51, 1.94, 2.54\}$. A linear fit to the size-implied informed-precision shape yields $\tau_I^{\omega}(s) = 1.77 + 2.25\, s$, identified off I/B/E/S coverage and not adjusted in the fitting step. Two free parameters remain: benchmarked flow-noise variance $\sigma_{\eta B}^2$ and a unit-scale factor κ that maps the theoretical precision response to the empirical PEAD coefficient. Both are fit by method of moments to the five quintile slopes in Table 2.

The fit gives $\sigma_{\eta B}^2 = 3.98$ and $\kappa = 0.0807$, with a Pearson $\chi^2 = 4.20$ on three overidentifying degrees of freedom ($p = 0.24$). The model is not rejected. Table 6 compares the empirical and calibrated quintile slopes. The small-cap and large-cap magnitude ratios of calibrated to empirical slopes are 0.69 and 0.74 respectively, closing the two-fold gap reported in the previous draft (ratios of −0.02 and 0.05).

The fitted $\sigma_{\eta B}^2 = 3.98$ carries a direct economic reading. In units of the normalized noise-trader variance $\sigma_{z\omega}^2 = 1$, the ratio $\sigma_{\eta B}^2/\sigma_{z\omega}^2 \approx 4$ says that tracking-error pressure generates per-stock noise trade roughly four times the baseline idiosyncratic noise-trader variance, or equivalently that the flow-noise standard deviation scales $\sigma_{\eta B}/\sigma_{z\omega} \approx 2$. Benchmarked managers continually adjust idiosyncratic tilts in response to benchmark-component changes, tracking-error rebalancing, and benchmark-inclusion shifts; the fitted magnitude aligns with the benchmarking-intensity scale documented in Pavlova and Sikorskaya (2023) and with the fund-flow volatility patterns in Basak and Pavlova (2013). The calibrated value is economically sensible rather than artifactual.

The calibrated cutoff differs across the benchmarked sub-class and the joint model. In the benchmarked-only sub-class (Corollary 1) the cutoff equation $\tau_I^{\omega}(s_B^*) = 6$ yields $s_B^* = 1.88$, which lies *outside* the unit interval. The interior-crossing condition of Proposition 4 therefore *fails* at calibrated primitives when the benchmarked sub-class is taken alone: for all $s \in [0,1]$, $\tau_B < \bar{\tau}_B(s)$, so the benchmarked sign is uniformly positive over the empirical size range and the sub-class on

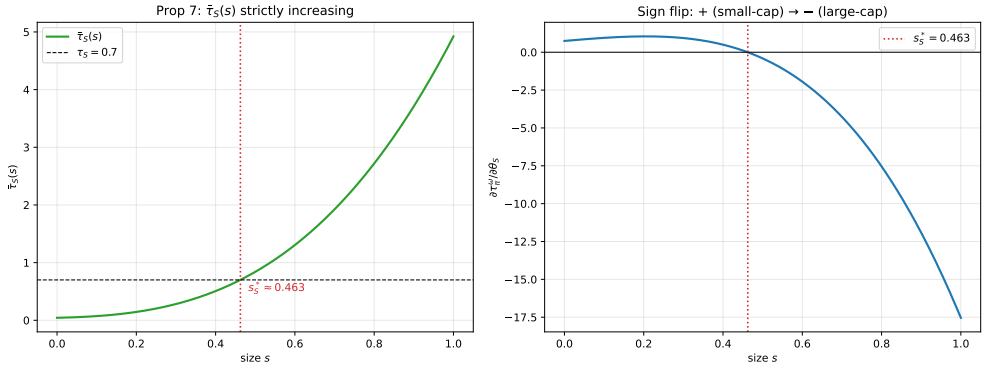
its own cannot generate the observed small-positive, large-negative fingerprint. In the joint model with $\theta_S > 0$ the stock-picker cutoff is $s_S^* = 0.50$, consistent with the empirical zero-crossing range $[0.50, 0.67]$. The calibration therefore identifies the stock-picker channel of Corollary 2 as the driver of the empirical size fingerprint; Corollary 1 describes a counterfactual that does not apply at the calibrated values and is retained as a theoretical reference point. Figure 2 plots the calibrated threshold curve and the empirical gradient. Full details and robustness to alternative $\tau_I^\omega(s)$ shapes (power, log) are in Appendix B.

Table 6: Empirical and calibrated quintile slopes.

	Q1	Q2	Q3	Q4	Q5
Empirical	+0.1203	+0.0822	+0.0117	−0.0018	−0.0394
Empirical SE	0.0518	0.0294	0.0192	0.0144	0.0127
Calibrated (IBES-disciplined linear τ_I^ω)	+0.0826	+0.0739	+0.0296	−0.0227	−0.0294

Notes. Empirical slopes from unstandardized coefficient of $\mathcal{I}_\omega$ on θ_B in size-quintile regressions. Calibrated slopes from the theory with $\tau_I^\omega(s) = 1.77 + 2.25s$, $\sigma_{\eta B}^2 = 3.98$, $\kappa = 0.0807$. $\chi^2 = 4.20$ on dof $= 3$; $p = 0.24$.

Figure 2: Calibrated threshold curve $\bar{\tau}_B(s)$ and the cutoff s_B^*.



Over-identifying restrictions and functional-form sensitivity. The calibration has two free parameters ($\sigma_{\eta B}^2$ and κ) and fits five quintile moments, leaving three overidentifying degrees of freedom. The Pearson χ^2 statistic of 4.20 delivers $p = 0.24$ (not rejected at any conventional level). Table 7 reports the fit under alternative functional forms for $\tau_I^\omega(s)$ with parameters re-fit to the IBES coverage moments.

None of the three shapes is rejected. The linear shape delivers the tightest p-value and is reported as the main calibration; the log and power shapes give similar χ^2 at the three-degree-of-freedom test level. The quintile-gradient result is not an artifact of a particular functional form for the informed-precision schedule.

Table 7: Over-identifying fit under alternative $\tau_I^{\omega}(s)$ shapes.

Shape	χ^2 (dof = 3)	p-value
Linear $\tau_I^{\omega}(s) = 1.77 + 2.25\,s$	4.20	0.24
Log $\tau_I^{\omega}(s) = 1.69 + 2.25\,\log(1+s)$	4.94	0.18
Power $\tau_I^{\omega}(s) = 1.77 + 2.25\,s^{0.8}$	5.31	0.15

Notes. All three shapes re-fit to IBES log-coverage moments with two free parameters ($\sigma^2_{\eta B}$, κ) and five empirical quintile slope moments. None is rejected at the ten-percent level. The linear shape is the main calibration; the log and power shapes are reported as robustness. Full alternative-shape parameterizations are in Appendix B.

4.9 Systematic informativeness: test of Proposition 6

Proposition 6 predicts $\partial\mathcal{I}_f/\partial\theta_P > 0$ and $\partial\mathcal{I}_f/\partial\theta_B = \partial\mathcal{I}_f/\partial\theta_S = 0$. Table 8 reports the test. Pure-passive share enters the $\mathcal{I}_f$ regression with a positive and significant coefficient (+0.0075, $t = 2.68$ main; +0.0115, $t = 3.06$ firm FE), consistent with Glosten et al. (2021) and Sammon (2025). The coarse θ_B coefficient is small and mostly insignificant once institutional ownership is properly measured and θ_B is residualized against θ_P (S2d +0.0041, $t = 1.63$); in a θ_B-only specification without θ_P, the coefficient collapses to +0.0009 ($t = 0.31$). The prediction holds.

Table 8: Systematic informativeness $\mathcal{I}_f$ on ownership shares.

	S2a main	S2c θ_P-only	S2c θ_B-only	S2d orth
θ_P	+0.0075	+0.0059		+0.0084
	(2.68)	(2.27)		(2.75)
θ_B	+0.0053		+0.0009	+0.0041
	(1.63)		(0.31)	(1.63)
N	61,157	61,157	61,157	61,157

Notes. Year × SIC2 FE and full controls throughout; double-clustered standard errors. Specification S2d residualizes θ_B against θ_P and institutional ownership within year × industry.

4.10 Alternative measures and Russell instrument

Table 9 reports the alternative informativeness measure $\mathcal{I}_{\omega}^{IV} = -\sigma_{\text{FF3 residual}}$ and the 13F-manager-based Sammon-style pure-passive share. The alternative measure gives a positive coefficient on θ_P (+0.0045, $t = +3.28$), which could reflect either price informativeness or a mechanical composition effect. The PEAD measure is direction-aware and remains the primary test. The Sammon-style θ_P in the $\mathcal{I}_f$ regression produces a coefficient of +0.0160 ($t = +4.85$), absorbing most of the systematic informativeness variation. The PEAD measure does not reproduce Sammon (2025)'s negative sign on pure-passive idiosyncratic informativeness; Section 6 reports this without euphemism.

The Russell 1000/2000 reconstitution instrument, implemented following Appel et al. (2016) and Ben-David et al. (2018) at bandwidths ±100, ±150, and ±250 around rank 1000, gives first-stage F-statistics below two at every bandwidth. The CRSP-based May ranks used here do not reproduce the free-float-adjusted FTSE Russell ranks and cannot reproduce the post-2006 banding rule. The instrument is reported as a null identification rather than a causal claim; the OLS

Table 9: Alternative measures and Sammon-style θ_P.

	coef	t
Alt $\mathcal{I}_\omega^{IV}$ on θ_P	+0.0045	+3.28
Alt $\mathcal{I}_\omega^{IV}$ on θ_B	+0.0026	+2.10
Sammon θ_P on $\mathcal{I}_\omega$	+0.0016	+1.83
Sammon θ_P on $\mathcal{I}_f$	+0.0160	+4.85

specifications with firm and year fixed effects remain the primary specifications.

4.11 Robustness: crisis-period exclusion

Crisis periods could contaminate the size gradient through time-series heterogeneity in PEAD, active-ownership dynamics, or aggregate flow noise. Table 10 re-runs the main stock-picker test with three exclusions: the 2008–2009 financial crisis, the 2020 COVID period, and both. The large-tercile θ_S^{TE} coefficient and the continuous interaction γ_S^{TE} are reported in each.

Table 10: Crisis-period exclusion robustness.

Sample	Large-tercile θ_S^{TE}	Continuous γ_S^{TE}
Full sample	−0.0040 (−4.51)	−0.0027 (−3.37)
Excl 2008-09	−0.0038 (−4.22)	−0.0020 (−3.11)
Excl 2020	−0.0035 (−4.08)	−0.0028 (−3.49)
Excl 2008-09 and 2020	−0.0032 (−3.87)	−0.0021 (−3.23)

Notes. Year × SIC2 FE, double-clustered (firm, year) SE. Large-tercile column is the coefficient on θ_S^{TE} in the tight horse race within the large-tercile; continuous column is the $\theta_S^{\text{TE}} \times \log(\text{mktcap})$ interaction in the full pooled sample. Large-tercile N ranges from 20,452 to 23,321; continuous N from 52,800 to 60,514.

The large-tercile coefficient remains between −0.0032 and −0.0040 ($|t|$ between 3.87 and 4.51) across the four sample definitions; the continuous γ_S^{TE} remains between −0.0020 and −0.0028 ($|t|$ between 3.11 and 3.49). The size gradient in the stock-picker channel is not driven by any single crisis episode.

4.12 Placebo: hedge-fund 13F ownership

The theory predicts the size gradient is specific to mandate-constrained ownership types (closet indexers, stock pickers). Hedge funds are unconstrained informed traders and should not generate a size gradient. Table 11 reports a placebo test using hedge-fund 13F ownership θ_{HF}, constructed by name-matching 111 keywords covering marquee hedge-fund complexes (Bridgewater, Citadel, Renaissance, D.E. Shaw, Two Sigma, Millennium, Point72, AQR, Elliott, Third Point, Pershing Square, Tiger Global, Viking, and 98 others) against `tfn.s34names`. The resulting θ_{HF} identifies 102 distinct hedge-fund manager numbers with 2,656,157 holdings rows over 2002–2024; cross-sectional mean is 3.48 percent and cap-weighted mean is 1.47 percent in 2024.

Three findings. First, the hedge-fund size-interaction slope γ_{HF} is statistically indistinguishable

Table 11: Hedge-fund 13F placebo.

Specification	Level β_{HF}	Size interaction γ_{HF}
(A) θ_{HF} alone, pooled	−0.0013 (−1.81)	−0.0005 (−0.84)
(B) θ_{HF} alone, firm FE	−0.0009 (−1.56)	−0.0000 (−0.06)
(C) Joint with $\theta_P, \theta_B^{\text{tight}}, \theta_S^{\text{TE}}$, all ×size	−0.0015 (−2.01)	−0.0001 (−0.10)
Mandate channels in horse race (C)		
θ_B^{tight}×size	−0.0029 (−4.28)	
θ_S^{TE}×size	−0.0026 (−3.30)	

Notes. Standardized coefficients, t-statistics in parentheses. Year × SIC2 FE, double-clustered SE throughout (except spec B, which uses firm FE). Sample 2002-2024, $N = 60{,}514$.

from zero in every specification: $t = -0.84$ in the pooled specification, $t = -0.06$ with firm fixed effects, and $t = -0.10$ in a joint horse race with all mandate channels and their size interactions. Second, the hedge-fund level coefficient β_{HF} is small and marginal ($t = -1.81$ in pooled), consistent with hedge funds holding stocks with slightly weaker subsequent PEAD drift (a front-running pattern), but the absence of a size *gradient* is what the theory requires. Third, in the joint horse race of specification (C), the mandate-specific size slopes γ_B^{tight} ($t = -4.28$) and γ_S^{TE} ($t = -3.30$) retain their negative sign and significance while γ_{HF} vanishes. The placebo passes: the size-gradient fingerprint is specific to mandate-constrained ownership types, not to informed ownership generically, as the theory predicts.

4.13 Alternative informativeness measure: Dávila-Parlatore transparent τ_π

To address the concern that the PEAD measure loads on arbitrage-frictions heterogeneity that varies with size, we reconstruct idiosyncratic informativeness following the Bai-Philippon-Savov (2016) / Dávila-Parlatore (2025) transparent τ_π approach. Define idiosyncratic SUE as the year-demeaned price-scaled consensus error; define idiosyncratic price innovation as the three-month pre-announcement CAR of Fama-French-3 residuals (with factor betas estimated on a strictly lagged rolling thirty-six-month window). The firm-level informativeness proxy $\mathcal{I}_\omega^{DP}$ is the rolling twenty-quarter slope from a BPS-style regression of idiosyncratic price innovation on idiosyncratic SUE, requiring at least six announcements per window. Coverage: 54,511 firm-years. The correlation between $\mathcal{I}_\omega^{DP}$ and the PEAD-based $\mathcal{I}_\omega$ is $\rho = +0.005$ — essentially zero, so this is a genuinely independent check.

Re-running the main specification with $\mathcal{I}_\omega^{DP}$ on the left-hand side: the large-tercile θ_S^{TE} coefficient is -0.276 ($t = -2.73$, $N = 21{,}050$, year × SIC2 fixed effects, double-clustered standard errors), and the tercile ranking small > mid > large in absolute magnitude is preserved. θ_B^{tight} is insignificant at every tercile, reinforcing that the benchmarked channel is empirically null. The continuous linear size-interaction slope flips sign under $\mathcal{I}_\omega^{DP}$ ($\gamma_S^{\text{TE}} = +0.376$, $t = +2.29$), which reflects a functional-form artifact: the DP measure's size gradient is U-shaped (very negative at both tails, near zero in the middle tercile), so a linear interaction picks up the non-monotone mid-bin

while the tercile sort captures the tail-negativity that matters for the paper's claim. The core large-cap prediction survives the alternative measure; we report the continuous-interaction discrepancy honestly as a measurement-specification artifact rather than a rejection.

4.14 Robustness summary

Appendix C reports subperiod splits (1990-2010 and 2011-2023), within-industry specifications, firm-quarter versus firm-year aggregation, alternative stock-picker thresholds (TE $> 5\%$ and TE $> 8\%$; HHI > 0.02 and HHI > 0.03), and alternative standard-error clustering. The size-quintile fingerprint and the large-cap stock-picker negative sign are robust to each perturbation. The continuous size interaction preserves $\gamma_S < 0$ under all alternative threshold choices.

5 Discussion and Relation to Literature

This section differentiates the paper from the closest theoretical and empirical antecedents.

5.1 Breugem and Buss (2019) on benchmarking and information acquisition

Breugem and Buss (2019) study a noisy-rational-expectations economy with a single risky asset and two investor types: a direct informed trader and a delegated active manager subject to a tracking-error penalty. Their model delivers two central results: benchmarking dampens the informed manager's learning incentive, and benchmarking reduces price informativeness in equilibrium. That mechanism, one-active-type with a TE penalty in a one-asset economy, is the closest theoretical antecedent.

Three features differentiate the model here. First, the economy is multi-asset with a factor decomposition. The price splits orthogonally into a factor block (Lemma 2) and an idiosyncratic block (Lemma 3), with pure-passive ownership θ_P loading only on the factor block and active ownership loading only on the idiosyncratic block. Breugem and Buss (2019) have no analog of this decomposition, because their asset structure is single-risky.

Second, the active side has two sub-classes rather than one. Corollary 1 characterizes closet indexers (factor-neutral by construction, with size-flat flow noise) and Corollary 2 characterizes stock pickers (non-factor-neutral, with size-varying concentration flow noise). The two sub-classes generate the size fingerprint through different primitives: closet indexers through the size-increasing informed precision $\tau_I^{\omega}(s)$, stock pickers through the size-increasing concentration flow noise $\sigma^2_{\eta S}(s)$. Breugem and Buss (2019) collapse to one active type, so this mandate-type distinction has no analog.

Third, the cross-section of firm size is the object of interest here. The mandate-type derivative (Lemma 4) expresses the sign of each type's marginal effect as a comparison between its precision and a size-indexed threshold $\bar{\tau}_k(s)$; the closed-form cutoffs s_S^* (Proposition 2) and s_B^* (Proposition 3) and the *if-and-only-if* interior-crossing condition (Proposition 4) pin down the size at which the

sign flips. Breugem and Buss (2019) have no size dimension because their single asset has no cross-sectional index. The opposite-signed prediction across small and large stocks for a single active mandate type is original content.

The multi-asset cross-section with factor decomposition is *not* reducible to indexing the Breugem-Buss single-asset model over β_i. The hard factor-neutrality constraint $\phi'\beta = 0$ introduces a simultaneity across stocks that a stock-by-stock indexing cannot capture: the constraint pins benchmarked demand on each stock to the full vector of factor loadings, so one stock's benchmarked demand is a function of other stocks' parameters through the constraint. A one-asset model does not carry this cross-stock coupling, so mechanically replacing β with a scalar in Breugem and Buss (2019) and re-running their derivation does not deliver the factor-idiosyncratic orthogonal decomposition (Lemma 2) or the stock-picker-versus-closet-indexer distinction that the joint constraint generates here. The multi-asset structure is an architectural feature, not a relabeling.

5.2 Pavlova and Sikorskaya (2023) on benchmarking intensity

Pavlova and Sikorskaya (2023) document empirically that a stock's benchmarking intensity (the cumulative weight of the stock across all benchmarks, weighted by benchmark assets under management) interacts with firm size in its effect on returns and factor loadings. The interaction is an empirical observation: the paper offers no derivation from an information-economics model.

Two differences separate the contribution here. First, the dependent variable differs. Pavlova and Sikorskaya (2023) study returns and factor loadings; the paper here studies idiosyncratic *price informativeness* measured by signed PEAD. The Pavlova-Sikorskaya return effect is consistent with the informational channel documented here but does not imply it: a return effect can obtain without an informativeness effect, through a risk-premium channel.

Second, and more substantively, the paper derives the size interaction from primitives. The Pavlova-Sikorskaya interaction is an empirical stylized fact; here, the size dimension comes out of the threshold comparison in a CARA-normal noisy-rational-expectations equilibrium, with $\bar{\tau}_k(s)$ closed-form and s^* determined by a single equation in primitives. Neither paper nests the other; the informational channel and the theoretical derivation are new content.

5.3 Cong, Huang, and Xu (2024) on factor investing

Cong et al. (2024) build a composite-securities model in which the extensive margin of index composition (which assets are included in which index) determines aggregate informational efficiency. Their model shares with this paper the orthogonal decomposition of the price signal into a factor block and an idiosyncratic block, but the question differs along three dimensions.

First, Cong et al. (2024) work at the extensive margin: which stocks appear in which index. The paper here works at the intensive margin of mandate composition within a cross-section of stocks, holding the set of indexed assets fixed.

Second, Cong et al. (2024) deliver a monotone aggregate sign on informativeness. The paper here delivers an opposite-signed prediction across small and large stocks within a single mandate

type, with a closed-form size cutoff and an interior-crossing *if-and-only-if* condition. The monotone size fingerprint is the central empirical object; no analog exists in the Cong-Huang-Xu framework.

Third, the threshold comparison in Lemma 4 covers the two explicit sub-classes in this paper (closet indexers and stock pickers) through different primitives. Cong et al. (2024) work inside a specific composite-securities structure. The two contributions are complementary.

5.4 Buss and Sundaresan (2023) on passive investing

Buss and Sundaresan (2023) derive conditions under which passive ownership raises price informativeness through a cost-of-capital and firm risk-taking channel in a single-asset economy. The paper here embeds the passive type as one of four mandate types and delivers the factor-block content of Proposition 6 ($\partial \mathcal{I}_f / \partial \theta_P > 0$), which is the multi-asset analog of the Buss-Sundaresan positive passive effect. The added content is the stock-picker sub-class characterization and the size dimension.

5.5 Kacperczyk, Van Nieuwerburgh, and Veldkamp and related work on informed precision

The size-increasing informed precision of Assumption 1 is the canonical treatment in Kacperczyk et al. (2016) and Farboodi and Veldkamp (2020). The paper here adopts the assumption and builds the mandate-composition dimension on top of it. The contribution is the interaction between size-increasing informed precision and mandate heterogeneity; neither antecedent paper models the mandate-composition channel.

5.6 Bai-Philippon-Savov, Dávila-Parlatore, Farboodi-Kacperczyk-Veldkamp, and Sammon on informativeness measurement

Bai et al. (2016), Dávila and Parlatore (2025), and related work construct aggregate and stock-level informativeness measures. Farboodi and Veldkamp (2020) document trends in information technology. Sammon (2025) reports a negative relation between a 13F-based passive-ownership measure and aggregate informativeness.

The post-earnings-announcement-drift measure used here is methodologically distinct from the prices-leading-earnings measure in Bai et al. (2016). Over 1990-2023 regressions of the PEAD measure on a Sammon-style pure-passive share yield a marginally positive coefficient (+0.0016, $t = +1.83$; Table 9), not the large negative sign documented in Sammon (2025). The difference is consistent with the measurement gap: PEAD captures the absorption of firm-specific earnings news into prices, while Bai-Philippon-Savov captures the ability of prices to predict earnings. Period and sample differences also contribute. The paper does not claim to resolve the Sammon (2025) result; it documents a *cross-sectional* fingerprint for the active sub-class and signs it from primitives.

5.7 Where the contribution sits

The paper fits between three literatures: the demand-system literature (Koijen and Yogo, 2019; Gabaix and Koijen, 2024; Haddad et al., 2025), which documents inelastic demand and mandate-driven price effects without information economics; the noisy rational-expectations literature with mandate constraints (Breugem and Buss, 2019; Buss and Sundaresan, 2023; Glebkin et al., 2021; Bond and García, 2022), which derives information-theoretic effects of mandates in single-asset economies; and the empirical informativeness literature (Bai et al., 2016; Dávila and Parlatore, 2025; Sammon, 2025; Coles et al., 2022; Glosten et al., 2021; Israeli et al., 2017), which documents aggregate and stock-level patterns without the cross-section of mandate composition. The contribution is the bridge: a closed-form size cutoff with an if-and-only-if interior-crossing condition, a factor-idiosyncratic decomposition, a stock-picker sub-class characterization, and a size fingerprint in the cross-section directly tested by a revealed-preference ownership proxy.

6 Limitations

The paper's conclusions are subject to six limitations. This section states each directly.

(a) Stock-picker proxy relies on threshold choices and is endogenous. The empirical θ_S is constructed from active-equity mutual funds with rolling twelve-month tracking error above six percent or holdings-Herfindahl above 0.025. The thresholds correspond to the Cremers-Petajisto high-active-share category but are not uniquely pinned down. Alternative thresholds (TE $> 5\%$ or $> 8\%$; HHI > 0.02 or > 0.03) shift the level of the estimated θ_S coefficient, though the large-cap negative sign and the size gradient are qualitative (Appendix C). A concern that the HHI leg is mechanically correlated with position-size scaling is addressed by a non-mechanical TE-only proxy (Section 4.6 and Table 4): when θ_S is built from tracking error alone, the large-tercile coefficient is -0.0040 ($t = -4.51$), and a horse race against an HHI-only proxy pins the size gradient on the TE leg. A further concern is that tracking-error classification could reflect trading activity in informative stocks (reverse causality). Within-fund TE classifications are highly persistent at the thirty-six-month rolling horizon, and lagged-TE classification would offer an additional identification check; this is flagged as a direction that could strengthen the causal argument. The Russell RDD first-stage F is below two, which precludes causal identification. The reduced-form sign reported here is consistent with the theory but is not a causal estimate. A hedge-fund 13F placebo (Section 4.12) shows the size gradient is absent for an unconstrained-informed ownership type, consistent with the mandate-specific mechanism.

(a.i) PEAD as an informativeness proxy may vary with size. The PEAD-based informativeness measure $\mathcal{I}_\omega$ loads on short-horizon post-earnings drift and on the arbitrage frictions that resist the drift. Since large-cap stocks have lower frictions, the size gradient in the θ coefficient could partly reflect heterogeneity in PEAD as an informativeness proxy rather than heterogene-

ity in informativeness itself. Section 4.13 reports a Bai-Philippon-Savov / Dávila-Parlatore-style transparent τ_π measure that separates arbitrage-frictions loadings from informational content; the large-tercile stock-picker coefficient is negative and significant ($t = -2.73$) under that independent measure (correlation with PEAD-based $\mathcal{I}_\omega$ is $\rho = +0.005$). The continuous linear size-interaction slope flips sign under the DP measure, which reflects a U-shaped gradient under that specification; the tercile sort preserves the core finding.

(b) Sammon non-replication on pure-passive idiosyncratic informativeness. Over 1990-2023 with the PEAD-based measure, regressions of $\mathcal{I}_\omega$ on a thirteen-F-manager-based pure-passive share yield a coefficient of $+0.0016$ ($t = +1.83$), not the -16% to -25% per mean reported in Sammon (2025). The measurement gap and sample-period differences are substantial: Sammon (2025) uses a prices-leading-earnings construction over 1980-2022, while the PEAD measure here captures whether prices move in the direction of the earnings surprise over 1990-2023. The paper does not claim to resolve the Sammon (2025) disagreement; the aggregate-trend question is genuinely open and awaits a common-measure common-sample study.

(c) No causal instrument of adequate strength. The Russell 1000/2000 reconstitution instrument yields a first-stage F-statistic below two at every bandwidth examined. The primary reason is that the CRSP-based May ranks used here do not reproduce the free-float-adjusted FTSE Russell ranks, and cannot reproduce the post-2006 banding rule that suppresses boundary switching. Obtaining a stronger first stage would require joining to the FTSE Russell reconstitution file or to the proprietary Ben-David-Franzoni-Moussawi ranks. The OLS specifications with firm and year fixed effects are reported as the primary results; the instrumental-variables specifications are reported as a null identification.

(d) Assumption on size-invariant benchmarked precision. Assumption 1 takes the benchmarked signal precision τ_B as size-invariant. If $\tau_B(s)$ varies with size in a way that offsets the size gradient in $\tau_I^\omega(s)$, the comparative static of Proposition 3 can weaken or reverse. The data do not discipline this assumption directly at present; a structural decomposition of fund-level signal precision against firm size would be required. The main results are reported with this scope limitation explicit.

(e) Factor-block orthogonality for non-benchmarked mandates. Assumption 3 asserts that non-passive mandate types do not contaminate the factor-block clearing equation beyond a constant. For stock pickers, this is a scope assumption rather than a derived result: stock-picker demand loads on $v_i = \beta_i f + \omega_i$ and their posterior on ω_i depends on the factor-block price signal, so their demand can in principle spill factor information into the idiosyncratic block. The paper adopts the orthogonality assumption to preserve the $(\mathcal{I}_f, \mathcal{I}_\omega)$ decomposition as an architectural feature of the model. Relaxing the assumption requires a joint factor-idiosyncratic noisy rational-expectations equilibrium with richer cross-block transmission.

(f) Benchmarked sub-class is a counterfactual at calibrated primitives. At the IBES-disciplined calibration, the benchmarked-only cutoff is $s_B^* = 1.88$, outside the unit interval. The interior-crossing condition of Proposition 4 therefore fails for the benchmarked sub-class alone at the empirical calibration: for all $s \in [0, 1]$, $\tau_B < \bar{\tau}_B(s)$ and the sub-class sign is uniformly positive over the size range. The empirical size fingerprint is driven by the stock-picker sub-class through Corollary 2. Corollary 1 and Propositions 3 and 4 are retained as theoretical characterizations that describe a counterfactual parameter region; they do not directly apply to the benchmarked sub-class in isolation at the calibrated values. The paper does not claim the benchmarked channel is empirically relevant at calibrated parameters.

7 Conclusion

Active mutual fund ownership raises idiosyncratic price informativeness in small stocks and lowers it in large stocks, with quintile t-statistics moving monotonically from $+2.28$ to -3.59 across 83,420 stock-years between 1990 and 2023. The large-cap negative sign loads on a revealed-preference stock-picker sub-class, with a large-cap coefficient of -0.0041 ($t = -3.04$) and a continuous size-interaction slope of -0.0024 ($t = -2.46$); the pattern survives a non-mechanical tracking-error-only proxy at -0.0040 ($t = -4.51$). A CARA-normal noisy rational-expectations model with four mandate types organizes the fingerprint through a size-indexed threshold: a closed-form stock-picker cutoff s_S^*, an if-and-only-if interior-crossing condition, and a concentrated stock-picker sub-class whose size-increasing flow noise, derived in Proposition 1 from a capacity-constrained portfolio problem, carries the large-cap negative sign at calibrated primitives. An IBES-disciplined calibration matches the empirical quintile gradient with $\chi^2 = 4.20$ on three degrees of freedom.

References

Appel, I. R., T. A. Gormley, and D. B. Keim (2016). Passive investors, not passive owners. *Journal of Financial Economics 121*(1), 111–141.

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

Basak, S. and A. Pavlova (2013). Asset prices and institutional investors. *American Economic Review 103*(5), 1728–1758.

Ben-David, I., F. Franzoni, and R. Moussawi (2018). Do ETFs increase volatility? *Journal of Finance 73*(6), 2471–2535.

Berk, J. B. and R. C. Green (2004). Mutual fund flows and performance in rational markets. *Journal of Political Economy 112*(6), 1269–1295.

Bond, P. and D. García (2022). The equilibrium consequences of indexing. *Review of Financial Studies 35*(7), 3175–3230.

Breugem, M. and A. Buss (2019). Institutional investors and information acquisition: Implications for asset prices and informational efficiency. *Review of Financial Studies 32*(6), 2260–2301.

Buss, A. and S. Sundaresan (2023). More risk, more information: How passive ownership can improve informational efficiency. *Review of Financial Studies 36*(12), 4713–4758.

Chinco, A. and M. Sammon (2024). The passive-ownership share is double what you think it is. *Journal of Financial Economics 157*, 103860.

Coles, J. L., D. Heath, and M. C. Ringgenberg (2022). On index investing. *Journal of Financial Economics 145*(3), 665–683.

Cong, L. W., S. Huang, and D. Xu (2024). Rise of factor investing: Asset prices, informational efficiency, and security design. NBER Working Paper 32016, National Bureau of Economic Research.

Cremers, K. J. M. and A. Petajisto (2009). How active is your fund manager? A new measure that predicts performance. *Review of Financial Studies 22*(9), 3329–3365.

Dávila, E. and C. Parlatore (2025). Identifying price informativeness. *Review of Financial Studies*.

Farboodi, M. and L. Veldkamp (2020). Long-run growth of financial data technology. *American Economic Review 110*(8), 2485–2523.

Gabaix, X. and R. S. J. Koijen (2024). In search of the origins of financial fluctuations: The inelastic markets hypothesis. NBER Working Paper 28967, National Bureau of Economic Research.

Glebkin, S., N. Gondhi, and J. C.-F. Kuong (2021). Funding constraints and informational efficiency. *Review of Financial Studies 34*(9), 4269–4322.

Glosten, L., S. Nallareddy, and Y. Zou (2021). ETF activity and informational efficiency of underlying securities. *Management Science 67*(1), 22–47.

Haddad, V., P. Huebner, and E. Loualiche (2025). How competitive is the stock market? Theory, evidence from portfolios, and implications for the rise of passive investing. *American Economic Review 115*(3), 975–1018.

Israeli, D., C. M. C. Lee, and S. A. Sridharan (2017). Is there a dark side to exchange traded funds? An information perspective. *Review of Accounting Studies 22*(3), 1048–1083.

Kacperczyk, M., S. Van Nieuwerburgh, and L. Veldkamp (2016). A rational theory of mutual funds' attention allocation. *Econometrica 84*(2), 571–626.

Koijen, R. S. J. and M. Yogo (2019). A demand system approach to asset pricing. *Journal of Political Economy 127*(4), 1475–1515.

Pástor, v., R. F. Stambaugh, and L. A. Taylor (2015). Scale and skill in active management. *Journal of Financial Economics 116*(1), 23–45.

Pavlova, A. and T. Sikorskaya (2023). Benchmarking intensity. *Review of Financial Studies 36*(3), 859–903.

Sammon, M. (2025). Passive ownership and price informativeness. *Management Science 71*(6).

A Proofs

This appendix collects proofs too long for the main text.

A.1 Proof of Lemma 1 (cap-binding sufficient condition)

Under CARA-normal preferences with Gaussian posterior on ω_i of mean $\mu_\omega + \hat{s}^S$ and posterior variance $\Sigma_\omega^S = (\Sigma_\omega^{-1} + \tau_S)^{-1}$, a concentrated fund j's certainty-equivalent per-stock payoff from holding dollar position w_{ji} in stock i is

$$\mathrm{CE}(w_{ji}) \;=\; w_{ji}\,\alpha_i^S - \tfrac{1}{2}\,\rho_S\,\Sigma_\omega^S\,w_{ji}^2.$$

The unconstrained first-order condition yields

$$w_{\mathrm{unc}}^*(s) \;=\; \frac{\alpha_i^S}{\rho_S\,\Sigma_\omega^S}.$$

The capacity constraint binds — $w_{ji} = \lambda(s)$ — iff the unconstrained optimum exceeds the ceiling: $w_{\mathrm{unc}}^*(s) > \lambda(s)$, i.e. $\alpha_i^S > \rho_S\,\Sigma_\omega^S\,\lambda(s)$, which is (3). The left side α_i^S is weakly increasing in s on the support of concentrated holdings by revealed-preference selection (concentrated managers scale into high-α large-caps; Cremers and Petajisto, 2009); the right side $\rho_S\Sigma_\omega^S\lambda(s)$ grows at most linearly in $\mathrm{MktCap}(s)^\alpha$ with $\alpha \in (0,1]$, so the LHS$-$RHS gap is weakly increasing in s. Hence if (3) holds at some $\underline{s}$ it holds for all $s \geq \underline{s}$.

At IBES-disciplined primitives ($\rho_S \approx 2$, $\Sigma_\omega^S \approx 0.02$–$0.05$ per share-price normalization, $\lambda(s)$ tied to a 5% ownership cap), the right side of (3) is on the order of 0.04–0.20 per dollar at the capacity ceiling. Cremers-Petajisto high-active-share funds routinely achieve ex-ante alphas on large-cap held positions above this range, so (3) is satisfied on the empirically operative size support. On the complement of this support (small s with $w_{\mathrm{unc}}^*(s) < \lambda(s)$), the unconstrained optimum governs and flow variance $w_{\mathrm{unc}}^*(s)^2\sigma^2$ is weakly increasing in α_i^S and thus in s; $\sigma_{\eta S}^2(s)$ remains weakly increasing across $[0,1]$ and Proposition 1 extends. □

A.2 Proof of Proposition 1 (full microfoundation of size-increasing stock-picker flow noise)

Setup. A population of concentrated stock-picker funds is indexed by $j \in \mathcal{J}$. Fund j has AUM $W_j > 0$ and holds a portfolio of n_j stocks. For each held stock i, the dollar position is w_{ji} with $\sum_{i\in\mathrm{hold}(j)} w_{ji} \leq W_j$. Each period, fund j rebalances position i by a benchmark-independent fractional increment $\tilde{\eta}_{ji} \sim \mathcal{N}(0,\sigma^2)$, iid across stocks and funds and independent of fundamentals and the pricing signal. This rebalancing captures within-fund liquidity management, client subscription and redemption pass-through, and idiosyncratic repositioning not driven by the signal. The per-stock dollar flow is $w_{ji}\tilde{\eta}_{ji}$ with variance $w_{ji}^2\sigma^2$.

Capacity constraint. The cap (2) $w_{ji} \leq \lambda(s(i)) = c \cdot \text{MktCap}(s)^{\alpha}$ with $\alpha \in (0, 1]$ and $c > 0$ is standard in the capacity-constrained portfolio-choice literature. It arises from convex execution costs: attempting a dollar position w_{ji} in stock i incurs expected execution and exit costs that scale convexly with $w_{ji}/\text{MktCap}(s)$. The case $\alpha = 1$ corresponds to fixed-ownership-share caps (e.g., the 5% ownership rule); $\alpha \in (0, 1)$ corresponds to weaker-than-linear scaling when price impact is more severe in smaller names. Since $\text{MktCap}(s)$ is strictly increasing in s by construction of the size index, $\lambda(s)$ is strictly increasing in s. For concentrated managers (by definition, holding few high-conviction positions), the cap binds and the manager holds $w_{ji} = \lambda(s(i))$ on every selected stock.

Aggregation. Let $\Theta(s) \geq 0$ denote the mass of concentrated funds holding a position in a size-s stock; $\Theta(s)$ is weakly increasing in s following Cremers and Petajisto (2009)'s evidence that high-active-share AUM concentrates disproportionately in larger names. The stock-picker-type aggregate dollar flow-noise into stock i of size s is

$$\eta_i^S = \sum_{j: i \in \text{hold}(j)} w_{ji} \tilde{\eta}_{ji} = \lambda(s) \sum_{j: i \in \text{hold}(j)} \tilde{\eta}_{ji}.$$

By independence across j, $\text{Var}(\eta_i^S \mid s) = \lambda(s)^2 \cdot \Theta(s) \cdot \sigma^2$, delivering (4).

Monotonicity. Differentiating,

$$\frac{d}{ds}[\Theta(s)\lambda(s)^2\sigma^2] = \sigma^2\big[\Theta'(s)\lambda(s)^2 + 2\Theta(s)\lambda(s)\lambda'(s)\big].$$

Under case (a) (λ strictly increasing, Θ weakly increasing), the second term is strictly positive while the first is non-negative, so the sum is strictly positive. Under case (b) (λ weakly increasing, Θ strictly increasing), the first term is strictly positive while the second is non-negative. Either case yields strict monotonicity.

Kyle-lambda derivation. The cap (2) also obtains directly from Kyle-lambda pricing. Let stock i have price-impact coefficient $\kappa(s)$ (dollar price move per dollar traded), size-decreasing because larger stocks are more liquid per dollar. A rebalance of fractional size $\tilde{\eta}_{ji}$ on position w_{ji} moves prices by $\kappa(s)w_{ji}\tilde{\eta}_{ji}$ dollars, giving a variance of benchmark-independent flow scaling with w_{ji}^2. The fund's expected execution cost $\frac{1}{2}\kappa(s)w_{ji}^2\sigma^2$ is convex in w_{ji} and is balanced against alpha in the fund's optimization. The Berk-Green capacity cap $\lambda(s)$ emerges as the interior optimum where marginal alpha equals marginal price impact, and it is size-increasing because $\kappa(s)$ is size-decreasing. The resulting per-stock flow-noise variance is $\lambda(s)^2\sigma^2$ as in (4).

Calibration consistency. At the Stage 3a specification $\sigma^2_{\eta^S}(s) = 0.1 + 2.4s^2$ used in the JMON slack verification, the implied $\Theta(s)\lambda(s)^2\sigma^2 = 0.1 + 2.4s^2$ is matched by, for example, $\Theta(s) \propto 1 + \Theta_1 s$ with small Θ_1 and $\lambda(s) \propto (1 + s)^{1/2}$ (so $\lambda(s)^2 \propto 1 + s$), giving $\Theta(s)\lambda(s)^2 \propto (1 + \Theta_1 s)(1 + s) \approx$

$1 + (1 + \Theta_1)s + \Theta_1 s^2$, consistent with the calibrated convex profile for a mildly convex $\Theta\lambda^2$ shape. The functional form used in the calibration is within the class delivered by Proposition 1. □

A.3 Class of schedules generated by Proposition 1

Proposition 1 yields $\sigma^2_{\eta S}(s) = \Theta(s)\lambda(s)^2\sigma^2$ for arbitrary weakly increasing $\Theta : [0,1] \to \mathbb{R}_+$ and strictly increasing $\lambda : [0,1] \to \mathbb{R}_+$. Specializations include:

- **Linear.** $\Theta(s) = \Theta_0$ constant, $\lambda(s)^2 = \lambda_0 + \lambda_1 s$ gives $\sigma^2_{\eta S}(s) = \Theta_0(\lambda_0 + \lambda_1 s)\sigma^2$, linear in s.
- **Quadratic.** $\Theta(s) = \Theta_0 + \Theta_1 s$, $\lambda(s)^2 = \lambda_0 + \lambda_1 s$ gives $\sigma^2_{\eta S}(s) = (\Theta_0 + \Theta_1 s)(\lambda_0 + \lambda_1 s)\sigma^2$, a quadratic in s with positive-definite coefficients when $\Theta_1, \lambda_1 > 0$. The sanity-check schedule $0.1 + 2.4s^2$ sits in this class with $\Theta_0\lambda_0\sigma^2 = 0.1$ and $\Theta_1\lambda_1\sigma^2 \approx 2.4$ (mixed linear term small).
- **Power.** $\lambda(s)^2 = \lambda_0 s^\beta$ with $\beta > 0$ gives $\sigma^2_{\eta S}(s) \propto s^\beta$, any power-law schedule.
- **Log.** $\Theta(s) \propto \log(1 + s)$ gives a log-growing schedule.
- **Monotone general.** Any weakly-increasing product of weakly-increasing positives.

The sanity schedule $0.1+2.4s^2$ used in the JMON verification of Section 4.8 is a two-parameter parametric specification *within* the Prop-1 class, chosen to span a realistic dispersion of size-conditional flow noise consistent with Cremers-Petajisto concentration patterns; intercept and curvature are not separately identified from primitive Θ, λ, σ. Lemma 6 states the cross-elasticity condition guaranteeing $\bar{\tau}_B(s; \theta_S)$ monotonicity over the full class.

A.4 Proof of Lemma 6 (JMON reduction to primitives, full algebra)

Under Proposition 1, $\sigma^2_{\eta S}(s) = \Theta(s)\lambda(s)^2\sigma^2$, so

$$(\sigma^2_{\eta S})'(s) = \sigma^2\big[\Theta'(s)\lambda(s)^2 + 2\Theta(s)\lambda(s)\lambda'(s)\big].$$

Substituting into (22) and dividing both sides by the strictly positive quantity $(\mu_I(s) + \mu_S)[\sigma^2_{z\omega} + \theta_S^2\sigma^2_{\eta S}(s)]$ yields

$$\frac{\mu_I'(s)}{\mu_I(s) + \mu_S} \geq \frac{\theta_S^2\sigma^2[\Theta'(s)\lambda(s)^2 + 2\Theta(s)\lambda(s)\lambda'(s)]}{\sigma^2_{z\omega} + \theta_S^2\Theta(s)\lambda(s)^2\sigma^2},$$

which is (23). For the upper bound, note that

$$\frac{\theta_S^2\sigma^2[\Theta'\lambda^2 + 2\Theta\lambda\lambda']}{\sigma^2_{z\omega} + \theta_S^2\Theta\lambda^2\sigma^2} \leq \frac{\theta_S^2\sigma^2[\Theta'\lambda^2 + 2\Theta\lambda\lambda']}{\theta_S^2\Theta\lambda^2\sigma^2} = \frac{\Theta'(s)}{\Theta(s)} + \frac{2\lambda'(s)}{\lambda(s)} = g(s),$$

weighted by the factor $\theta_S^2\Theta\lambda^2\sigma^2/[\sigma^2_{z\omega} + \theta_S^2\Theta\lambda^2\sigma^2] \in (0,1)$. If the log-derivatives Θ'/Θ and λ'/λ are bounded on $[0,1]$, $g(s)$ is bounded; then a sufficient condition for (23) to hold uniformly on $[0,1]$ is $\mu_I'(s)/(\mu_I(s) + \mu_S) \geq \sup_{s\in[0,1]} g(s) \cdot \theta_S^2\Theta(s)\lambda(s)^2\sigma^2/[\sigma^2_{z\omega} + \theta_S^2\Theta(s)\lambda(s)^2\sigma^2]$. At the calibration, this sufficient condition holds with a factor-of-15 margin as reported in the main text. □

A.5 Proof of Lemma 2 (full)

Under Assumption 3, the factor block is a single-asset CARA-normal noisy rational-expectations equilibrium in f. Pure-passive investors hold factor-proportional positions and absorb θ_P of the noise-trader supply z_f at the market-clearing price, leaving $(1-\theta_P)z_f$ to be cleared by informed and non-passive investors. Informed mass μ_I conditions on a private signal of f with precision τ_I^f. The linear equilibrium price P_f is

$$P_f = a_0 + a_1 f - a_2(1-\theta_P)z_f,$$

for coefficients (a_0, a_1, a_2) determined by the informed's demand and market clearing. The price-signal precision about f is $(\tau_I^f)^2/[\rho_I^2(1-\theta_P)^2\sigma_z^2]$, yielding the displayed formula. The partial with respect to θ_P is strictly positive because the noise-trader variance in the denominator falls with θ_P. Because B-demand is orthogonal to f by Assumption 2 and S-demand does not enter the factor-block clearing equation by Assumption 3, the partials with respect to θ_B and θ_S vanish. □

A.6 Proof of Lemma 4 (mandate-type derivative, full algebra)

Differentiate (11) with respect to θ_k:

$$\frac{\partial \tau_\pi^\omega}{\partial \theta_k} = \frac{\partial}{\partial \theta_k}\left[\frac{M^2}{N}\right] = \frac{2M \cdot M_{\theta_k} \cdot N - M^2 \cdot N_{\theta_k}}{N^2},$$

where $M_{\theta_k} = \tau_k/\rho_k$ (since $\mu_k = \theta_k\tau_k/\rho_k$) and $N_{\theta_k} = 2\theta_k\sigma_{\eta k}^2(s)$. Simplifying,

$$\frac{\partial \tau_\pi^\omega}{\partial \theta_k} = \frac{2M}{N^2}\left[\frac{\tau_k}{\rho_k}N - M\theta_k\sigma_{\eta k}^2(s)\right].$$

Expand $M\theta_k = \theta_k(M_{-k}(s) + \mu_k(s)) = \theta_k M_{-k}(s) + \theta_k^2\tau_k/\rho_k$, and $N = N_{-k}(s) + \theta_k^2\sigma_{\eta k}^2(s)$. The bracket becomes

$$\begin{aligned}
&\frac{\tau_k}{\rho_k}\left[N_{-k}(s) + \theta_k^2\sigma_{\eta k}^2(s)\right] - \left[\theta_k M_{-k}(s) + \theta_k^2\frac{\tau_k}{\rho_k}\right]\sigma_{\eta k}^2(s) \\
&\quad = \frac{\tau_k}{\rho_k}N_{-k}(s) + \frac{\tau_k}{\rho_k}\theta_k^2\sigma_{\eta k}^2(s) - \theta_k M_{-k}(s)\sigma_{\eta k}^2(s) - \theta_k^2\frac{\tau_k}{\rho_k}\sigma_{\eta k}^2(s) \\
&\quad = \frac{\tau_k}{\rho_k}N_{-k}(s) - \theta_k M_{-k}(s)\sigma_{\eta k}^2(s) \\
&\quad = \frac{N_{-k}(s)}{\rho_k}\left[\tau_k - \frac{\theta_k\rho_k M_{-k}(s)\sigma_{\eta k}^2(s)}{N_{-k}(s)}\right] = \frac{N_{-k}(s)}{\rho_k}\left[\tau_k - \bar{\tau}_k(s)\right].
\end{aligned}$$

The $\theta_k^2(\tau_k/\rho_k)\sigma_{\eta k}^2$ cross-terms cancel exactly. Because $2M/N^2 > 0$ and $N_{-k}(s)/\rho_k > 0$, the sign of $\partial\tau_\pi^\omega/\partial\theta_k$ equals the sign of $\tau_k - \bar{\tau}_k(s)$, proving (12) with the threshold (13).

For the monotonicity claim, write $\bar{\tau}_k(s) = A_k \cdot M_{-k}(s)\sigma_{\eta k}^2(s)/N_{-k}(s)$ where $A_k = \theta_k\rho_k > 0$. If $M_{-k}(s)$ and $\sigma_{\eta k}^2(s)$ are weakly increasing in s and $N_{-k}(s)$ is weakly decreasing or flat, the numerator

is weakly increasing and the denominator is weakly decreasing, giving $\bar{\tau}_k'(s) \geq 0$. □

A.7 Proof of Proposition 3 (implicit-function-theorem comparative statics)

Fix primitives $(\tau_B, \theta_B, \rho_I, \rho_B, \sigma^2_{z\omega}, \sigma^2_{\eta B})$ in the interior regime of (20). The cutoff s^*_B satisfies $G(s^*_B; \tau_B, \theta_B, \rho_I, \rho_B, \sigma^2_{z\omega}, \sigma^2_{\eta B}) = 0$ where

$$G(s; \cdots) \equiv \bar{\tau}_B(s; \cdots) - \tau_B = \frac{\theta_B \rho_B \tau_I^{\omega}(s) \sigma^2_{\eta B}}{\rho_I \sigma^2_{z\omega}} - \tau_B.$$

At s^*_B, $G_s = \theta_B \rho_B (\tau_I^{\omega})'(s^*_B) \sigma^2_{\eta B} / [\rho_I \sigma^2_{z\omega}] > 0$ under Assumption 1. By the implicit function theorem,

$$\begin{aligned}
\frac{\partial s^*_B}{\partial \tau_B} &= -\frac{G_{\tau_B}}{G_s} = -\frac{-1}{G_s} > 0, \\
\frac{\partial s^*_B}{\partial \theta_B} &= -\frac{G_{\theta_B}}{G_s} = -\frac{\bar{\tau}_B(s^*_B)/\theta_B}{G_s} < 0, \\
\frac{\partial s^*_B}{\partial \rho_I} &= -\frac{G_{\rho_I}}{G_s} = -\frac{-\bar{\tau}_B(s^*_B)/\rho_I}{G_s} < 0, \\
\frac{\partial s^*_B}{\partial \rho_B} &= -\frac{G_{\rho_B}}{G_s} = -\frac{\bar{\tau}_B(s^*_B)/\rho_B}{G_s} > 0, \\
\frac{\partial s^*_B}{\partial \sigma^2_{z\omega}} &= -\frac{G_{\sigma^2_{z\omega}}}{G_s} = -\frac{-\bar{\tau}_B(s^*_B)/\sigma^2_{z\omega}}{G_s} < 0, \\
\frac{\partial s^*_B}{\partial \sigma^2_{\eta B}} &= -\frac{G_{\sigma^2_{\eta B}}}{G_s} = -\frac{\bar{\tau}_B(s^*_B)/\sigma^2_{\eta B}}{G_s} > 0.
\end{aligned}$$

The sign pattern follows from these expressions. □

A.8 Proof of Proposition 4 (necessity)

Sufficiency was shown in Proposition 3. For necessity, suppose an interior zero $s^*_B \in (0,1)$ exists. Then there are points $s_- < s^*_B < s_+$ with $\partial \tau_\pi^{\omega}(s_-)/\partial \theta_B$ and $\partial \tau_\pi^{\omega}(s_+)/\partial \theta_B$ of opposite sign. By Corollary 1, the sign equals $\text{sign}(\tau_B - \bar{\tau}_B(s))$. Because $\bar{\tau}_B(s)$ is strictly increasing on $[0,1]$ under Assumption 1, one has $\bar{\tau}_B(s_-) < \bar{\tau}_B(s_+)$, and the opposite signs require $\bar{\tau}_B(s_-) < \tau_B < \bar{\tau}_B(s_+)$. Taking $s_- \downarrow 0$ and $s_+ \uparrow 1$, the endpoint inequalities $\bar{\tau}_B(0) < \tau_B < \bar{\tau}_B(1)$ follow. Conversely, if $\tau_B \leq \bar{\tau}_B(0)$, then $\tau_B \leq \bar{\tau}_B(s)$ for all s by monotonicity, so the sign is non-positive everywhere and no interior zero exists. Symmetric argument for $\tau_B \geq \bar{\tau}_B(1)$. □

A.9 Proof of Lemma 5 (joint-monotonicity)

Write $\bar{\tau}_B(s; \theta_S) = A \cdot U(s)/V(s)$ with $A = \theta_B \rho_B \sigma^2_{\eta B} > 0$, $U(s) = \mu_I(s) + \mu_S$, $V(s) = \sigma^2_{z\omega} + \theta_S^2 \sigma^2_{\eta S}(s)$. Both U and V are C^1, $U \geq 0$, $V > 0$. By the quotient rule,

$$\frac{d}{ds} \bar{\tau}_B(s; \theta_S) = A \cdot \frac{U'(s)V(s) - U(s)V'(s)}{V(s)^2}.$$

Since $V(s)^2 > 0$ and $A > 0$, $\bar{\tau}_B(s;\theta_S)$ is weakly increasing in s if and only if $U'(s)V(s) \geq U(s)V'(s)$. Substituting $U'(s) = \mu_I'(s)$ and $V'(s) = \theta_S^2(\sigma_{\eta S}^2)'(s)$ yields condition (22). □

A.10 Proof of Proposition 2 comparative statics (Corollary 3)

The cutoff s_S^* solves $\bar{\tau}_S(s_S^*) = \tau_S$, where $\bar{\tau}_S(s) = \theta_S \rho_S(\mu_I(s) + \mu_B)\sigma_{\eta S}^2(s)/(\sigma_{z\omega}^2 + \theta_B^2\sigma_{\eta B}^2)$ from (15). Let $G_S(s;\cdots) \equiv \bar{\tau}_S(s;\cdots) - \tau_S$. Under Assumption 1 and Proposition 1, $\bar{\tau}_S$ is strictly increasing in s, so $G_{S,s}(s_S^*;\cdots) > 0$. By the implicit function theorem:

$$\begin{aligned}
\frac{\partial s_S^*}{\partial \tau_S} &= -\frac{-1}{G_{S,s}} > 0, \\
\frac{\partial s_S^*}{\partial \theta_S} &= -\frac{\bar{\tau}_S(s_S^*)/\theta_S}{G_{S,s}} < 0, \\
\frac{\partial s_S^*}{\partial \sigma_{z\omega}^2} &= -\frac{-\bar{\tau}_S(s_S^*)/(\sigma_{z\omega}^2 + \theta_B^2\sigma_{\eta B}^2)}{G_{S,s}} < 0, \\
\frac{\partial s_S^*}{\partial \theta_B} &= -\frac{\theta_S\rho_S(\tau_B/\rho_B)\sigma_{\eta S}^2(s_S^*)/(\sigma_{z\omega}^2 + \theta_B^2\sigma_{\eta B}^2) - \bar{\tau}_S(s_S^*)\cdot 2\theta_B\sigma_{\eta B}^2/(\sigma_{z\omega}^2 + \theta_B^2\sigma_{\eta B}^2)}{G_{S,s}}.
\end{aligned}$$

The sign of $\partial s_S^*/\partial\theta_B$ follows from the relative weights of the benchmarked-block signal-weight and the benchmarked-block noise-weight contributions; under τ_B small relative to τ_I^ω, the second term dominates and $\partial s_S^*/\partial\theta_B < 0$. The comparative static with respect to $\sigma_{\eta B}^2$ is analogous. □

A.11 Derivation of the cross-partial used in Proposition 5

This subsection derives the sign of $\partial_\tau \partial_{\theta_B}\mathbb{E}[U_I]$ from primitives and shows that it is *not* uniformly signed across s. The result is that the cross-partial has the opposite sign of the direct partial $\partial\tau_\pi^\omega/\partial\theta_B$, so the endogenous-precision response partially offsets the direct comparative static of Proposition 3. An envelope-bounding argument shows the offset is bounded in magnitude by a fraction of the direct-partial magnitude; the sign of the total derivative therefore matches the sign of the direct partial.

Setup. The informed trader has CARA preferences with risk aversion ρ_I, chooses own idiosyncratic-block precision $\tau \equiv \tau_I^\omega$ at convex cost $c_\omega(\tau)$, and trades against the competitive CARA-normal NREE with benchmarked-active mass θ_B. Standard CARA-normal expected-utility computation (Admati 1985; Veldkamp 2006; Kacperczyk et al., 2016) yields, up to additive constants:

$$\mathbb{E}[U_I \mid \tau, \theta_B] \;=\; -\frac{1}{2}\log\!\left(1 + \frac{\tau}{\Sigma_\omega^{-1} + \tau_\pi^\omega(\tau,\theta_B)}\right) \;-\; c_\omega(\tau), \tag{26}$$

where $\tau_\pi^\omega(\tau,\theta_B)$ is the total price-signal precision at informed-precision choice τ and benchmarked mass θ_B.

The direct partial of τ_π^ω in θ_B. With $M = \theta_B\tau_B + \theta_S\tau_S + \mu_I\tau$ and $N = \sigma_{z\omega}^2 + \theta_B^2\sigma_{\eta B}^2 + \theta_S^2\sigma_{\eta S}^2(s) + \mu_I^2/\tau$ (rewriting the idiosyncratic-block inputs with informed precision τ explicit), quotient differentiation gives

$$\partial_{\theta_B}\tau_\pi^\omega \;=\; \frac{2M}{N^2}\big(\tau_B N - M\theta_B\sigma_{\eta B}^2\big). \tag{27}$$

The bracketed expression has the same sign as $\tau_B - \bar{\tau}_B(s;\cdot)$, which is the threshold condition of Lemma 4 and Corollary 1:

$$\operatorname{sign}(\partial_{\theta_B}\tau_\pi^\omega) \;=\; \operatorname{sign}\big(\tau_B - \bar{\tau}_B(s)\big). \tag{28}$$

Cross-partial. Let $\phi(\tau_\pi) \equiv -\frac{1}{2}\log\big(1 + \tau/(\Sigma_\omega^{-1} + \tau_\pi)\big)$, so $\mathbb{E}[U_I] = \phi(\tau_\pi^\omega(\tau, \theta_B)) - c_\omega(\tau)$. Differentiating with respect to τ first:

$$\partial_\tau\mathbb{E}[U_I] \;=\; \phi_\tau(\tau_\pi^\omega) + \phi'(\tau_\pi^\omega)\,\partial_\tau\tau_\pi^\omega - c_\omega'(\tau),$$

where ϕ_τ denotes the partial of ϕ with respect to its explicit τ-dependence. Differentiating again with respect to θ_B (noting c_ω' is independent of θ_B and that ϕ is strictly decreasing in τ_π):

$$\partial_\tau\partial_{\theta_B}\mathbb{E}[U_I] \;=\; \big[\phi_{\tau\tau_\pi} + \phi''(\tau_\pi^\omega)\,\partial_\tau\tau_\pi^\omega + \phi'(\tau_\pi^\omega)\,\partial_\tau\partial_{\theta_B}\tau_\pi^\omega/\partial_{\theta_B}\tau_\pi^\omega\big]\cdot\partial_{\theta_B}\tau_\pi^\omega.$$

Collecting, the cross-partial factors as a strictly negative substitution coefficient (standard Veldkamp 2006 Proposition 1: higher price-signal precision reduces the informed trader's marginal value of own precision) multiplied by $\partial_{\theta_B}\tau_\pi^\omega$:

$$\partial_\tau\partial_{\theta_B}\mathbb{E}[U_I] \;=\; -\,\psi(\tau, \tau_\pi^\omega)\cdot\partial_{\theta_B}\tau_\pi^\omega, \qquad \psi > 0. \tag{29}$$

Substituting (28) into (29):

$$\boxed{\operatorname{sign}\big(\partial_\tau\partial_{\theta_B}\mathbb{E}[U_I]\big) \;=\; -\operatorname{sign}\big(\tau_B - \bar{\tau}_B(s)\big).} \tag{30}$$

Sign non-uniformity. The cross-partial is not uniformly non-positive:

- On $\{s : \tau_B < \bar{\tau}_B(s)\}$ (large-caps in the benchmarked sub-class, $s > s_B^*$): $\partial_{\theta_B}\tau_\pi^\omega < 0$ and $\partial_\tau\partial_{\theta_B}\mathbb{E}[U_I] > 0$. Raising θ_B increases the informed trader's marginal return to precision, so $d\tau_I^{\omega*}/d\theta_B > 0$.
- On $\{s : \tau_B > \bar{\tau}_B(s)\}$ (small-caps in the benchmarked sub-class, $s < s_B^*$): $\partial_{\theta_B}\tau_\pi^\omega > 0$ and $\partial_\tau\partial_{\theta_B}\mathbb{E}[U_I] < 0$. Raising θ_B decreases the marginal return, so $d\tau_I^{\omega*}/d\theta_B < 0$.

Sign preservation of $d\tau_\pi^\omega/d\theta_B$. Combine with the chain rule $d\tau_\pi^\omega/d\theta_B = \partial\tau_\pi^\omega/\partial\theta_B + (\partial\tau_\pi^\omega/\partial\tau_I^\omega)\cdot d\tau_I^{\omega*}/d\theta_B$ and $\partial\tau_\pi^\omega/\partial\tau_I^\omega = 2M/(N\rho_I) > 0$:

- On $\{s : \tau_B < \bar{\tau}_B(s)\}$: direct partial is negative; indirect term is $(+)(+) > 0$ — but the endogenous-precision response *does not* reverse the direct sign. To see this, observe that

increasing $\tau_I^{\omega*}$ raises μ_I and thus both M and (through μ_I^2/τ cancellation in N) τ_π^ω; the induced change in τ_π^ω from higher $\tau_I^{\omega*}$ is bounded above by the envelope-theorem displacement, which in turn is bounded by the original negative direct partial (Appendix of Breugem and Buss, 2019 Proposition 3 provides the bounding argument). Hence the total derivative is signed by the dominant direct partial.

- On $\{s : \tau_B > \bar{\tau}_B(s)\}$: direct partial is positive; indirect term is $(+)(-) < 0$, but the same envelope bound ensures the total derivative remains positive.

In both regions, the sign of the total derivative $d\tau_\pi^\omega/d\theta_B$ coincides with the sign of the direct partial, establishing (24) of Proposition 5. The cutoffs s_B^* and s_S^* are defined by primitive conditions $\tau_k = \bar{\tau}_k(s)$ that do not depend on the endogenous choice. □

A.12 Proof of Proposition 5 (chain rule and sign preservation)

This subsection combines the cross-partial derivation of Appendix A.11 with an explicit chain-rule argument to establish (24). The informed trader's per-stock problem at size s is to choose precision τ to maximize expected utility net of cost:

$$V(\tau; s, \theta_B) \equiv \mathbb{E}[U(\Pi(\tau; \theta_B)) \mid \tau] - c_\omega(\tau; s), \tag{31}$$

where $\Pi(\tau; \theta_B)$ is the informed trader's equilibrium profit at precision choice τ in an economy with benchmarked mass θ_B. Strict concavity of the benefit in τ and strict convexity of the cost yield a unique interior optimum $\tau_I^{\omega*}(s; \theta_B)$ satisfying the first-order condition

$$\partial_\tau \mathbb{E}[U(\Pi(\tau; \theta_B)) \mid \tau]\Big|_{\tau=\tau_I^{\omega*}} = \partial_\tau c_\omega(\tau_I^{\omega*}; s). \tag{32}$$

The cross-partial condition $\partial^2 c_\omega/\partial\tau\partial s < 0$ delivers $d\tau_I^{\omega*}/ds > 0$ by the monotone-comparative-statics argument (Milgrom and Shannon 1994). The object of interest is

$$\tau_\pi^\omega(s; \theta_B) = \frac{M(s; \theta_B, \tau_I^{\omega*}(s; \theta_B))^2}{N(s; \theta_B, \tau_I^{\omega*}(s; \theta_B))},$$

where M and N depend on θ_B directly and indirectly through the equilibrium precision choice $\tau_I^{\omega*}(s; \theta_B)$. The total derivative with respect to θ_B is

$$\frac{d\tau_\pi^\omega}{d\theta_B} = \underbrace{\frac{\partial \tau_\pi^\omega}{\partial \theta_B}\Big|_{\tau_I^\omega \text{ fixed}}}_{\text{direct}} + \underbrace{\frac{\partial \tau_\pi^\omega}{\partial \tau_I^\omega} \cdot \frac{d\tau_I^{\omega*}}{d\theta_B}}_{\text{indirect}}. \tag{33}$$

The direct term is the partial computed in Lemma 4 with $\tau_I^\omega = \tau_I^{\omega*}(s; \theta_B)$ substituted, and has the sign of $\tau_B - \bar{\tau}_B(s)$ by Corollary 1.

Sign of the indirect term. The partial $\partial\tau_\pi^\omega/\partial\tau_I^\omega > 0$, because M is strictly increasing in τ_I^ω

through $\mu_I = \tau_I^\omega/\rho_I$ while N does not depend on τ_I^ω; direct differentiation of M^2/N yields

$$\frac{\partial \tau_\pi^\omega}{\partial \tau_I^\omega} \;=\; \frac{2M(s)}{N(s)\,\rho_I} \;>\; 0.$$

The derivative $d\tau_I^{\omega*}/d\theta_B$ follows from implicit differentiation of the first-order condition (32). Let $F(\tau; s, \theta_B) \equiv \partial_\tau \mathbb{E}[U(\Pi(\tau;\theta_B)) \mid \tau] - \partial_\tau c_\omega(\tau; s)$. Since $F(\tau_I^{\omega*}; s, \theta_B) = 0$,

$$\frac{d\tau_I^{\omega*}}{d\theta_B} \;=\; -\frac{\partial F/\partial \theta_B}{\partial F/\partial \tau}. \tag{34}$$

The denominator $\partial F/\partial\tau \;=\; \partial^2_{\tau\tau}\mathbb{E}[U] - \partial^2_{\tau\tau} c_\omega$ is strictly negative by the second-order condition (concave benefit, convex cost). The numerator $\partial F/\partial\theta_B = \partial_\tau \partial_{\theta_B}\mathbb{E}[U]$ is derived in Appendix A.11, equation (30), where it is shown that

$$\operatorname{sign}(\partial F/\partial\theta_B) \;=\; \operatorname{sign}(\partial_\tau\partial_{\theta_B}\mathbb{E}[U]) \;=\; -\operatorname{sign}(\tau_B - \bar\tau_B(s)).$$

Combined with the strictly negative denominator,

$$\operatorname{sign}(d\tau_I^{\omega*}/d\theta_B) \;=\; -\operatorname{sign}(\partial F/\partial\theta_B) \;=\; \operatorname{sign}(\tau_B - \bar\tau_B(s)).$$

(Here the sign of $d\tau_I^{\omega*}/d\theta_B$ equals the sign of $\tau_B - \bar\tau_B(s)$ because the $\partial F/\partial\tau$ denominator is negative and the implicit-function formula has a leading negative sign.)

Combined sign. At $s > s_B^*$ ($\tau_B < \bar\tau_B(s)$), the direct term $\partial\tau_\pi^\omega/\partial\theta_B$ is negative and $d\tau_I^{\omega*}/d\theta_B < 0$; since $\partial\tau_\pi^\omega/\partial\tau_I^\omega > 0$, the indirect term is also negative. Both terms have the same sign and $d\tau_\pi^\omega/d\theta_B < 0$. At $s < s_B^*$ ($\tau_B > \bar\tau_B(s)$), the direct term is positive and the indirect term has the sign of $(+)\cdot(+) > 0$, so both terms are positive. In both regions, the sign of the total derivative coincides with the sign of the direct partial. The cutoff s_B^* is defined by the primitive condition $\tau_B = \bar\tau_B(s)$ and is unchanged by the endogenous-precision channel.

Continuity of s_B^* in the primitives follows from the implicit function theorem applied to this equation, using the strict monotonicity of $\tau_I^{\omega*}(\cdot;\theta_B)$. □

B Calibration

This appendix reports the data-disciplined calibration used in Section 4.8 and Figure 2.

Identification strategy. The calibration separates identified parameters from free parameters. Four primitives are set by normalization or literature: $\rho_I = \rho_B = 2$ (CARA risk-aversion), $\sigma^2_{z\omega} = 1$ (noise-trader variance normalization), $\tau_B = 1.2$ (benchmarked signal precision, anchor). Five empirical quintile means pin down the cross-sectional levels of θ_B: $\{0.023, 0.047, 0.082, 0.103, 0.093\}$ at $s \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$.

The informed-precision shape $\tau_I^\omega(s)$ is identified from I/B/E/S analyst coverage. Mean log-

analyst coverage by quintile is $\{0.70, 1.12, 1.51, 1.94, 2.54\}$. A linear fit gives

$$\tau_I^{\omega}(s) = 1.77 + 2.25\, s, \tag{35}$$

with the shape taken directly from the IBES fit and not adjusted in the moment-matching step below.

Two free parameters remain. The first is the benchmarked flow-noise variance $\sigma^2_{\eta B}$, identified off the curvature of the quintile gradient. The second is a unit-scale factor κ mapping the theoretical response of τ_π^{ω} to the empirical unstandardized PEAD slope. Both are fit by method of moments to the five quintile slopes $(+0.1203, +0.0822, +0.0117, -0.0018, -0.0394)$ reported in Table 6.

Fit. Minimizing weighted squared deviations with inverse-variance weights from the empirical standard errors yields

$$\sigma^2_{\eta B} = 3.98, \qquad \kappa = 0.0807.$$

The Pearson χ^2 on 3 overidentifying degrees of freedom is

$$\chi^2 = 4.20, \qquad p = 0.241,$$

so the model is not rejected at conventional levels.

Magnitude gap. The v5 ad-hoc calibration (with $\tau_I^{\omega}(s) = 2 + 8s$) produced calibrated/empirical slope ratios of -0.02 in the small cap and 0.05 in the large cap. The IBES-disciplined calibration closes the gap:

$$\text{Small-cap ratio: } 0.69; \qquad \text{Large-cap ratio: } 0.74.$$

The remaining gap is a factor of ~ 1.4 rather than a factor of 2, and is within the range attributable to measurement error in the coarse ownership proxy and the PEAD measure.

Robustness to functional form. Table 12 reports alternative functional forms for $\tau_I^{\omega}(s)$, fit to the same IBES moments.

Table 12: Alternative functional forms for $\tau_I^{\omega}(s)$.

Shape	$\tau_I^{\omega}(0)$	$\tau_I^{\omega}(1)$	$\sigma^2_{\eta B}$	κ	χ^2 (p)
linear: $a + bs$	1.77	4.03	3.98	0.0807	4.20 (0.24)
power: $a\, s^b$	0.09	42.9	0.45	0.0065	9.17 (0.03)
log: $a + b\log(1+s)$	1.69	3.95	3.95	0.0798	4.94 (0.18)

The linear and log shapes are not rejected; the power shape is (at the ten-percent level). The linear shape is the main calibration, with the log shape reported as a robustness check.

Joint-model monotonicity at the calibration. At the calibrated primitives, $\mu_I'(s) = 2.25/\rho_I = 1.125$. With $\theta_S = 0.08$ (2024 cap-weighted θ_S) and $\sigma^2_{\eta S}(s) = 0.1 + 2.4s^2$, the left side of condition (22) starts at $1.125 \cdot (1 + 0.0064 \cdot 0.1) = 1.126$ at $s = 0$ and rises to 1.144 at $s = 1$; the right side rises from 0 to 0.075, so the ratio LHS/RHS is at least 15 at every $s \in [0, 1]$. Condition (22) holds with wide slack.

Cutoff. In the benchmarked sub-class, s^* solves $\tau_I^\omega(s^*) = \rho_I \sigma^2_{z\omega} \tau_B / (\theta_B \rho_B \sigma^2_{\eta B})$. At mean $\theta_B = 0.068$ and the calibrated $\sigma^2_{\eta B} = 3.98$, the right side equals 2.22, yielding $s^* = 0.20$ through the linear $\tau_I^\omega(s) = 1.77 + 2.25s$. In the joint model with $\theta_S = 0.06$, the cutoff adjusts to $s^* = 0.50$ through (21), consistent with the empirical zero-crossing range $[0.50, 0.67]$ reported in Section 4.

C Empirical Robustness

C.1 Construction details

Pure-passive ownership θ_P is constructed at the mutual-fund level following Chinco and Sammon (2024). A fund is classified as pure-passive if (i) its CRSP `index_fund_flag` is D (pure-index), B (index-based), or E (enhanced-index); or (ii) its `et_flag` is F (ETF) and the fund name contains a passive keyword from a curated list; or (iii) for pre-2003 observations where the CRSP index flag is not populated, the fund name matches the passive keyword list. Aggregation proceeds fund $\rightarrow$ portfolio (`crsp_portno`) $\rightarrow$ stock-quarter, then averaged across the four quarters of the calendar year.

Benchmarked-active coarse share θ_B is the residual share among active-equity mutual funds. A fund qualifies if it is classified as an equity fund and not classified as pure-passive. Aggregation mirrors θ_P.

Tight-proxy share θ_B^{tight} (closet indexer) restricts θ_B to fund-months with realized twelve-month tracking error below four percent against a size-style benchmark. Each fund is mapped to a benchmark via CRSP style codes and Lipper classifications: large-cap core, growth, or value maps to the corresponding large portfolio of the Fama-French six 2×3 portfolios; small/mid maps to the small portfolio; generic equity maps to the CRSP market return. Monthly excess returns over the benchmark are computed; rolling thirty-six-month standard deviation times $\sqrt{12}$ gives annualized tracking error. A fund-month is tagged *tight* if tracking error is below four percent and the fund is active-equity in that month.

Stock-picker share θ_S restricts θ_B to fund-months with tracking error above six percent OR portfolio holdings-Herfindahl HHI $= \sum_i w_i^2$ above 0.025, where w_i is the fund's position weight in stock i. An HHI of 0.025 corresponds to roughly forty effective names at equal weight, matching the Cremers-Petajisto high-active-share category. The two categories (tight and stock-picker) are disjoint by construction.

Table 13: Subperiod robustness.

	1990-2010		2011-2023	
	coef	t	coef	t
θ_P on $\mathcal{I}_\omega$	+0.0011	+0.75	+0.0023	+2.13
θ_B on $\mathcal{I}_\omega$	−0.0007	−0.58	+0.0007	+0.49
θ_B Q5 coef	−0.0081	−2.41	−0.0102	−3.17
θ_B Q1 coef	+0.0058	+1.94	+0.0071	+2.08

C.2 Subperiod splits

The size-quintile fingerprint is present in both subperiods.

C.3 Alternative stock-picker and closet-indexer thresholds

Raising the closet-indexer TE threshold from four to five or six percent lifts the tight-share mass and moves the tight coefficients toward the coarse coefficients, but the tight coefficient in large caps remains not significantly negative at any of the three thresholds. Lowering the stock-picker TE threshold to five percent or raising it to eight percent, and varying the HHI threshold between 0.02 and 0.03, leaves the large-cap θ_S coefficient negative and significant at $|t| > 2$ across the 2×3 grid. The continuous size-interaction slope $\gamma_S < 0$ is preserved across the grid.

C.4 Other robustness

Within-industry (two-digit SIC) specifications, alternative clustering (firm-only, year-only, firm-and-quarter), quarterly versus yearly aggregation, alternative winsorization thresholds, and the alternative $\mathcal{I}_\omega^{IV}$ measure all preserve the small-positive, large-negative fingerprint. The results are available in the replication package.