Journal of Financial Economics 179 (2026) 104257

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/finec

Which investors drive anomaly returns and how?[☆]

Yizhang Li [a], Stanislav Sokolinski [b],[*], Andrea Tamoni [c]

[a] Two Sigma Investments, LP., 100 Avenue of the Americas, 16th Floor, New York, NY 10013, United States of America
[b] Broad College of Business, Michigan State University, United States of America
[c] Mendoza College of Business, University of Notre Dame., 304 Mendoza College of Business, Notre Dame, IN 46556, United States of America

ARTICLE INFO

Dataset link: Replication Package for "Which Investors Drive Anomaly Returns and How?" (Original data)

JEL classification:
G11
G12
G23

Keywords:
Factor investing
Anomalies
Institutional investors
Portfolio choice

ABSTRACT

We decompose variation in anomaly returns across investor types and trading motives, linking returns to holdings within a demand system. Fundamental-based trading is the primary driver of anomalies, and both this channel and non-fundamental trading are driven disproportionately by small investors. Institution-led channels — flows and industry tilts — are secondary and concentrated in specific anomalies. These patterns favor theories with heterogeneous investors and distinct trading behaviors over pure flow- or risk-only explanations. Small investors account for a large share of the cash-flow news component of fundamentals-based trading, suggesting that biases in processing information play a role in anomaly returns.

1. Introduction

Asset pricing anomalies are return patterns that basic models, like the CAPM, fail to explain. Despite decades of research, we still lack clear evidence on who drives these patterns and how. Competing explanations — priced risk, institutional flows, and heterogeneous-investor behavior — point to different investor groups and trading motives, but separating their contributions in the data has proved difficult. We address this gap with an empirical framework that jointly estimates the effect of each investor type and trading motive on anomaly returns, allowing to distinguish among competing theories and to clarify the conditions under which each mechanism operates.

We link anomaly returns to investor holdings through a structural model in which shifts in investor demand determine equilibrium returns. The design enables counterfactuals that perturb one demand driver at a time for a given investor type, hold all else fixed, re-solve for implied returns, and record the resulting price effect. This approach yields direct and comparable estimates of how much of anomaly return is attributable to each investor type and each trading motive, allowing us to apportion total returns along both dimensions.

After constructing the investor–motive breakdown of returns, we apply a canonical variance decomposition to attribute time-series variation in anomaly returns to those channels. In the spirit of traditional approaches that separate movements driven by fundamentals from those driven by discounting, we adapt the framework to our setting: instead of cash-flow and discount-rate components, we decompose by *who trades and how*, using investor types and trading motives as the units of analysis for their contributions to return variation.[1] This decomposition is intentionally theory-neutral: rather than imposing a

☆ Nikolai Rousannov was the editor for this article. We are grateful to Shuaiyu Chen for the independent replication of the main findings in this article. We thank anonymous referees, Azi Ben-Rephael, Stefano Cassella, Fotis Grigoris, Paul Huebner, Kristy Jansen, Sophia Li, Ralph Koijen, Rob Richmond, Nick Roussanov (the editor), Petra Vokata, Julian Thimme, Christian Wagner, Sterling Yan, Motohiro Yogo, and seminar and conference participants at University of Central Florida, 6th Conference on "Recent Advances in Mutual Fund and Hedge Fund Research", University of Notre Dame, Indiana University, Northeastern University, 10th SAFE Asset Pricing Workshop (Goethe University), UT Dallas 2023 Fall Finance Conference, Helsinki Finance Summit 2023, Tilburg Finance Summit 2023, Esade Spring Workshop 2023, Adam Smith Spring 2023 Workshop, Citi Quantitative Research Conference 2023, MFA 2023, 11th MSUFCU Conference on Financial Institutions, and Rutgers Business School for valuable comments.

* Corresponding author.
E-mail addresses: yizhang.li.research@gmail.com (Y. Li), sokolins@msu.edu (S. Sokolinski), atamoni@nd.edu (A. Tamoni).

[1] For cash-flow and discount-rate decompositions of market-wide, stock-level, and anomaly returns, see, e.g., Campbell (1991), Vuolteenaho (2002), Cohen et al. (2003), Chen et al. (2013), Lochstoer and Tetlock (2020).

https://doi.org/10.1016/j.jfineco.2026.104257
Received 23 January 2025; Received in revised form 10 February 2026; Accepted 11 February 2026

specific explanation ex ante, we let the data indicate which channels are most important.

An important advantage of our approach is the head-to-head comparability across theories that make distinct predictions about investors and motives. Consider momentum: one view attributes it to underreaction by less sophisticated investors (Hong and Stein, 1999) while another links it to institutional flow-induced trading (Lou, 2012). By decomposing momentum — and other anomalies — into investor–motive components, we quantify which channels dominate in the data and where each theory fits (or fails). More broadly, the framework benchmarks competing drivers on a common scale and reveals when a force is economically material versus negligible, always through an explicit investor lens.[2]

To implement the decomposition, we estimate investor demand for U.S. equities in the spirit of Koijen and Yogo (2019), Koijen et al. (2023), but with a novel two-tier nested industry–stock structure that flexibly accommodates substitution both across and within industries, conditional on stock characteristics. We adopt a nested, rather than simple, logit to account for realistic substitution patterns — capital typically rotates toward close substitutes (Chaudhary et al., 2025) — to obtain more credible elasticity estimates (Fuchs et al., 2025) and to reconcile portfolio- and stock-level evidence (Li and Lin, 2025).[3]

The decomposition attributes anomaly-return variation to four motives — demand for stock characteristics (fundamentals-based), latent demand (non-fundamentals-based), flow-induced trading, and an industry term that captures outer-nest rotation. We examine how these motives contribute across investor types — small investors (defined below) and major 13F institutions, including mutual funds, investment advisors, banks, insurers, and pension funds — and apply this investor–motive breakdown to 47 anomaly portfolios spanning seven categories: value, momentum, profitability, investment, issuance, size, and tangibility.

Our analysis yields two main findings. First, fundamental-based trading is the primary engine of anomaly returns, explaining about 40% of variance across portfolios on average. Non-fundamental-based trading is the next meaningful layer, accounting for roughly 20%. Industry effects are more modest at about 13%, and flow-induced trading is small on average at around 3%, with meaningful contributions concentrated in a few categories. The ordering of these channels is stable and statistically tight.

Fundamentals dominate in categories such as momentum, profitability, and size; non-fundamentals matter more where fundamentals recede (notably value and investment); flows register mainly in size and issuance; and industry effects peak in value and tangibility. Dispersion across individual anomalies is economically meaningful, but it does not overturn the ranking of trading motives.

Second, investor identity is central. Small investors — defined as non-13F, non-blockholder (stake $< 5\%$), long-only owners — transmit most of the fundamental-based component (about 27% of total variance, roughly two thirds of the 40% fundamentals share) and are also the single largest contributor to the non-fundamental layer (about 7%, compared with 5% for mutual funds and 3% for banks). By contrast, institutional channels show up where expected: institutions account for most of the industry margin (about 8% for 13F investors as a whole vs. 3% for small investors, on average across anomalies), and mutual funds and investment advisors carry the flow channel.

We interpret the "small investor" category cautiously. It is not the mechanical $100\% - 13\text{F}$ residual; our construction is more granular, excluding disclosed blockholders and positions attributable to short sellers. Data limitations prevent a clean separation of households, so some non-13F institutions may remain in the residual. Even so, retail-tilt diagnostics across multiple signals (lottery-like payoffs, idiosyncratic volatility, firm age, turnover) align with the small-investor type and point toward a household interpretation. Taken together, we read the evidence as indicating that anomalies are shaped chiefly by fundamental-based trading with a meaningful non-fundamentals layer, transmitted disproportionately by small investors, while institution-led flows and industry tilts play targeted but secondary roles.[4]

We link our findings to three main explanations for anomalies. First, flow-based explanations emphasize how institutional trading responds to fund flows and shapes return patterns — momentum (Lou, 2012; Ben-David et al., 2022), size and value (Li, 2022), and broader anomalies (Akbas et al., 2015). Our results refine this view by benchmarking flows against other motives and by allocating them across investor types: where flow effects arise, they are primarily institutional — consistent with delegated rebalancing and benchmark pressure by mutual funds and investment advisors — yet on average they are modest relative to fundamental- and non-fundamental-based trading. We interpret this as flows tending to propagate or amplify pressures that originate elsewhere, while not ruling out episodes in which flows play a more prominent, localized role.

Second, risk-based explanations link anomalies to priced risk (Berk et al., 1999; Zhang, 2005; Lettau and Wachter, 2007; Santos and Veronesi, 2010; Gabaix, 2012; Kogan and Papanikolaou, 2013). In an investor-demand framework, movements in coefficients on observable characteristics translate signals about expected payoffs and risk exposures into portfolio weights (Koijen et al., 2023), so the prominence of fundamental-based trading is consistent with a risk channel. Two tensions remain for a purely risk-based account. First, most risk-based models feature a representative investor and therefore cannot speak to the distinct effects across investor types that we document. Second, a sizable non-fundamental component emerges precisely where characteristics explain less. Neither point rules out risk-based explanations, but together they suggest risk alone is unlikely to explain the full pattern across motives and investor types.

Third, heterogeneous-investor and mispricing explanations offer a natural organizing lens (De Long et al., 1990; Shleifer and Vishny, 1997; Barberis et al., 1998; Hong and Stein, 1999; Daniel et al., 2001). In these models, misinterpretation of fundamentals (over- or underreaction, coarse use of information, mistaken precision) and non-fundamental trading (sentiment, liquidity, financing frictions) generate persistent pressures that rational but constrained arbitrageurs only partially offset. Our results line up with this structure: fundamental-based trading does most of the work and is transmitted disproportionately by small investors; a meaningful non-fundamental layer also tilts toward small investors; and institution-led forces are most visible in flows and industry tilts. We remain theory-neutral ex ante, but the joint pattern is most consistent with mechanisms that combine investor heterogeneity with both fundamental and non-fundamental forces.

To tighten the connection between our evidence and existing anomaly theories, we map investor–motive channels onto cash-flow (CF) and discount-rate (DR) news, following the variance-decomposition approach of Chen et al. (2013) and Lochstoer and

[2] Most prior studies focus on a single investor type, a single motive, or a narrow set of anomalies, limiting head-to-head comparisons. Notable exceptions include Edelen et al. (2016) and McLean et al. (2022), who study multiple anomalies and investor groups. Their emphasis, however, is on which investors exploit anomalies, not on apportioning the drivers of anomalies across investor types and motives as we do here.

[3] Koijen and Yogo (2025) use a nested-logit structure in cross-country settings, and Bretscher et al. (2025) apply it to estimate a demand system in the corporate bond market.

[4] Our findings on small investors align with recent studies that highlight their influence on asset prices. Balasubramaniam et al. (2023) analyze Indian retail investors' stock holdings, while Betermier et al. (2024) show that Norwegian individual investor portfolios explain stock return cross-sections. Gabaix et al. (2022) also find that U.S. ultra-high-net-worth households help stabilize market fluctuations. Unlike these studies, our work specifically quantifies small investors' contributions to anomalies, assesses their trading motives, and compares their impact to that of large institutions.

Tetlock (2020). This analysis quantifies how specific investors and trading motives load on CF versus DR forces.

Two main findings emerge. First, we confirm the central result of Lochstoer and Tetlock (2020): CF news are the dominant source of anomaly-return variation, accounting for 62% of variance on average across categories. Our investor–motive decomposition sharpens this conclusion. Nearly half of this CF-driven variation is attributable to small investors, in line with heterogeneous-investor mechanisms in which less sophisticated traders act on (and sometimes misinterpret) firm-level signals (Barberis et al., 1998; Daniel et al., 2001). Second, only about 25% of the variance explained by the non-fundamental layer loads on DR news. This leaves room for noise-trading dynamics (De Long et al., 1990), yet most of the time variation in this layer remains CF-related. Together, these patterns reinforce a CF-centric view of anomalies and clarify who transmits CF and DR news, and through which motives.

Our results place new empirical constraints on which investors and trading motives drive anomalies. Prior work links ownership shifts to future returns — e.g., Edelen et al. (2016) and McLean et al. (2022) study who exploits predictability — but does not connect holdings to anomaly returns through a structural model or compare investors and motives head-to-head. Their conclusions, like Lewellen (2011), complement ours by suggesting that institutions exert limited influence on anomalies, a view reinforced by Davis (2024) who argues that low price elasticity among sophisticated investors curbs their ability to trade away mispricings. Haddad and Muir (2021) use cross-asset return predictability to decompose aggregate risk-premium variation and bound the shares attributable to households versus intermediaries. In contrast, we work at the anomaly level rather than the aggregate market, use a structural demand system rather than predictive regressions, and identify investor roles directly from holdings.

Our study also relates to demand-system approaches that trace return drivers from investor behavior. Haddad et al. (2024) analyze institutional reactions to passive investing, Koijen et al. (2023) show how investors shape valuation ratios, Huebner (2024) highlights momentum, and Han et al. (2022) examines mutual funds' role in low-risk anomalies. We offer a broader perspective: a unified decomposition across investors, motives, and many anomalies, rather than focusing on a single anomaly or investor group.

2. Variance decomposition of anomaly returns

2.1. A nested-logit demand system for equities

We model portfolio choice with a two-tier nested-logit demand system, akin to Koijen and Yogo (2025) in the cross-country setting. Below, we briefly outline the methodology; full details are in Appendix B.1. At the outer nest, investors allocate wealth across G industries. We use the Fama–French 5 industry classification — Manufacturing, Durables, Hi-Tech, Health Care, and a residual "Others" — validated in Appendix B.2. "Others" serves as the reference sector for normalizing outer-nest asset share equations. Conditional on industry γ, the inner nest allocates that sector's stake across stocks $n \in \gamma$, with each industry including an outside asset $n = 0 \mid \gamma$.

Let $w_{i,t}(\gamma)$ denote the share of investor i's wealth in industry γ at time t. Within that industry, the investor chooses conditional weights $w_{i,t}(n \mid \gamma)$ for stocks $n \in \gamma$. The total investor i's portfolio weight on stock n is then

$$w_{i,t}(n) = w_{i,t}(n \mid \gamma)\, w_{i,t}(\gamma)\ . \tag{1}$$

The inner nest $w_{i,t}(n \mid \gamma)$ models how an investor substitutes across stocks within an industry, while the outer nest $w_{i,t}(\gamma)$ captures how an investor substitutes across industries. This two-tier nested logit structure relaxes the homogeneous cross-asset substitution restriction imposed by the standard logit model.

Below, we describe how we model and estimate investor choice within and across industries. In the inner nest, preferences for stock characteristics are industry-specific—the coefficient vector varies with γ, capturing that the relevance of characteristics differs across sectors. In the outer nest, we estimate investor-specific, time-invariant substitution parameters governing how investors allocate wealth across industries.

Inner-Nest Estimation. Within each industry γ, we observe the conditional weights $w_{i,t}(n \mid \gamma)$, i.e. the fraction of investor i's allocation to industry γ that is invested in stock n. For example, if investor i allocates half of total wealth to Manufacturing and then distributes that amount 40–40–20 across three stocks A, B, C, our model must explain the triplet $(0.40, 0.40, 0.20)$ as the outcome of a within-industry demand system.

We assume the conditional weights follow a logit structure:

$$w_{i,t}(n \mid \gamma) = \frac{\exp\left\{b_{i,\gamma,t} + \beta_{0,i,\gamma,t}\, p_t(n) + \beta_{x,i,\gamma,t}^{\top} x_t(n) + \epsilon_{i,t}(n \mid \gamma)\right\}}{1 + \sum_{m \in \mathcal{N}_{i,\gamma}} \exp\left\{b_{i,\gamma,t} + \beta_{0,i,\gamma,t}\, p_t(m) + \beta_{x,i,\gamma,t}^{\top} x_t(m) + \epsilon_{i,t}(m \mid \gamma)\right\}}, \tag{2}$$

where $\mathcal{N}_{i,\gamma}$ is investor i's stock choice set within industry γ, $p_t(n)$ is the price of asset n, $x_t(n)$ is a vector of firm characteristics, and $\epsilon_{i,t}(n \mid \gamma)$ is an idiosyncratic shock. The "1" in the denominator captures an inside-industry outside option, representing the share of the industry budget allocated to stocks not included in the tracked universe.

Recasting the logit in the usual share-ratio form,

$$\ln \frac{w_{i,t}(n \mid \gamma)}{w_{i,t}(0 \mid \gamma)} = b_{i,\gamma,t} + \beta_{0,i,\gamma,t}\, p_t(n) + \beta_{x,i,\gamma,t}^{\top} x_t(n) + \epsilon_{i,t}(n \mid \gamma), \tag{3}$$

clarifies estimation: both the price coefficient β_0 and the characteristic-tastes β_x vary by industry, accommodating the fact that an investor may care more about, say, profitability in Utilities than in Tech.

Endogeneity of $p_t(n)$ is tackled with a within-industry analogue of the (Koijen and Yogo, 2019) "counterfactual market equity" instrument. We find the counterfactual market equity, at the market clearing price, if other investors (excluding i and stripping small investors) were to hold an equal-weighted portfolio within their *industry-specific* investment universe. This delivers a price that is orthogonal to i's own demand shock. We estimate the share-ratio equation by nonlinear GMM on a quarterly cross-section, applying this instrument to every stock in the industry.

Inclusive Value. After estimating the within-industry parameters $(\beta_{0,i,\gamma,t}, \beta_{x,i,\gamma,t})$ for investor i, we collapse the entire menu of stocks in industry γ into a single statistic,

$$\mathcal{I}_{i,t}(\gamma) = 1 + \sum_{m \in \mathcal{N}_{i,\gamma}} \exp\left\{b_{i,\gamma,t} + \beta_{0,i,\gamma,t}\, p_t(m) + \beta_{x,i,\gamma,t}^{\top} x_t(m) + \epsilon_{i,t}(m \mid \gamma)\right\}. \tag{4}$$

The "inclusive value" equals the inverse of investor i's outside portfolio weight in industry γ, $\mathcal{I}_{i,t}(\gamma) = 1/w_{i,t}(0 \mid \gamma)$. A higher $\mathcal{I}_{i,t}(\gamma)$ therefore indicates that the tracked stocks in industry γ are more attractive relative to the untracked securities within the same sector. In the outer nest, $\mathcal{I}_{i,t}(\gamma)$ serves as a sufficient statistic for industry γ's overall appeal when the investor decides how much total wealth to allocate to that industry.

Outer-Nest Estimation. Having turned each industry's menu of stocks into a single inclusive value $\mathcal{I}_{i,t}(\gamma)$, we model the share of total wealth that investor i allocates in industry γ as

$$w_{i,t}(\gamma) = \frac{\exp\{\alpha_{i,\gamma} + \eta_{i,t}(\gamma)\}\, \mathcal{I}_{i,t}(\gamma)^{\sigma_{i,\gamma}}}{\sum_{\gamma'} \exp\{\alpha_{i,\gamma'} + \eta_{i,t}(\gamma')\}\, \mathcal{I}_{i,t}(\gamma')^{\sigma_{i,\gamma'}}}, \qquad \gamma = 1, \ldots, G,$$

where $\eta_{i,t}(\gamma)$ is an industry-specific latent demand shock and $\sigma_{i,\gamma}$ governs how aggressively the investor substitutes toward industry γ when its inclusive value rises. It gauges how much a change in $\mathcal{I}_{i,t}(\gamma)$ spills over into investor i's relative demand for the *other* industries.

- If $\sigma_{i,\gamma} = 1$ for all industries γ, the system collapses to a standard logit with perfect substitution across industries: the elasticity of substitution within- and across-industry is identical, so a shock in Tech affects Utilities just as much as it affects another Tech stock.
- If $\sigma_{i,\gamma} = 0$ for all industries γ, there is no substitution across industries, meaning industries are fully segmented. The allocation across industries then becomes independent of the inclusive value, effectively eliminating substitution effects.

Because reallocating capital within an industry is typically easier than jumping across unrelated sectors, economic priors place $\sigma_{i,\gamma}$ in the open interval $(0, 1)$. Estimating $\sigma_{i,\gamma}$ therefore tells us exactly how much cross-industry rebalancing accompanies a given demand shock in the inner nest: higher values imply that shocks transmit strongly across sectors, whereas lower values indicate that most adjustment occurs inside the shocked industry itself.

Dividing the share for industry γ by the share for "Others" and taking logs yields

$$\ln \frac{w_{i,t}(\gamma)}{w_{i,t}(\text{Others})} = \alpha_{i,\gamma} - \alpha_{i,\text{Others}} + \eta_{i,t}(\gamma) - \eta_{i,t}(\text{Others}) + \sigma_{i,\gamma} \ln \mathcal{I}_{i,t}(\gamma) - \sigma_{i,\text{Others}} \ln \mathcal{I}_{i,t}(\text{Others}). \tag{5}$$

Because we cannot difference out an investor–industry fixed effect, $\alpha_{i,\gamma} - \alpha_{i,\text{Others}}$ is retained as an intercept. Eq. (5) is then estimated investor by investor on a panel that stacks investor i's portfolio weight in the four industries (each relative to "Others") over all time periods in the sample. For each regression we include four regressors $\ln \mathcal{I}_{i,t}(\gamma)$ (one per industry relative to "Others"); the remaining inclusive-value columns are masked so that, in any given row, only the relevant industry's regressor is active.

The inclusive value $\mathcal{I}_{i,t}(\gamma)$ embeds contemporaneous prices, and it is endogenous to the same shocks that drive industry-level demand. Estimating (5) would therefore bias the substitution-elasticity $\sigma_{i,\gamma}$. To remove this bias we follow the inner-loop strategy and construct an instrument by replacing the observed market equity $p_t(m)$ with its Koijen and Yogo (2019)'s instrument $\hat{p}_t(m)$.

Plugging $\hat{p}_t(m)$ into (4) yields the predicted inclusive value

$$\hat{\mathcal{I}}_{i,t}(\gamma) = 1 + \sum_{m \in \mathcal{N}_{i,\gamma}} \exp\Big\{ b_{i,\gamma,t} + \beta_{0,i,\gamma,t}\, \hat{p}_t(m) + \beta_{x,i,\gamma,t}^{\top} x_t(m) \Big\}, \tag{6}$$

which is orthogonal to the latent demand shocks by construction. We compute $\hat{\mathcal{I}}_{i,t}(\gamma)$ for each of the four industries and an analogous $\hat{\mathcal{I}}_{i,t}(\text{Others})$ for the outside option.

We then instrument both $\ln \mathcal{I}_{i,t}(\gamma)$ and $\ln \mathcal{I}_{i,t}(\text{Others})$ with their predicted counterparts. Using the same price instrument in the inner and outer nests keeps the two layers internally consistent and alleviates endogeneity concerns.

2.2. Implementation

2.2.1. Definitions and data

Stock Characteristics. We select a streamlined set of characteristics, beginning with the standard set $x_{k,t}(n)$ from Koijen and Yogo (2019): log market equity, log book equity, dividends to book equity, profitability, investment, and market beta. Given the inclusion of a short-selling sector in our model, we also add accruals as an additional characteristic, following Sloan (1996). By using this established baseline set, we ensure comparability with recent demand-based asset pricing studies (Han et al., 2022; Choi et al., 2023; Haddad et al., 2024), while the inclusion of accruals addresses the documented relationship between shorting demand and discretionary accruals ((Kolasinski et al., 2013); Mainardi (2021)). Full details on the definition of stock characteristics are provided in Appendix A.2.

Industry Choice. We adopt a five-industry outer nest because industries align with the substitution margin we observe in the data — the anomaly portfolios we study are tilted toward a subset of industries. Appendix Table B.1 shows that, under the Fama–French 10-industry classification, the mean Herfindahl–Hirschman Index (HHI) across anomaly portfolios is 22% (relative to a 10% industry-neutral benchmark). When collapsed to the FF-5 classification, Appendix Table B.2 shows that the mean HHI rises to about 31% (relative to a 20% neutral benchmark), indicating that within-industry rotation is the important margin of substitution.

We stop at five industries to preserve statistical power and ensure stability in the outer nest. Appendix Table B.3 shows that, under the FF-10 classification, per-industry cells become too thin — many investor–industry panels have only double-digit stock counts, rendering the nested-logit likelihood ill-conditioned. In contrast, Appendix Table B.4 shows that the FF-5 classification raises even the smallest cell sizes well above workable thresholds across investor types.

We also ensure that each industry has a comparable outside option using the (Koijen and Yogo, 2019) definition of outside asset, combined with the (Koijen et al., 2023) tail filter. This prevents a few sectors from absorbing nearly all substitution. This design keeps the σ estimates well-behaved and preserves their economic interpretation. In short, the FF-5 classification is the finest partition that captures meaningful substitution patterns without starving the data. A detailed discussion of the industry choice and all related results is provided in Appendix B.1.

Investors and Holdings. We construct an investor-holdings panel from standard datasets and SEC filings; Appendix A.3 provides full details. We obtain quarterly holdings of all 13F institutional filers from 1980–2022 from the Thomson Reuters Institutional Ownership (S34) database.[5] Following (Koijen and Yogo, 2019), we apply the same filters and classify 13F filers into six groups: investment advisers, mutual funds, banks, insurance companies, pension funds, and other 13F institutions (e.g., endowments, foundations, and nonfinancial corporations).

We refine the residual ownership beyond the standard "households" label used in the literature (e.g., Cohen et al. (2002); DeVault et al. (2019)), which mixes small retail investors with other non-13F holders whose behavior can differ. We follow (Schwartz-Ziv and Volkova, 2024) to identify non-13F blockholders (stakes > 5%) from Forms 13D/13G and amendments on EDGAR, and we remove overlaps across sources to avoid double counting. Because Forms 13D/13G became mandatory in 1999, we restrict the final sample to 1999–2022 so blockholder coverage aligns with our holdings. These data allow us to reassign residual ownership to blockholders where applicable.[6]

We treat short sellers as a separate investor category because their positions are linked to anomaly returns (Hanson and Sunderam, 2014) and should not inflate long-only ownership. We collect stock-level short interest from Compustat and compute the quarter-end dollar value of outstanding shorts. Rather than folding these negative positions into the long side, we exclude them from the residual sector so borrowed-and-sold shares are not attributed to small investors.

After removing non-13F blockholders, and short sellers, we define "small" investors as non-13F, non-blockholder, long-only holders. Separating short positions lowers the small-investor share and yields a conservative allocation of residual ownership.[7]

[5] We use the updated, partially regenerated S34 data with corrections noted by researchers (e.g., Ben-David et al. (2021)). See https://wrds-www.wharton.upenn.edu/documents/952/S12_and_S34_Regenerated_Data_2010-2016.pdf for details.

[6] Schwartz-Ziv and Volkova (2024) also identify corporate insiders as a separate investor group. We fold insiders into the small-investor group: in the early years (1999–2002) they hold fewer than fifty stocks per industry, well below the threshold for reliable estimation.

[7] As an illustration, in 2021Q4, blockholders account for about 5% of ownership across all U.S. stocks (3.3% within our estimation sample), while short sellers hold on average 3.5% per stock. From a 37.5% residual ownership that might otherwise be labeled "households", deducting 3.3% (blockholders) and 3.5% (shorts) leaves 30.7% assigned to small investors.

2.2.2. Demand estimation

Our estimation framework follows Koijen and Yogo (2019), Koijen et al. (2023). Each quarter, within a given investor type (e.g., banks, mutual funds, insurers), we sort accounts by the number of nonzero holdings. Investors with more than 1000 positions are estimated as stand-alone entities. Smaller accounts are pooled until each pool contains at least 2000 stock holdings. This design preserves identification for large institutions and stabilizes estimation for smaller ones.

For stand-alone investors, we estimate the nested-logit demand system at the inner level using (3). Using the fitted inner-nest coefficients, we construct inclusive values and proceed to the outer nest.

For pooled investors, we apply a two-step procedure. Step 1 estimates inner-nest logits at the group level for each pooled set. Step 2 refines these into investor-level elasticities via ridge-penalty shrinkage toward the group estimate, allowing controlled heterogeneity within investor types. With inner nests in place, we estimate the outer-nest substitution parameters $\sigma_{i,\gamma}$ via (5), instrumenting $\ln \mathcal{I}_{i,t}(\gamma)$ with predicted inclusive values (6) constructed from the counterfactual prices.

For the non-13F categories — small investors, short sellers, and blockholders — individual holdings are either unobserved (small investors, short sellers) or too sparse (blockholders). We therefore estimate demand at the group level for each category. We impose $\beta_{0,t} > 1$ for short sellers to ensure an upward-sloping demand curve and $\beta_{0,t} \le 1$ for all long-only investors. This approach captures rich heterogeneity across institutional investors in both nests, while the contributions of non-13F categories should be viewed as conservative lower bounds due to within-group aggregation.[8]

2.3. Variance decomposition of anomaly returns

2.3.1. Decomposition of stock returns under nested logit

We decompose quarterly log returns using the nested-logit price system introduced above. Let $\mathbf{p}_t$ denote the $N \times 1$ vector of equilibrium log prices, $\mathbf{s}_t$ the $N \times 1$ vector of log shares outstanding, and $\mathbf{X}_t$ the matrix of firm characteristics. Let $\mathbf{A}_t$ be the vector of investors' assets under management (AUM), $\boldsymbol{\beta}_t$ the collection of inner-nest coefficients on stock characteristics, and ϵ_t the vector of within-industry shocks. The outer nest enters through the industry block $\mathcal{J}_t \equiv (\boldsymbol{\eta}_t, \boldsymbol{\alpha}_t, \boldsymbol{\sigma}_t)$, which stacks industry-specific latent demand $\boldsymbol{\eta}_t$ and the (investor-weighted) intercepts $\boldsymbol{\alpha}_t$ and substitution elasticities $\boldsymbol{\sigma}_t$.[9] We summarize the pricing map as

$$\mathbf{p}_t = \mathbf{g}\big(\mathbf{s}_t,\, \mathbf{X}_t,\, \mathbf{A}_t,\, \boldsymbol{\beta}_t,\, \epsilon_t\,;\, \mathcal{J}_t\big). \tag{7}$$

Define log returns by

$$\mathbf{r}_t = \mathbf{p}_t - \mathbf{p}_{t-1} + \mathbf{v}_t, \tag{8}$$

where $\mathbf{v}_t = \log\big(1 + \exp(\mathbf{d}_t \oslash \mathbf{p}_t)\big)$ is the vector of dividend yields and $\oslash$ denotes elementwise division.

Two-loop counterfactual sequence. To attribute $\mathbf{p}_t - \mathbf{p}_{t-1}$ to economically distinct drivers, we sequentially update inputs in two loops that mirror the model's nested structure.

Outer loop: We first shift the entire industry block $\mathcal{J}_t$ (i.e., all $\eta_{i,t}(\gamma)$ simultaneously, along with the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\sigma}$). Because the five industries compete in a single share system, moving them jointly preserves adding-up constraints and avoids scale effects.

Inner loop: holding $\mathcal{J}_t$ fixed, we update (i) shares outstanding $\mathbf{s}_t$, (ii) characteristics $\mathbf{X}_t$, (iii) AUM $\mathbf{A}_t$, and (iv) inner-nest taste coefficients $\boldsymbol{\beta}_t$. Finally, we update (v) idiosyncratic shocks ϵ_t. At each step we solve for the counterfactual equilibrium using a Newton method adapted to the nested system; Appendix B.1.5 provides implementation details.

Decomposition. Let $\Theta_{t-1} \equiv (\mathbf{s}_{t-1}, \mathbf{X}_{t-1}, \mathbf{A}_{t-1}, \boldsymbol{\beta}_{t-1}, \epsilon_{t-1}; \mathcal{J}_{t-1})$. We write the capital gain as

$$\mathbf{p}_t - \mathbf{p}_{t-1} = \Delta\mathbf{p}_t(\mathcal{J}) + \Delta\mathbf{p}_t(\mathbf{s}) + \Delta\mathbf{p}_t(\mathbf{X}) + \Delta\mathbf{p}_t(\mathbf{A}) + \Delta\mathbf{p}_t(\boldsymbol{\beta}) + \Delta\mathbf{p}_t(\epsilon) + \boldsymbol{\psi}_t, \tag{9}$$

with components defined in the two-loop order:

$$\begin{aligned}
\Delta\mathbf{p}_t(\mathcal{J}) &= \mathbf{g}\big(\mathbf{s}_{t-1}, \mathbf{X}_{t-1}, \mathbf{A}_{t-1}, \boldsymbol{\beta}_{t-1}, \epsilon_{t-1}; \mathcal{J}_t\big) - \mathbf{g}\big(\mathbf{s}_{t-1}, \mathbf{X}_{t-1}, \mathbf{A}_{t-1}, \boldsymbol{\beta}_{t-1}, \epsilon_{t-1}; \mathcal{J}_{t-1}\big),\\
\Delta\mathbf{p}_t(\mathbf{s}) &= \mathbf{g}\big(\mathbf{s}_t, \mathbf{X}_{t-1}, \mathbf{A}_{t-1}, \boldsymbol{\beta}_{t-1}, \epsilon_{t-1}; \mathcal{J}_t\big) - \mathbf{g}\big(\mathbf{s}_{t-1}, \mathbf{X}_{t-1}, \mathbf{A}_{t-1}, \boldsymbol{\beta}_{t-1}, \epsilon_{t-1}; \mathcal{J}_t\big),\\
\Delta\mathbf{p}_t(\mathbf{X}) &= \mathbf{g}\big(\mathbf{s}_t, \mathbf{X}_t, \mathbf{A}_{t-1}, \boldsymbol{\beta}_{t-1}, \epsilon_{t-1}; \mathcal{J}_t\big) - \mathbf{g}\big(\mathbf{s}_t, \mathbf{X}_{t-1}, \mathbf{A}_{t-1}, \boldsymbol{\beta}_{t-1}, \epsilon_{t-1}; \mathcal{J}_t\big),\\
\Delta\mathbf{p}_t(\mathbf{A}) &= \mathbf{g}\big(\mathbf{s}_t, \mathbf{X}_t, \mathbf{A}_t, \boldsymbol{\beta}_{t-1}, \epsilon_{t-1}; \mathcal{J}_t\big) - \mathbf{g}\big(\mathbf{s}_t, \mathbf{X}_t, \mathbf{A}_{t-1}, \boldsymbol{\beta}_{t-1}, \epsilon_{t-1}; \mathcal{J}_t\big),\\
\Delta\mathbf{p}_t(\boldsymbol{\beta}) &= \mathbf{g}\big(\mathbf{s}_t, \mathbf{X}_t, \mathbf{A}_t, \boldsymbol{\beta}_t, \epsilon_{t-1}; \mathcal{J}_t\big) - \mathbf{g}\big(\mathbf{s}_t, \mathbf{X}_t, \mathbf{A}_t, \boldsymbol{\beta}_{t-1}, \epsilon_{t-1}; \mathcal{J}_t\big),\\
\Delta\mathbf{p}_t(\epsilon) &= \mathbf{g}\big(\mathbf{s}_t, \mathbf{X}_t, \mathbf{A}_t, \boldsymbol{\beta}_t, \epsilon_t; \mathcal{J}_t\big) - \mathbf{g}\big(\mathbf{s}_t, \mathbf{X}_t, \mathbf{A}_t, \boldsymbol{\beta}_t, \epsilon_{t-1}; \mathcal{J}_t\big),\\
\boldsymbol{\psi}_t &= \mathbf{p}_t - \mathbf{g}\big(\mathbf{s}_t, \mathbf{X}_t, \mathbf{A}_t, \boldsymbol{\beta}_t, \epsilon_t; \mathcal{J}_t\big).
\end{aligned} \tag{10}$$

Economic interpretation. The industry term $\Delta\mathbf{p}_t(\mathcal{J})$ captures the outer-nest reallocation channel. The next four terms reflect supply and within-industry demand components: $\Delta\mathbf{p}_t(\mathbf{s})$ captures changes in shares outstanding (issuances and repurchases); $\Delta\mathbf{p}_t(\mathbf{X})$ reflects changes in stock characteristics; $\Delta\mathbf{p}_t(\mathbf{A})$ captures shifts in the scale of investor demand through AUM; and $\Delta\mathbf{p}_t(\boldsymbol{\beta})$ captures changes in demand for characteristic inside industries.[10] We then update ϵ_t to capture residual within-industry shifts not explained by observable state variables. The final term, $\boldsymbol{\psi}_t$, reconciles model-implied prices with the observed ones and absorbs numerical and higher-order terms (see Appendix C.3). We refer to the demand-side terms — $\Delta\mathbf{p}_t(\mathcal{J})$, $\Delta\mathbf{p}_t(\mathbf{A})$, $\Delta\mathbf{p}_t(\boldsymbol{\beta})$, and $\Delta\mathbf{p}_t(\epsilon)$ — as trading motives.

We do not decompose changes in $\mathbf{A}_t$ into return- and flow-driven parts. Instead, we treat the total change as a single block $\Delta\mathbf{p}_t(\mathbf{A})$, interpreted as flow-induced trading. A finer decomposition would require model-dependent imputations of mechanical capital gains across the inner and outer nests. Because such mechanical gains do not involve trading, attributing all AUM variation to a flow pressure channel provides a transparent upper bound on flow effects.

This choice also avoids counting the same shock twice. The industry component $\Delta\mathbf{p}_t(\mathcal{J})$ already re-prices portfolios through cross-industry reallocations in the nested logit, and those price changes mechanically alter investors' AUM. Introducing a separate "return-driven AUM" term would therefore reassign part of the $\mathcal{J}$ shock to $\mathbf{A}$. By applying a single AUM block after the industry update, we ensure that each price movement is credited exactly once: $\Delta\mathbf{p}_t(\mathcal{J})$ captures reallocation effects across industries, and $\Delta\mathbf{p}_t(\mathbf{A})$ captures only the incremental scale effect beyond that step.

Decomposition by Investors. To attribute price movements to investor groups, we vary a single trading motive for all investors in group g, updating it from its values at $t-1$ to t, holding all other groups at their $t-1$ values. After each update, we re-solve the nested-logit equilibrium to isolate the group's contribution. For example, consider the characteristic-demand channel. Let $\boldsymbol{\beta}_{i,\cdot,t} \equiv (\boldsymbol{\beta}_{i,\gamma,t})_{\gamma=1}^{G}$ denote investor i's inner-nest preference coefficients stacked across industries. If g = **Banks** represents a set of $I_{\mathbf{Banks}}$ investors, the contribution of banks' characteristic demand is obtained by updating their coefficients from

[8] Heterogeneity within small investors likely increases the variation attributable to this group; see Betermier et al. (2017) for household heterogeneity in portfolio tilts.

[9] In the outer nest, we include three components in the industry effect: industry-specific latent demand $\boldsymbol{\eta}_t$, the intercepts $\boldsymbol{\alpha}_t$, and the substitution elasticities $\boldsymbol{\sigma}_t$. While $\boldsymbol{\alpha}$ and $\boldsymbol{\sigma}$ are estimated as time-invariant at the investor level, aggregate (investor-weighted) values can shift between $t-1$ and t when some investors enter or exit the sample. This affects the decomposition through changes in the effective intercepts and elasticities. We therefore aggregate these three pieces into one "industry" term to report the overall effect of introducing industries in the nested framework.

[10] The supply channel $\Delta\mathbf{p}_t(\mathbf{s})$ embeds the price impact of corporate actions such as stock issuance or buybacks. For interactions between firm investment and investor demand, see Choi et al. (2023).

$\boldsymbol{\beta}_{i,\cdot,t-1} \to \boldsymbol{\beta}_{i,\cdot,t}$ for all $i \in g$. For a group represented by a single aggregate investor (e.g., $g =$ **Small Investors**), we perform the analogous update $\boldsymbol{\beta}_{\textbf{Small Investors},\cdot,t-1} \to \boldsymbol{\beta}_{\textbf{Small Investors},\cdot,t}$.

More generally, we define group-level effects by switching only group g's component of the relevant block from $t-1$ to t and keeping all other groups at $t-1$:

$$\begin{aligned}
\Delta \mathbf{p}_t\left(\mathcal{J}^g\right) &= \mathbf{g}\Big(\mathbf{s}_t, \mathbf{X}_t, \mathbf{A}_{t-1}, \boldsymbol{\beta}_{t-1}, \boldsymbol{\epsilon}_{t-1} \,;\, \mathcal{J}^{(-g)} \oplus \mathcal{J}^{(g)}\Big) \\
&\quad - \mathbf{g}\Big(\mathbf{s}_t, \mathbf{X}_t, \mathbf{A}_{t-1}, \boldsymbol{\beta}_{t-1}, \boldsymbol{\epsilon}_{t-1} \,;\, \mathcal{J}_{t-1}\Big), \\
\Delta \mathbf{p}_t\left(\mathbf{A}^g\right) &= \mathbf{g}\Big(\mathbf{s}_t, \mathbf{X}_t, \mathbf{A}_{t-1}^{(-g)} \oplus \mathbf{A}_t^{(g)}, \boldsymbol{\beta}_{t-1}, \boldsymbol{\epsilon}_{t-1} \,;\, \mathcal{J}_{t-1}\Big) \\
&\quad - \mathbf{g}\Big(\mathbf{s}_t, \mathbf{X}_t, \mathbf{A}_{t-1}, \boldsymbol{\beta}_{t-1}, \boldsymbol{\epsilon}_{t-1} \,;\, \mathcal{J}_{t-1}\Big), \\
\Delta \mathbf{p}_t\left(\boldsymbol{\beta}^g\right) &= \mathbf{g}\Big(\mathbf{s}_t, \mathbf{X}_t, \mathbf{A}_{t-1}, \boldsymbol{\beta}_{t-1}^{(-g)} \oplus \boldsymbol{\beta}_t^{(g)}, \boldsymbol{\epsilon}_{t-1} \,;\, \mathcal{J}_{t-1}\Big) \\
&\quad - \mathbf{g}\Big(\mathbf{s}_t, \mathbf{X}_t, \mathbf{A}_{t-1}, \boldsymbol{\beta}_{t-1}, \boldsymbol{\epsilon}_{t-1} \,;\, \mathcal{J}_{t-1}\Big), \\
\Delta \mathbf{p}_t\left(\boldsymbol{\epsilon}^g\right) &= \mathbf{g}\Big(\mathbf{s}_t, \mathbf{X}_t, \mathbf{A}_{t-1}, \boldsymbol{\beta}_{t-1}, \boldsymbol{\epsilon}_{t-1}^{(-g)} \oplus \boldsymbol{\epsilon}_t^{(g)} \,;\, \mathcal{J}_{t-1}\Big) \\
&\quad - \mathbf{g}\Big(\mathbf{s}_t, \mathbf{X}_t, \mathbf{A}_{t-1}, \boldsymbol{\beta}_{t-1}, \boldsymbol{\epsilon}_{t-1} \,;\, \mathcal{J}_{t-1}\Big).
\end{aligned} \tag{11}$$

Here $(\cdot)^{(g)}$ and $(\cdot)^{(-g)}$ denote the motive-specific components for group g and for all other groups, respectively, and $\oplus$ combines the updated motive for group g with the unchanged motives of the remaining investors.

In all cases, we resolve the full nested equilibrium after the group-specific switch. Industry shares and inclusive values adjust endogenously through $\mathcal{J}$, and the aggregate demand elasticity for each stock — defined as the dollar-weighted sum of investor-level elasticities — updates with the implied holdings. Thus, each group-level contribution reflects both the direct effect of the targeted motive and the indirect effect operating through market-wide price sensitivity.[11]

2.3.2. Decomposition of returns on anomaly portfolios

Portfolio return decomposition. We map the nested-logit stock-level components in (10) into portfolio space. Define the set of price-change drivers as $\mathcal{K} = \{\mathcal{J}, \mathbf{A}, \boldsymbol{\beta}, \boldsymbol{\epsilon}, \mathbf{s}, \mathbf{X}, v\}$ and let $\mathcal{I}$ denote the set of investor groups.[12] The simple portfolio return induced by updating motive $k \in \mathcal{K}$ for investor $i \in \mathcal{I}$ is:

$$R_t^{(k_i)} = \sum_{n=1}^{N} w_{n,t-1}\left(e^{\Delta p_{n,t}(k_i)} - 1\right), \tag{12}$$

where $w_{n,t-1}$ is the portfolio weight of stock n at $t-1$, and $\Delta p_{n,t}(k_i)$ is the stock-level price change resulting from updating the motive-investor component (k,i) in the nested system (for v, define $\Delta p_{n,t}(v) \equiv v_{n,t}$). Let $r_t^{(k_i)} \equiv \log(1 + R_t^{(k_i)})$. The total portfolio log return is

$$r_t = \sum_{k\in\mathcal{K}} \sum_{i\in\mathcal{I}} r_t^{(k_i)} + u_t, \tag{13}$$

where u_t is the portfolio-level residual.[13]

Anomaly portfolios. We construct 47 anomaly portfolios — covering value, momentum, profitability, investment, issuance, size, and tangibility categories — starting from Freyberger et al. (2020) with modifications detailed in Appendix A.1 to align portfolio construction with quarterly 13F holdings data. Following Lochstoer and Tetlock (2020), each anomaly return is the value-weighted return of the highest characteristic quintile minus the lowest, using NYSE breakpoints each June and annual rebalancing. We report results both for individual portfolios (e.g., book-to-market and cash-to-assets within value) and for broader anomaly groups (e.g., value) by averaging across portfolios within each group.[14]

2.3.3. Time-series variance decomposition

We decompose the time-series variance of portfolio log returns into the covariances of the component-induced returns with the total return:

$$\operatorname{Var}(r_t) = \operatorname{Cov}\Big(\sum_{k\in\mathcal{K}}\sum_{i\in\mathcal{I}} r_t^{(k_i)} + u_t,\ r_t\Big) = \sum_{k\in\mathcal{K}}\sum_{i\in\mathcal{I}} \operatorname{Cov}\left(r_t^{(k_i)}, r_t\right) + \operatorname{Cov}(u_t, r_t). \tag{14}$$

We estimate the contribution of each motive-investor component (k,i) with the time-series regression

$$r_t^{(k_i)} = a_{k_i} + b_{k_i} r_t + z_{k_i,t}, \qquad b_{k_i} = \frac{\operatorname{Cov}\left(r_t^{(k_i)}, r_t\right)}{\operatorname{Var}(r_t)}. \tag{15}$$

A positive b_{k_i} indicates that the component increases return variation, while a negative b_{k_i} implies that it offsets variation. By construction, the b_{k_i} across all (k,i), together with the residual's share, sum to one, providing a complete decomposition of the variance of anomaly portfolio returns.

Decomposition by Trading Motive. For some analysis we attribute variation to motives without distinguishing investors by summing investor–motive shares. For each motive–investor pair (k,i), let b_{k_i} be the time–series share from (15). The motive–level effect is

$$b_k = \sum_i b_{k_i}. \tag{16}$$

[15] For reporting, define $r_t^{(k)} \equiv \sum_i r_t^{(k_i)}$. We group smaller channels — changes in shares outstanding ($\mathbf{s}$), changes in characteristics ($\mathbf{X}$), and dividends (v) — into "Other Effects". Incorporating these definitions into (14) yields

$$\begin{aligned}
\operatorname{Var}(r_t) = \; & \underbrace{\operatorname{Cov}\left(r_t^{(\boldsymbol{\beta})}, r_t\right)}_{\text{Demand for Characteristics}} + \underbrace{\operatorname{Cov}\left(r_t^{(\boldsymbol{\epsilon})}, r_t\right)}_{\text{Latent Demand}} \\
& + \underbrace{\operatorname{Cov}\left(r_t^{(\mathbf{A})}, r_t\right)}_{\text{Flow-Induced Trading}} + \underbrace{\operatorname{Cov}\left(r_t^{(\mathcal{J})}, r_t\right)}_{\text{Industry Effects}} \\
& + \underbrace{\operatorname{Cov}\left(r_t^{(\mathbf{s})}, r_t\right) + \operatorname{Cov}\left(r_t^{(\mathbf{X})}, r_t\right) + \operatorname{Cov}\left(r_t^{(v)}, r_t\right)}_{\text{Other Effects}} + \underbrace{\operatorname{Cov}\left(u_t, r_t\right)}_{\text{Residual}}.
\end{aligned} \tag{17}$$

3. Results

3.1. Which trading motives drive anomaly returns?

Fig. 1 displays the nested-logit variance decomposition across seven

[11] We compare the demand elasticities implied by our nested-logit demand system to the simple-logit elasticities under Koijen and Yogo (2019)'s approach in Appendix B.1.4.

[12] Formally, for investor-dependent trading motives $k \in \{\mathcal{J}, \mathbf{A}, \boldsymbol{\beta}, \boldsymbol{\epsilon}\}$, $\mathcal{I}(k)$ is the set of investor groups, while for $k \in \{\mathbf{s}, \mathbf{X}, v\}$, $\mathcal{I}(k)$ is a singleton. We slightly abuse notation for simplicity.

[13] The residual u_t bundles (i) the portfolio analog of $\boldsymbol{\psi}_t$ from (9)–(10) and (ii) the log transformation difference; see Appendix C.3.

[14] To gauge the breadth of systematic tilts, we map investors' quarterly holdings onto 100 industry-decile portfolios (value, investment, profitability). The exposures fit actual positions closely: AUM-weighted R^2 exceeds 80% for banks, insurers, pensions, and mutual funds, and reaches 95% for small investors, with only advisers and insiders near 50% (Appendix Table C.4). Market-wide fit stays above 0.73 and averages 0.80 over 2000–2022 (Appendix Figure C.1).

[15] We refer to this as a bottom-up aggregation: we first map investor-specific log price changes for motive k into simple returns at the investor level, form investor-level portfolio components, and then compute the motive-level effects. In contrast, an aggregate-first procedure would first add investor-level log shocks within each stock and only then convert the sum back to simple returns before estimating a single motive-level component. As shown in Appendix C.4, this aggregate-first construction tends to assign a larger fraction of variation to motive k and a smaller fraction to the residual. We therefore adopt the bottom-up convention throughout, as it provides a more conservative decomposition.

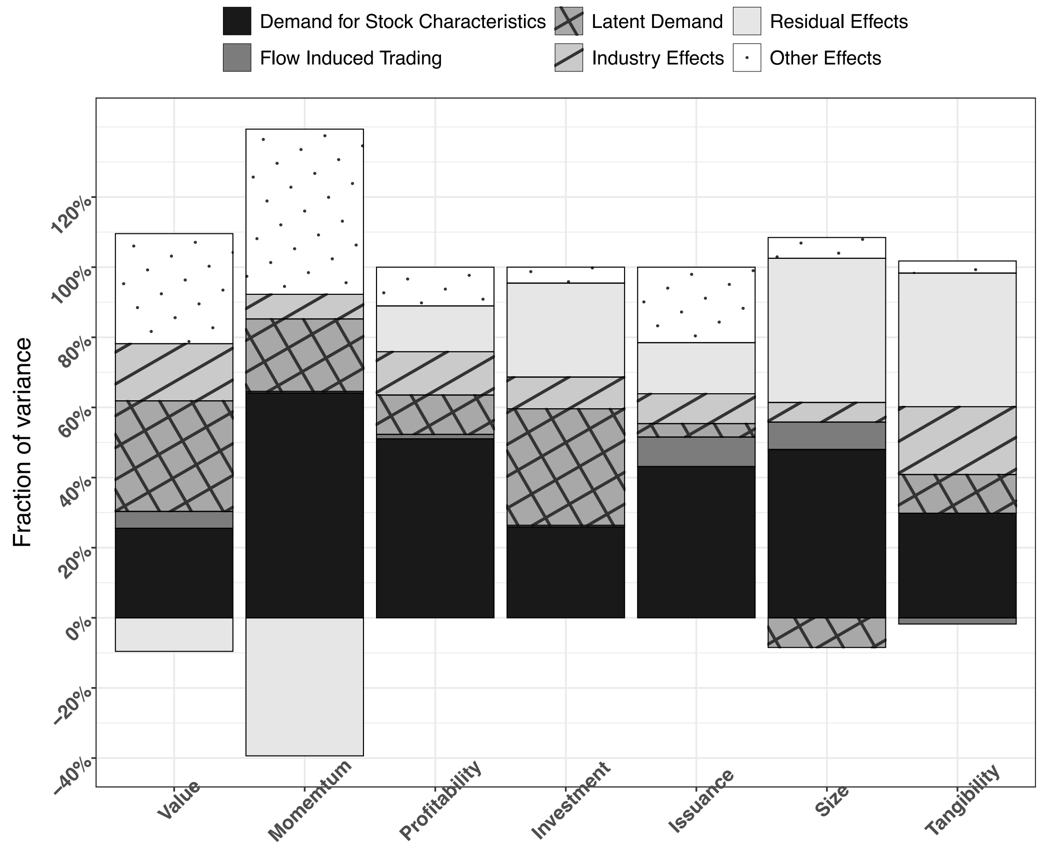


Fig. 1. Variance Decomposition by Trading Motives. Each bar represents a specific anomaly group. Appendix Table A.1 lists the anomalies included in each group. We aggregate individual anomalies into groups using the method outlined in Section 2.3.2. The segments within bars illustrate the proportions of total anomaly return variance attributed to demand for stock characteristics, latent demand, flow-induced trading, industry effects, other effects, and residual effects. "Industry effects" from the outer nest capture industry-specific latent demand, investor intercepts, and substitution elasticities. "Other Effects" encompass changes in shares outstanding, changes in stock characteristics, and dividend yield. "Residual Effects" include portfolio adjustment impacts and model error. Detailed information can be found in Section 2. The total sum of all effects is 100%.

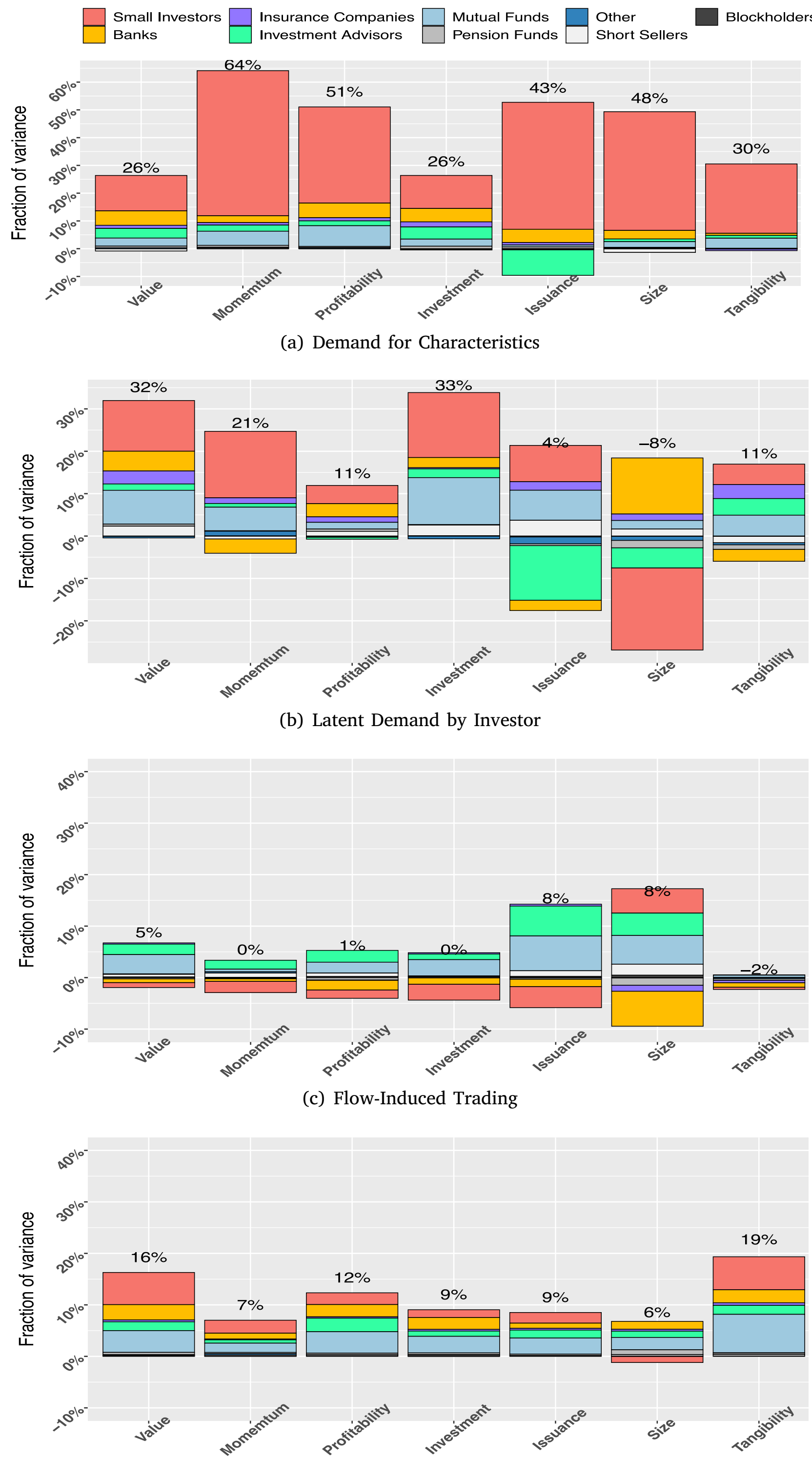
anomaly categories, with bars summing to 100%. Two forces dominate. Demand for characteristics leads in most places — most visibly for momentum (64%), and still high for profitability (51%), size (48%), and issuance (43%). Latent demand is strongest where characteristics play a smaller role, with sizable contributions in value (32%) and investment (33%). Flow-induced trading is small on average (3%) and becomes material only in issuance and size (8% each). Industry effects are secondary (mean 13%), peaking in tangibility (19%) and value (16%). A few components are negative — latent demand for size (−8%) and flows for tangibility (−2%) — so in those cases specific motives dampen rather than amplify variance.

Three patterns organize the cross-category variation. First, a chara cteristic-driven demand cluster, led by momentum and extending to profitability and size, exhibits large shares relative to other motives. Second, a latent-demand tilt emerges for value and investment, where unobserved tastes or constraints account for a greater share of the variation. Third, mixed cases underscore the influence of additional forces: tangibility and value have the largest industry footprints, while size and issuance are the only categories where flows attain a sizable magnitude, though still below that of the leading motives.

Table 1 reports the variance decomposition at the individual anomaly level, the within-category means with their statistical significance, and the overall mean, along with the 25th–75th percentile range of the variance decomposition across all anomalies. Effects are statistically significant almost everywhere, with the few exceptions concentrated in components that are close to zero or slightly negative. The statistical picture matches the visual one: characteristic-driven demand dominates broadly, latent demand matters most where characteristics recede, and flows rise to prominence in a small number of cases.

Dispersion across strategies is economically meaningful. The interquartile ranges across all anomalies are 21%–53% for demand for stock characteristics, 0%–6% for flows, -3%–36% for latent demand, and 6%–18% for industry effects. Taken together, these ranges reveal a clear hierarchy and distinct structure. Characteristic-driven and latent demand are the most variable components, with the former dominating in magnitude and the latter exhibiting uneven patterns — its importance rising where demand for characteristics effects weaken. Flow contributions are generally small, with only occasional pockets of relevance, while industry effects are steady but modest.[16]

[16] Appendix Figures C.2 and C.3 present the variance decomposition separately for the long and short legs. Three patterns emerge. First, across anomalies, latent demand is higher in the short leg than in the long or in the long-minus-short portfolio (32% vs. 26% and 15%), whereas demand for characteristics is higher in the long and long-minus-short portfolios than in the short leg (40% and 41% vs. 33%). Second, within categories, value exhibits high latent demand in both legs (45% long; 38% short) but a smaller share in the long-minus-short portfolio (32%), while size shows negative latent demand in the long leg and, to a lesser extent, in the short leg. Third, momentum displays a pronounced asymmetry: characteristic-driven demand is stronger for winners than for losers (61% vs. 22%), whereas latent demand is higher on the short side than on the long side (36% vs. 12%).



(a) Demand for Characteristics

(b) Latent Demand by Investor

(c) Flow-Induced Trading

(d) Industry Effects

Fig. 2. Variance Decomposition by Investor and Trading Motive Panel (a) shows the effects of demand for stock characteristics by investor group. Panel (b) shows the effects of latent demand by investor group. Each bar represents a specific anomaly group. Appendix Table A.1 lists the anomalies included in each group. We combine individual anomalies into groups using the method described in Section 2.3.2. The bars display the fractions of total anomaly return variance explained by the effects of each investor type. The numbers above each bar indicate the total effects of each demand component from Fig. 1.

Table 1
Variance Decomposition of Anomaly Returns.
This table reports the variance decomposition of the 47 anomaly portfolios, grouped into 7 categories. Appendix Table A.1 lists the anomalies included in each group. The estimates represent the proportions of total anomaly return variance attributed to demand for stock characteristics, flow-induced trading, latent demand, and industry effects. Detailed information can be found in Section 2. We estimate standard errors for the cross-anomaly means using a seemingly unrelated regressions (SUR) framework with Newey–West (one-lag) corrections to account for within-anomaly serial correlation. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Anomalies	Demand for stock characteristics	Flow-induced trading	Latent demand	Industry effects
A. Value				
Total assets over market equity	12%	4%	15%	19%
Book equity over market equity	17%	4%	51%	13%
Industry-adjusted book-to-market ratio	24%	6%	72%	4%
Cash and short-term investments over total assets	19%	4%	34%	18%
Cash flow to total liabilities	32%	0%	4%	12%
Log change in the shares outstanding	53%	7%	15%	15%
Debt over market equity	11%	6%	20%	17%
Income over market equity	12%	11%	48%	15%
Cash flow to book equity	70%	2%	−2%	14%
Net payout ratio	31%	7%	34%	26%
Payout ratio	32%	8%	31%	24%
Tobin’s Q	20%	3%	50%	11%
Sales over market equity	17%	3%	46%	18%
Sales growth	7%	1%	24%	22%
Mean	26%***	5%***	32%***	16%***
B. Momentum				
12-month lagged returns, held for 12 months	53%	1%	28%	6%
6-month lagged returns, held for 6 months	65%	0%	19%	9%
12-month lagged returns, held for 3 months	75%	0%	15%	6%
Mean	64%***	0%***	21%***	7%***
C. Profitability				
Sales over lagged net operating assets	44%	9%	32%	24%
Sales over lagged total assets	24%	3%	12%	11%
Differences in growth of gross margin and sales	3%	2%	72%	9%
Earnings per share	50%	6%	−3%	12%
Pre-tax income over sales	70%	3%	−7%	5%
Price-to-cost margin	26%	0%	−1%	22%
Profit margin	42%	−1%	38%	6%
Industry-adjusted profit margin	14%	6%	11%	36%
Gross profitability over book equity	116%	0%	−3%	12%
Return on net operating assets	81%	0%	−25%	16%
Return on assets	84%	−4%	−9%	1%
Return on capital	−3%	8%	45%	25%
Return on equity	79%	1%	1%	0%
Return on invested capital	77%	−8%	−5%	3%
Sales over cash and short-term investments	25%	3%	38%	17%
Sales over total assets	41%	−2%	15%	5%
Quality score from (Asness et al., 2019)	95%	−4%	−22%	5%
Mean	51%***	1%	11%	12%***
D. Investment				
% Change in total assets	29%	4%	3%	31%
% Change in book equity	5%	−3%	66%	−7%
Change in PP&E over lagged total assets	21%	5%	66%	12%
Change in inventories over average total assets	21%	3%	42%	2%
Net operating assets over lagged total assets	53%	−6%	−11%	7%
Mean	26%***	0%	33%**	9%
E. Issuance				
1-year net-share issuance	50%	8%	22%	6%
5-year composite-share issuance	36%	9%	−14%	11%
Mean	43%***	8%***	4%	9%***
F. Size				
Total assets	42%	12%	−26%	6%
Market equity	46%	2%	−17%	5%
Industry-adjusted market equity	56%	9%	18%	6%
Mean	48%***	8%***	−8%	6%***
G. Tangibility				
Absolute value of operating accruals	30%	−2%	21%	28%

(*continued on next page*)

Table 1 (*continued*).

Operating leverage	28%	0%	−9%	6%
Operating accruals	32%	−3%	21%	23%
Mean	30%***	−2%**	11%	19%***
Mean (All Anomalies)	40%***	3%**	19%***	13%***
25th percentile (All Anomalies)	21%	0%	−3%	6%
75th percentile (All Anomalies)	53%	6%	36%	18%

3.2. Which investors drive anomaly returns?

We next decompose the contribution of each trading motive to anomaly variance by investor type. Building on the previous section, we focus on the four main demand drivers — demand for characteristics, latent demand, flow-induced trading, and industry effects — which together account for 75% of the average anomaly's variance. Fig. 2 shows the investor breakdown for each motive: bar labels indicate the total motive effect (from Fig. 1), while stacked colors allocate that total across investor groups. This layout offers a clear view of who drives each motive and how the investor mix varies across anomaly categories.

3.2.1. Demand for stock characteristics by investor types

Panel A in Fig. 2 shows how the characteristic-driven demand component splits across investor types by anomaly category. Across categories, the small-investor slice consistently represents the largest share. It is most pronounced where demand for characteristics is strong — momentum, profitability, and size — and remains the leading component even when the share of characteristic-driven demand is lower, such as for value and investment. Variance contributions from banks and mutual funds are small and stable, thinner for investment advisors, and near zero for short sellers and blockholders.

The cross-category contrast is sharp. When the variance contribution from demand for stock characteristics is high, small investors account for most of the effect; even when it is low — as in value and investment — they remain the largest contributors, with institutions splitting the remainder in roughly equal proportions. Overall, demand for characteristics is predominantly a small-investor phenomenon rather than an institutional one.

Table 2 confirms this pattern with cross-anomaly means and their statistical significance. Small-investor contributions are statistically significant across categories and materially larger than those of any other group. Institutional contributions are small and often indistinguishable from zero, with only modest positive averages for banks and mutual funds and negligible averages for the remaining groups.

Dispersion across individual anomalies reinforces the dominance of small investors in characteristic-driven demand. Their interquartile range is wide (6%–44%), indicating substantial cross-anomaly variation. By contrast, banks and mutual funds are tightly clustered (2%–6% and 1%–7%, respectively), and the remaining groups sit at or near zero throughout the distribution. Taken together, both the levels and the spreads show that small investors drive the bulk of characteristic-driven demand and account for most of its variation across anomalies.

3.2.2. Latent demand by investor types

Panel B in Fig. 2 shows that latent demand is largely an investor-mix phenomenon. When latent demand is high (above 20%), small investors dominate; when it weakens, contributions become more evenly distributed across investor types. The composition therefore shifts with the strength of latent demand. Small investors' variance contributions are sizeable for value, investment, and momentum, but much smaller for profitability and tangibility. Latent demand is negative for size, driven primarily by an attenuation from small investors and investment advisors. Mutual funds contribute relatively evenly across anomalies, at about 6% on average, while banks account for most of the remaining gap.

Table 3 aligns with the graphical evidence and quantifies investor-level contributions. Small investors account for the largest share of latent-demand means in value (12%), investment (15%), and momentum (16%), followed by mutual funds (8%, 11%, and 5%, respectively). Size is the main exception: the mean contribution from small investors is negative (–19%), largely offset by banks (13%). Issuance shows only a modest latent-demand component (4%), driven by positive contributions from small investors and mutual funds (each around 8%) and a negative contribution from investment advisors (–13%). Tangibility has an 11% latent-demand contribution, with small investors and mutual funds each contributing roughly 5%. Most category means are statistically significant, and the few insignificant estimates occur precisely where total effects are small.

Dispersion is far greater across categories than within them. Latent-demand variance contributions range from –8% for size to 33% for investment, whereas within-category variation is much more limited. Across all anomalies, the interquartile ranges are 0%–15% for small investors, −−5%–7% for banks, and 1%–9% for mutual funds, with other investor groups showing similarly tight spreads centered near zero. In sum, latent demand is primarily driven by small investors in the categories where it matters most (value, investment, and momentum), while institutional investors contribute more episodically — banks in size and mutual funds in value and investment.[17]

3.2.3. Flow-induced trading by investor types

Panel C of Fig. 2 reports the flow-induced trading effects, which display a markedly different mix from characteristic and latent demand. Small investors contribute little, while institutions — especially mutual funds and investment advisors — dominate. Flow-induced trading variance is sizable mainly in issuance and size, and to a lesser extent in value. In issuance, mutual funds and advisors each account for roughly 7%; in size, each contributes about 5%; and in value, about 3% each. Category totals are near zero elsewhere, with tangibility showing a slightly negative effect. Visually, the small-investor slice is thin or negative across most categories, and whenever the bar height is nontrivial, the standout contributions belong to mutual funds and investment advisors, with mutual funds generally slightly larger.

Table 4 confirms this institutional tilt. Across anomalies, mutual funds (3%) and investment advisors (2%) are the only groups with consistently positive means, while small investors average about –1% and banks –2%. Statistical tests across categories reinforce this pattern: institutional effects from mutual funds and investment advisors are frequently significant, while the averages for small investors and most other groups are statistically indistinguishable from zero.

Dispersion is tight and reinforces the institutional channel. Interquartile ranges cluster around 1%–5% for mutual funds and 0%–4% for investment advisors, compared with -3%–2% for small investors and -3%–0% for banks; other groups sit effectively at zero. In short, flow-induced trading is modest in level, concentrated in issuance, size, and

[17] For small investors, characteristic-driven effects contribute positively to the size anomaly, while latent-demand effects contribute negatively, suggesting that fundamental-driven variation may be more important in smaller, less-followed firms, whereas non-fundamental variation is more pronounced in large stocks that are exposed to broad sentiment waves (e.g., tech or mega-cap cycles).

Table 2
The Effects of Demand for Stock Characteristics By Investor.
This table reports the decomposition of the effects of demand for stock characteristics by investor across 47 anomaly portfolios, grouped into 7 categories. The estimates represent the proportions of total anomaly return variance attributed to demand of each investor type. Detailed information can be found in Section 2. We estimate standard errors for the cross-anomaly means using a seemingly unrelated regressions (SUR) framework with Newey–West (one-lag) corrections to account for within-anomaly serial correlation. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Anomalies	Small investors	Banks	Insurance companies	Investment advisors	Mutual funds	Pension funds	Other investors	Short sellers	Blockholders
A. Value									
Total assets over market equity	7%	2%	1%	1%	2%	0%	0%	−1%	0%
Book equity over market equity	−3%	8%	1%	6%	7%	0%	0%	−2%	0%
Industry-adjusted book-to-market ratio	−10%	12%	2%	9%	11%	2%	0%	−2%	0%
Cash and short-term investments over total assets	−2%	2%	2%	10%	6%	1%	0%	0%	0%
Cash flow to total liabilities	28%	3%	1%	−3%	2%	0%	1%	0%	0%
Log change in the shares outstanding	40%	6%	1%	2%	3%	1%	0%	−1%	0%
Debt over market equity	7%	2%	1%	2%	1%	0%	0%	−1%	0%
Income over market equity	5%	3%	1%	5%	0%	0%	0%	−1%	0%
Cash flow to book equity	56%	8%	1%	−2%	5%	1%	1%	0%	0%
Net payout ratio	25%	4%	1%	1%	0%	1%	0%	0%	0%
Payout ratio	25%	4%	1%	1%	0%	1%	0%	−1%	0%
Tobin's Q	−4%	8%	2%	9%	6%	1%	1%	−3%	0%
Sales over market equity	6%	6%	0%	5%	−1%	0%	1%	0%	0%
Sales growth	0%	2%	1%	2%	1%	0%	0%	0%	0%
Mean	13%**	5%***	1%***	3%***	3%***	1%***	0%***	−1%***	0%
B. Momentum									
12-month lagged returns, held for 12 months	38%	4%	1%	3%	5%	1%	0%	0%	0%
6-month lagged returns, held for 6 months	55%	1%	1%	1%	6%	1%	0%	0%	0%
12-month lagged returns, held for 3 months	63%	2%	1%	2%	4%	1%	0%	0%	0%
Mean	52%***	2%***	1%***	2%***	5%***	1%***	0%	0%***	0%***
C. Profitability									
Sales over lagged net operating assets	25%	4%	2%	5%	7%	0%	0%	0%	0%
Sales over lagged total assets	11%	4%	2%	2%	5%	1%	0%	−1%	0%
Differences in growth of gross margin and sales	−6%	4%	0%	1%	4%	0%	0%	0%	0%
Earnings per share	37%	4%	1%	0%	6%	0%	1%	0%	0%
Pre-tax income over sales	62%	6%	1%	−5%	5%	0%	1%	1%	0%
Price-to-cost margin	27%	2%	1%	−5%	1%	1%	0%	−1%	0%
Profit margin	42%	5%	0%	−2%	−3%	−1%	0%	1%	0%
Industry-adjusted profit margin	9%	3%	1%	1%	1%	0%	0%	0%	0%
Gross profitability over book equity	56%	19%	2%	7%	30%	1%	0%	0%	0%
Return on net operating assets	47%	8%	3%	11%	12%	1%	0%	0%	0%
Return on assets	68%	5%	2%	0%	7%	1%	1%	1%	0%
Return on capital	−28%	7%	2%	12%	4%	1%	0%	−1%	0%
Return on equity	62%	5%	1%	−1%	10%	−1%	1%	1%	0%
Return on invested capital	59%	5%	1%	0%	12%	0%	1%	0%	0%
Sales over cash and short-term investments	20%	−1%	0%	2%	4%	1%	0%	0%	0%
Sales over total assets	23%	3%	0%	3%	11%	0%	0%	−1%	0%
Quality score from (Asness et al., 2019)	74%	6%	2%	−1%	11%	1%	1%	1%	0%
Mean	35%***	5%***	1%***	2%	7%***	0%**	0%***	0%	0%
D. Investment									
% Change in total assets	5%	8%	2%	7%	5%	1%	1%	0%	0%
% Change in book equity	11%	1%	0%	−3%	−4%	1%	0%	0%	0%
Change in PP&E over lagged total assets	16%	1%	3%	6%	−4%	0%	0%	0%	0%
Change in inventories over average total assets	1%	7%	3%	3%	7%	2%	−1%	−1%	0%
Net operating assets over lagged total assets	27%	7%	1%	8%	9%	1%	0%	0%	0%
Mean	12%**	5%***	2%**	4%**	3%	1%***	0%	0%*	0%
E. Issuance									
1-year net-share issuance	39%	5%	1%	2%	2%	1%	0%	−1%	0%
5-year composite-share issuance	52%	5%	1%	−21%	−1%	0%	0%	0%	0%
Mean	46%***	5%***	1%**	−10%	1%	1%	0%	−1%	0%
F. Size									
Total assets	39%	4%	−1%	−1%	1%	0%	0%	0%	0%
Market equity	39%	1%	0%	4%	3%	0%	0%	0%	0%
Industry-adjusted market equity	51%	5%	0%	0%	3%	0%	1%	−3%	0%
Mean	43%***	3%***	0%	1%	2%***	0%	0%	−1%	0%
G. Tangibility									
Absolute value of operating accruals	33%	1%	0%	−2%	−2%	0%	0%	0%	0%
Operating leverage	3%	3%	−1%	7%	17%	0%	0%	−1%	0%

(*continued on next page*)

Table 2 (*continued*).

Operating accruals	39%	−2%	0%	−2%	−3%	0%	0%	1%	0%
Mean	25%**	1%	0%*	1%	4%	0%	0%	0%	0%
Mean (All Anomalies)	27%***	5%***	1%***	2%**	5%***	1%***	0%***	0%**	0%
25th percentile (All Anomalies)	6%	2%	0%	−1%	1%	0%	0%	−1%	0%
75th percentile (All Anomalies)	44%	6%	2%	5%	7%	1%	1%	0%	0%

Table 3
The Effects of Latent Demand By Investor.
This table reports the decomposition of the latent demand by investor across 47 anomaly portfolios, grouped into 7 categories. The estimates represent the proportions of total anomaly return variance attributed to demand of each investor type. Detailed information can be found in Section 2. We estimate standard errors for the cross-anomaly means using a seemingly unrelated regressions (SUR) framework with Newey–West (one-lag) corrections to account for within-anomaly serial correlation. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Anomalies	Small investors	Banks	Insurance companies	Invest-ment advisors	Mutual funds	Pension funds	Other investors	Short sellers	Block-holders
A. Value									
Total assets over market equity	3%	7%	1%	0%	2%	1%	0%	2%	0%
Book equity over market equity	20%	8%	4%	3%	10%	2%	0%	4%	0%
Industry-adjusted book-to-market ratio	52%	1%	4%	2%	9%	1%	1%	1%	0%
Cash and short-term investments over total assets	3%	11%	3%	6%	7%	1%	1%	2%	0%
Cash flow to total liabilities	17%	−8%	0%	−3%	−1%	0%	−1%	0%	0%
Log change in the shares outstanding	6%	4%	4%	−14%	12%	−1%	−2%	6%	0%
Debt over market equity	0%	7%	2%	4%	4%	1%	0%	2%	0%
Income over market equity	5%	13%	3%	8%	16%	1%	0%	3%	0%
Cash flow to book equity	1%	−5%	5%	−1%	−1%	0%	−2%	1%	0%
Net payout ratio	15%	6%	2%	−2%	12%	0%	0%	2%	0%
Payout ratio	8%	4%	3%	3%	14%	0%	−1%	1%	0%
Tobin’s Q	20%	7%	3%	5%	9%	1%	1%	4%	0%
Sales over market equity	12%	6%	3%	10%	11%	0%	0%	4%	0%
Sales growth	5%	4%	7%	−2%	8%	1%	0%	2%	0%
Mean	12%***	5%***	3%***	1%	8%***	1%**	0%	2%***	0%***
B. Momentum									
12-month lagged returns, held for 12 months	16%	0%	1%	3%	8%	0%	1%	0%	0%
6-month lagged returns, held for 6 months	19%	−7%	2%	2%	2%	0%	1%	−1%	0%
12-month lagged returns, held for 3 months	12%	−3%	1%	−2%	6%	0%	1%	−1%	0%
Mean	16%***	−3%	1%**	1%	5%**	0%	1%***	−1%**	0%***
C. Profitability									
Sales over lagged net operating assets	12%	8%	−3%	3%	12%	1%	0%	0%	0%
Sales over lagged total assets	14%	6%	−2%	−3%	−1%	0%	0%	−1%	0%
Differences in growth of gross margin and sales	49%	15%	2%	1%	1%	0%	1%	2%	0%
Earnings per share	1%	−2%	3%	−4%	−2%	0%	−1%	2%	0%
Pre-tax income over sales	1%	−11%	6%	−3%	0%	0%	−2%	2%	0%
Price-to-cost margin	−43%	27%	−2%	9%	2%	1%	1%	4%	0%
Profit margin	7%	22%	6%	−1%	2%	0%	0%	1%	0%
Industry-adjusted profit margin	−2%	6%	1%	0%	5%	1%	0%	0%	0%
Gross profitability over book equity	−2%	1%	5%	0%	−9%	1%	1%	0%	0%
Return on net operating assets	−7%	−18%	−7%	−1%	5%	1%	0%	0%	0%
Return on assets	0%	−8%	2%	−5%	2%	0%	−1%	2%	0%
Return on capital	16%	6%	5%	0%	12%	2%	1%	3%	0%
Return on equity	7%	−11%	4%	4%	−3%	0%	−2%	3%	0%
Return on invested capital	4%	−2%	2%	−6%	−2%	0%	0%	0%	0%
Sales over cash and short-term investments	5%	15%	0%	10%	8%	0%	0%	0%	0%
Sales over total assets	11%	6%	−2%	2%	0%	0%	0%	−1%	0%
Quality score from (Asness et al., 2019)	0%	−7%	2%	−14%	−4%	0%	−1%	2%	0%
Mean	4%	3%	1%	0%	2%	1%***	0%	1%***	0%***
D. Investment									
% Change in total assets	−6%	−5%	6%	−1%	8%	0%	−1%	3%	0%
% Change in book equity	25%	9%	2%	−5%	30%	1%	0%	4%	0%
Change in PP&E over lagged total assets	52%	−2%	3%	10%	4%	−1%	−2%	2%	0%
Change in inventories over average total assets	6%	16%	−2%	7%	9%	1%	1%	4%	0%
Net operating assets over lagged total assets	−2%	−6%	−7%	0%	3%	0%	−1%	1%	0%
Mean	15%	2%	0%	2%	11%**	0%	−1%	3%***	0%*
E. Issuance									
1-year net-share issuance	4%	7%	5%	−15%	16%	0%	−3%	7%	0%
5-year composite-share issuance	13%	−12%	−1%	−11%	−2%	−1%	0%	0%	0%

(*continued on next page*)

Table 3 (*continued*).

Mean	8%*	−2%	2%	−13%***	7%	0%	−2%	4%	0%***
F. Size									
Total assets	−12%	−8%	7%	−13%	2%	−1%	−3%	2%	0%
Market equity	−20%	4%	−1%	−2%	3%	−1%	0%	0%	0%
Industry-adjusted market equity	−25%	43%	−1%	0%	1%	−3%	0%	3%	0%
Mean	−19%***	13%	2%	−5%	2%***	−2%***	−1%	2%*	0%**
G. Tangibility									
Absolute value of operating accruals	15%	−5%	6%	2%	6%	−1%	−1%	−2%	0%
Operating leverage	−16%	−2%	−1%	10%	1%	0%	0%	−1%	0%
Operating accruals	15%	−1%	5%	−1%	8%	−2%	−1%	−2%	0%
Mean	5%	−3%**	3%	4%	5%***	−1%	0%***	−2%***	0%
Mean (All Anomalies)	7%**	3%**	2%***	0%	5%***	0%	0%**	1%***	0%***
25th percentile (All Anomalies)	0%	−5%	0%	−2%	1%	0%	−1%	0%	0%
75th percentile (All Anomalies)	15%	7%	4%	3%	9%	1%	0%	3%	0%

value, and transmitted primarily through mutual funds and investment advisors rather than small investors.[18]

3.2.4. Industry effects by investor types

Panel D of Fig. 2 shows that industry effects are largely driven by institutions. Mutual funds make sizable contributions across anomalies — about 4% on average — followed by investment advisors and banks at roughly 2% each, providing steady support. The industry component is material in categories such as value and tangibility (16% and 19%) and smaller in momentum and size (7% and 6%). Small-investor contributions are generally thin but become nontrivial in certain anomalies — reaching about 6% for value and tangibility.

Table 5 confirms this split. Cross-anomaly means for mutual funds, banks, investment advisors, and small investors are statistically different from zero, whereas pension funds, short sellers, blockholders, and the residual group hover near zero. Dispersion is limited relative to other motives: interquartile ranges cluster around 3%–6% for mutual funds, 1%–4% for banks, 1%–3% for advisors, and −2%–7% for small investors. In short, industry effects are smaller than characteristic or latent-demand forces, concentrated in a few categories, and transmitted chiefly by institutions, with small investors playing a secondary role.[19]

3.3. Robustness

Why small investors matter. Small investors' dominance is not mechanical. They control under 32% of anomaly-portfolio dollars and hold only slightly more names than institutions, yet their demand explains most of the variance. This gap between footprint and influence reflects differences in trading behavior rather than sheer size or breadth (Appendix Section C.1). Taken together, these diagnostics indicate that the small-investor effect reflects genuine demand differences, not a byproduct of broader holdings. In our framework, demand for characteristics is where this difference shows up most clearly: small investors place more weight on observable signals, and the resulting demand is strong enough to dominate anomalies' variance despite their modest capital share.

Varying market conditions. State dependence does not overturn the core message. Splitting the sample by volatility (Appendix Table C.5), small investors remain the leading force in both regimes: their variance share dominates in calm periods (52%) and is roughly on par with mutual funds in high-volatility quarters (about 30% each). In general, mutual funds and advisors gain influence when uncertainty is elevated. Demand for characteristics strengthens in turbulent markets (about 46% vs. 29%), and both latent and industry components also increase (about 38% vs. 32% and 13% vs. 6%).

Partitioning by valuation spreads (Appendix Table C.6) tells a similar story. Small investors remain the leading force but are more influential when spreads are low (about 40%) than when spreads are high (about 33%); mutual funds and advisors together gain ground when spreads widen (about 39% vs. 18%). Demand for characteristics stays prominent across spread regimes, while latent demand rises sharply in high-spread quarters.

In sum, small investors and characteristic-driven demand consistently play a large and relatively stable role, whereas institutional activity and latent demand vary with market conditions.

Substitution margin and the nested-logit design. We use reduced-form flow-multiplier regressions à la Chaudhary et al. (2025) as an alternative test of the importance of within-industry reallocation (Appendix Section B.2). At the single-stock level, the baseline own multiplier is small (about 0.18) and declines only slightly once we absorb five Fama–French industries, with little further attenuation at finer industry cuts. For our anomaly portfolios, the baseline own multiplier is large (about 2.5) but falls to roughly 1.6 after partialing out exposure-weighted five-industry returns. The implied industry pass-through is strong, and moving to 10 or 49 industries adds little. This evidence indicates that most substitution occurs within the broad five-industry categories. Including style substitutes barely nudges the estimates, and IV–2SLS confirms a sizeable, precisely estimated industry channel.

These results point to meaningful heterogeneity in substitutability across industries. The distinct substitute clusters validate the five-industry nested-logit structure as the appropriate level of aggregation: the outer nest captures heterogeneous demand elasticities across industries, while the inner tier model demand for individual names within industries, where more homogeneous substitutability across securities is likely to hold.

Factor-asset robustness. We also estimate the (non-nested) Koijen and Yogo (2019) system on "factor assets" formed by interacting ten Fama–French industries with book-to-market deciles — and, in extensions, with investment and profitability deciles — treating them as quasi-fixed assets (Appendix Section B.3). The resulting variance decompositions mirror the stock-level findings: demand for characteristics, along with a smaller latent component, continues to explain meaningful shares, while flow effects remain secondary. Because the

[18] Flow-induced trading by small investors can be interpreted as the effect of proportional allocation across all the stocks in an investor's portfolio. For example, suppose an investor holds $100 in stocks at time $t-1$, with $30 in stock A and $70 in stock B. Assume that between time $t-1$ and t, the investor increases their total stock allocation to $110, resulting in a $10 "flow" into the portfolio. The effects of these flows on prices are computed exactly as for institutional investors. In this example, stock A experiences a flow of $3, and stock B experiences a flow of $7.

[19] We re-estimated standard errors for all cross-anomaly means using an alternative seemingly unrelated regression (SUR) specification that permits arbitrary cross-anomaly correlation in addition to within-anomaly serial correlation. The key results remain statistically significant at the 5% level or below: (i) small-investors drive demand for stock characteristics, (ii) flow-induced trading is driven by mutual funds and investment advisors, and (iii) industry demand is driven by banks, investment advisors, and mutual funds.

Table 4
The Effects of Flow-Induced Trading By Investor.
This table reports the decomposition of the effects of flow-induced trading by investor across 47 anomaly portfolios, grouped into 7 categories. The estimates represent the proportions of total anomaly return variance attributed to flow-induced trading of each investor type. Detailed information can be found in Section 2. We estimate standard errors for the cross-anomaly means using a seemingly unrelated regressions (SUR) framework with Newey–West (one-lag) corrections to account for within-anomaly serial correlation. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Anomalies	Small investors	Banks	Insurance companies	Investment advisors	Mutual funds	Pension funds	Other investors	Short sellers	Block-holders
A. Value									
Total assets over market equity	4%	−1%	0%	0%	0%	0%	0%	0%	0%
Book equity over market equity	−2%	0%	0%	2%	4%	0%	0%	0%	0%
Industry-adjusted book-to-market ratio	0%	−1%	0%	1%	5%	0%	0%	0%	0%
Cash and short-term investments over total assets	−1%	3%	1%	0%	3%	0%	0%	−1%	0%
Cash flow to total liabilities	2%	−1%	0%	0%	−1%	0%	0%	1%	0%
Log change in the shares outstanding	−4%	−1%	0%	5%	5%	0%	0%	1%	0%
Debt over market equity	7%	−1%	0%	0%	0%	0%	0%	0%	0%
Income over market equity	4%	1%	0%	1%	4%	0%	0%	0%	0%
Cash flow to book equity	−11%	−4%	0%	8%	9%	−1%	0%	1%	0%
Net payout ratio	−2%	−1%	0%	4%	6%	0%	0%	0%	0%
Payout ratio	3%	−2%	0%	3%	4%	0%	0%	0%	0%
Tobin's Q	−3%	−1%	0%	1%	5%	0%	0%	1%	0%
Sales over market equity	−5%	0%	0%	1%	5%	0%	0%	2%	0%
Sales growth	−3%	0%	1%	1%	3%	0%	0%	0%	0%
Mean	−1%	−1%	0%***	2%***	4%***	0%***	0%***	0%**	0%***
B. Momentum									
12-month lagged returns, held for 12 months	−3%	0%	0%	1%	1%	0%	0%	1%	0%
6-month lagged returns, held for 6 months	−2%	−1%	0%	2%	1%	0%	0%	1%	0%
12-month lagged returns, held for 3 months	−2%	−1%	0%	2%	0%	0%	0%	1%	0%
Mean	−2%***	−1%**	0%**	2%***	1%*	0%***	0%***	1%***	0%*
C. Profitability									
Sales over lagged net operating assets	3%	0%	1%	3%	3%	0%	0%	−1%	0%
Sales over lagged total assets	3%	−1%	0%	0%	1%	0%	0%	0%	0%
Differences in growth of gross margin and sales	2%	0%	0%	0%	−1%	0%	0%	0%	0%
Earnings per share	−6%	−3%	0%	6%	8%	0%	0%	1%	1%
Pre-tax income over sales	−11%	−3%	0%	7%	7%	−1%	0%	3%	0%
Price-to-cost margin	−2%	−1%	0%	1%	1%	0%	0%	1%	0%
Profit margin	−8%	−3%	0%	5%	5%	−1%	0%	1%	0%
Industry-adjusted profit margin	1%	1%	0%	0%	3%	0%	0%	0%	0%
Gross profitability over book equity	5%	−3%	−1%	1%	−1%	−1%	0%	0%	0%
Return on net operating assets	2%	−3%	0%	1%	−1%	0%	0%	0%	0%
Return on assets	−8%	−3%	0%	3%	3%	−1%	0%	2%	0%
Return on capital	4%	2%	0%	0%	2%	0%	0%	0%	0%
Return on equity	−10%	−4%	−1%	6%	7%	−1%	0%	2%	0%
Return on invested capital	−2%	−5%	0%	1%	−2%	−1%	0%	1%	0%
Sales over cash and short-term investments	3%	1%	0%	0%	0%	0%	0%	−1%	0%
Sales over total assets	0%	−1%	0%	0%	−1%	0%	0%	0%	0%
Quality score from (Asness et al., 2019)	−5%	−5%	0%	4%	0%	−1%	0%	3%	0%
Mean	−2%	−2%**	0%	2%***	2%***	0%***	0%	1%**	0%***
D. Investment									
% Change in total assets	−1%	−1%	0%	1%	3%	0%	0%	0%	0%
% Change in book equity	−3%	−1%	0%	−1%	2%	0%	0%	0%	0%
Change in PP&E over lagged total assets	−5%	0%	1%	3%	5%	0%	0%	0%	0%
Change in inventories over average total assets	−3%	0%	0%	2%	3%	0%	0%	0%	0%
Net operating assets over lagged total assets	−3%	−5%	0%	0%	2%	−1%	0%	0%	0%
Mean	−3%***	−1%	0%	1%	3%**	0%	0%	0%	0%***
E. Issuance									
1-year net-share issuance	−1%	−1%	0%	5%	4%	0%	0%	1%	0%
5-year composite-share issuance	−7%	−1%	1%	7%	10%	0%	0%	1%	0%
Mean	−4%	−1%***	0%	6%***	7%**	0%**	0%	1%*	0%***
F. Size									
Total assets	2%	−7%	−1%	8%	9%	−1%	0%	1%	1%
Market equity	11%	−7%	−1%	0%	0%	−2%	0%	1%	0%
Industry-adjusted market equity	1%	−7%	−1%	5%	7%	−1%	0%	4%	0%
Mean	5%	−7%***	−1%***	4%	5%**	−1%***	0%	2%**	0%***
G. Tangibility									
Absolute value of operating accruals	−3%	−2%	−1%	1%	1%	0%	0%	1%	0%

(*continued on next page*)

Table 4 (*continued*).

Operating leverage	1%	3%	−1%	7%	17%	0%	0%	−1%	0%
Operating accruals	1%	−2%	−1%	−1%	0%	0%	0%	0%	0%
Mean	0%	−1%	−1%	0%	0%	0%	0%***	0%	0%
Mean (All Anomalies)	−1%	−2%***	0%	2%***	3%***	0%***	0%	1%***	0%***
25th percentile (All Anomalies)	−3%	−3%	0%	0%	1%	0%	0%	−1%	0%
75th percentile (All Anomalies)	2%	0%	0%	4%	5%	0%	0%	1%	0%

Table 5
The Effects of Industry Demand By Investor.
This table reports the decomposition of the effects of industry demand by investor across 47 anomaly portfolios, grouped into 7 categories. The estimates represent the proportions of total anomaly return variance attributed to demand of each investor type. Detailed information can be found in Section 2. We estimate standard errors for the cross-anomaly means using a seemingly unrelated regressions (SUR) framework with Newey–West (one-lag) corrections to account for within-anomaly serial correlation. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Anomalies	Small investors	Banks	Insurance companies	Invest-ment advisors	Mutual funds	Pension funds	Other investors	Short sellers	Block-holders
A. Value									
Total assets over market equity	5%	4%	0%	2%	6%	1%	0%	0%	0%
Book equity over market equity	6%	3%	0%	0%	3%	0%	0%	0%	0%
Industry-adjusted book-to-market ratio	1%	1%	0%	0%	1%	0%	0%	0%	0%
Cash and short-term investments over total assets	8%	4%	0%	2%	3%	0%	0%	0%	0%
Cash flow to total liabilities	1%	4%	0%	2%	5%	0%	0%	0%	0%
Log change in the shares outstanding	4%	3%	1%	2%	6%	0%	0%	0%	0%
Debt over market equity	3%	4%	0%	2%	6%	1%	0%	0%	0%
Income over market equity	6%	3%	0%	1%	3%	0%	0%	0%	0%
Cash flow to book equity	6%	1%	0%	3%	3%	0%	0%	0%	0%
Net payout ratio	13%	4%	0%	3%	6%	1%	0%	0%	0%
Payout ratio	11%	3%	0%	3%	6%	1%	0%	0%	0%
Tobin's Q	5%	3%	0%	−1%	3%	0%	0%	0%	0%
Sales over market equity	7%	2%	1%	2%	6%	0%	0%	0%	0%
Sales growth	11%	4%	0%	2%	4%	0%	0%	0%	0%
Mean	6%***	3%***	0%***	2%***	4%***	0%***	0%***	0%***	0%***
B. Momentum									
12-month lagged returns, held for 12 months	1%	1%	0%	1%	2%	0%	0%	0%	0%
6-month lagged returns, held for 6 months	4%	1%	0%	0%	2%	0%	0%	0%	0%
12-month lagged returns, held for 3 months	2%	1%	0%	1%	2%	0%	0%	0%	0%
Mean	2%***	1%***	0%***	1%***	2%***	0%***	0%***	0%***	0%***
C. Profitability									
Sales over lagged net operating assets	14%	2%	0%	2%	4%	0%	0%	0%	0%
Sales over lagged total assets	−1%	4%	0%	3%	4%	0%	0%	0%	0%
Differences in growth of gross margin and sales	0%	3%	0%	3%	3%	0%	0%	0%	0%
Earnings per share	3%	2%	0%	3%	4%	0%	0%	0%	0%
Pre-tax income over sales	−3%	0%	0%	4%	3%	0%	0%	0%	0%
Price-to-cost margin	7%	4%	1%	3%	7%	0%	0%	0%	0%
Profit margin	−4%	1%	0%	5%	3%	0%	0%	0%	0%
Industry-adjusted profit margin	18%	5%	1%	2%	10%	1%	0%	0%	0%
Gross profitability over book equity	−2%	4%	0%	4%	4%	0%	0%	1%	0%
Return on net operating assets	8%	2%	0%	2%	6%	0%	0%	0%	0%
Return on assets	−5%	1%	0%	2%	6%	0%	0%	0%	0%
Return on capital	13%	5%	0%	1%	4%	1%	0%	0%	0%
Return on equity	−6%	1%	0%	2%	3%	0%	0%	0%	0%
Return on invested capital	−3%	1%	0%	2%	3%	0%	0%	0%	0%
Sales over cash and short-term investments	7%	4%	1%	2%	3%	0%	0%	0%	0%
Sales over total assets	−5%	2%	0%	3%	3%	0%	0%	0%	0%
Quality score from (Asness et al., 2019)	−2%	1%	0%	3%	2%	0%	0%	0%	0%
Mean	1%	2%***	0%***	2%***	3%***	0%***	0%	0%***	0%***
D. Investment									
% Change in total assets	20%	3%	1%	1%	6%	0%	0%	0%	0%
% Change in book equity	−8%	1%	0%	1%	0%	0%	0%	0%	0%
Change in PP&E over lagged total assets	2%	4%	0%	1%	5%	0%	0%	0%	0%
Change in inventories over average total assets	−4%	2%	0%	0%	3%	0%	0%	0%	0%
Net operating assets over lagged total assets	−2%	3%	0%	2%	3%	0%	0%	0%	0%
Mean	2%	3%***	0%**	1%***	3%***	0%***	0%	0%***	0%
E. Issuance									
1-year net-share issuance	−1%	1%	0%	1%	4%	0%	0%	0%	0%
5-year composite-share issuance	5%	1%	0%	2%	3%	0%	0%	0%	0%

(continued on next page)

Table 5 (*continued*).

Mean	2%	1%***	0%**	2%***	4%***	0%***	0%	0%	0%
F. Size									
Total assets	−1%	2%	0%	3%	2%	1%	0%	0%	0%
Market equity	−1%	1%	0%	2%	2%	1%	0%	0%	0%
Industry-adjusted market equity	−1%	1%	1%	−1%	4%	1%	0%	1%	0%
Mean	−1%***	2%***	0%**	1%	3%***	1%***	0%	0%	0%**
G. Tangibility									
Absolute value of operating accruals	14%	3%	0%	2%	8%	0%	0%	0%	0%
Operating leverage	−7%	2%	1%	4%	6%	0%	0%	1%	0%
Operating accruals	11%	3%	1%	0%	8%	0%	0%	1%	0%
Mean	6%	3%**	1%***	2%	7%***	0%***	0%	0%***	0%***
Mean (All Anomalies)	3%***	2%***	0%***	2%***	4%***	0%***	0%***	0%***	0%***
25th percentile (All Anomalies)	−2%	1%	0%	1%	3%	0%	0%	0%	0%
75th percentile (All Anomalies)	7%	4%	0%	3%	6%	0%	0%	0%	0%

asset definition embeds both characteristic and industry rotation, this factor-asset exercise is a stringent check that suggests our stock-level nested-logit design is not driving the results.[20]

Across these robustness exercises, the relevance of small investors, the stability of the ranking of motives, and the centrality of within-industry homogeneous substitution all prove robust to alternative market states, asset definitions, and specification choices.

4. Interpretation and connection to anomaly theories

We distill the results into a small set of empirical facts about trading motives and investor types and use them to interpret the sources of anomaly returns. We then relate these facts to risk-, flow-, and heterogeneous-investor explanations and, where useful, combine our nested-logit decomposition with the cash-flow vs. discount-rate (CF–DR) news breakdown to map investor–motive channels onto theory-relevant mechanisms.

4.1. Interpretation of key estimates

Interpretation of Trading Motives. Within our framework, stock characteristics are observable fundamentals that guide portfolio choice. We therefore label changes in portfolio weights traced to the estimated coefficients on characteristics as *fundamental-based trading*. In line with Koijen et al. (2023), these coefficients reflect how investors translate signals about expected payoffs and risk into demand — potentially mixing both channels — and we remain agnostic about whether the underlying beliefs are fully rational or partly behavioral. Our estimates capture the strength of this mapping from fundamentals to portfolios rather than a sharp decomposition of investors' beliefs about profitability or risk exposures.

By contrast, latent demand is the component of demand not explained by contemporaneous characteristics, which we interpret as *non-fundamental-based trading*. This residual can reflect investor-specific or marketwide sentiment, liquidity or financing shocks, or plain noise. In the cross-section, this interpretation is most transparent because latent demand is constructed to be orthogonal to current characteristics. In the time series, two caveats motivate caution. First, latent demand may contain information about *future* fundamentals — public or private — not yet captured by contemporaneous characteristics. Second, latent demand may proxy for *past* fundamentals if investors adjust slowly or rely on lagged signals. In short, non-fundamental-based trading is a useful descriptive label in our setting, but we do not assume it is entirely information-free.

We address these concerns with machine-learning forecasting exercises that nest latent demand alongside characteristics when predicting subsequent fundamentals (Appendix C.7). Across specifications, adding latent demand does not meaningfully improve out-of-sample predictive accuracy for standard fundamentals; modest gains appear in a few cases (e.g., dividend-to-book), but accuracy does not improve for market beta, operating profitability, or asset growth and often declines for accruals. The reverse experiment tells the same story: a broad set of historical fundamentals has little power to predict latent demand.

Taken together, these tests suggest that the empirical link between latent demand and fundamentals is weak in our data, consistent — though not conclusive — with interpreting latent demand as predominantly non-fundamental-based trading. It is important to note, however, that latent demand may also embed slow-moving, long-horizon information that can take several years to fully materialize in fundamentals.[21]

Interpretation of Small Investors. A standard interpretation of non-13F ownership is that it proxies for individual (household) demand Cohen et al. (2002), DeVault et al. (2019). To assess this interpretation, we use market-wide proxies for retail appetite drawn from established evidence: retail trading tilts toward lottery-like payoffs, high idiosyncratic volatility, young firms, and high turnover (Bali et al., 2011; Han and Kumar, 2013; Laarits and Sammon, 2025; Barber and Odean, 2008). Each quarter, we sort the universe into High-Retail and Low-Retail portfolios based on these signals and re-estimate the variance decomposition across six partitions: (i) High vs. Low Retail, and (ii) within each, contributions of *Small Investors*, *Financial Intermediaries* (Banks, Pension Funds, Insurers), and *Delegated Asset Management* (Mutual Funds, Investment Advisors). Table C.3 summarizes the gaps: the Small − Other difference widens sharply in the high-retail quintiles — regardless of whether the benchmark is Intermediaries or Delegated Asset Management — while it narrows in the low-retail quintiles. The pattern is robust across signal definitions.

Operationally, our small-investor residual is *not* the simple 100% − 13F measure used in prior work: we exclude disclosed blockholders

[20] Factor-level decompositions can be inherently noisy: projecting investor portfolios onto characteristic- (and industry-) sorted factors and aggregating to investor-factor cells introduces measurement error and reduces the effective number of observations. The resulting residual reflects these additional forces.

[21] We summarize industry effects with a single term, $\Delta p_t(\mathcal{J})$, which shifts the entire outer-nest block $\mathcal{J}_t \equiv (\boldsymbol{\eta}_t, \boldsymbol{\alpha}_t, \boldsymbol{\sigma}_t)$ and captures industry rotation while preserving adding-up in the five-industry share system. Economically, this channel can mix fundamental-based forces — effective changes in intercepts and substitution elasticities $(\boldsymbol{\alpha}, \boldsymbol{\sigma})$ that govern how investors substitute across industries — and non-fundamental-based movements — industry-level latent demand. Disentangling these components within $\Delta p_t(\mathcal{J})$ is not feasible without additional instruments or structure: aggregate (investor-weighted) $(\boldsymbol{\alpha}, \boldsymbol{\sigma})$ can shift as investors enter/exit the sample, and the adding-up constraints limit separate identification from $\boldsymbol{\eta}$. We therefore treat the industry term as a composite channel plausibly reflecting both fundamental and non-fundamental forces, without imposing a finer decomposition that the data and identification do not support.

and positions attributed to short sellers from the residual. We therefore use the label "small investors" — rather than "non-13F" — to avoid confusion with that accounting residual, while acknowledging classification limits: the residual may still include some non-13F institutions (e.g., small hedge funds, foreign vehicles, family offices) whose incentives can differ from typical households. We retain the "small investor" label but frame conclusions cautiously and emphasize behavior over identity. In our setting, these investors disproportionately drive fundamental-based trading (and, in select categories, non-fundamental-based trading) in ways predicted by the retail literature; excluding blockholders and short sellers helps align the residual with household behavior by removing types more likely driven by control, governance, inventory, or hedging motives. We stop short of a literal household-only claim: the evidence aligns investor-type patterns with canonical retail proxies and remains stable under robustness checks, making a household-led channel a plausible reading of the data.

4.2. Evidence and theory: What drives anomalies

Key facts. *First*, fundamental-based trading is the primary driver of anomaly variance, and it is transmitted primarily by small investors. *Second*, where fundamentals recede, non-fundamental-based trading becomes important and the investor mix tilts toward small investors. *Third*, flow-induced and industry channels are smaller in level and predominantly institution-led, with mutual funds and investment advisors carrying most of these effects.

Flow-based explanations. A recent literature links institutional flows to anomaly returns — momentum (Lou, 2012; Ben-David et al., 2022), size and value (Li, 2022), and broader patterns (Akbas et al., 2015). Our results refine this view by benchmarking flows against other motives and by splitting them across investors. Where flows are material (notably in issuance and size, and to a lesser extent value), the investor composition is institutional: mutual funds and investment advisors dominate — consistent with delegated rebalancing and benchmark pressure. Outside those categories, aggregate flow contributions are modest. More generally, across anomalies they are small relative to fundamental-based demand. In this sense, the data are consistent with flows *amplifying* or *propagating* existing pressures rather than *originating* the bulk of anomaly variance. That said, our evidence does not rule out episodic flow-driven dynamics; it indicates that, on average, flow-based explanations are secondary to fundamentals.

Risk-based explanations. Risk-based theories link anomalies to priced risk or cash-flow dynamics under an exogenous SDF (Berk et al., 1999; Zhang, 2005; Lettau and Wachter, 2007; Santos and Veronesi, 2010; Gabaix, 2012; Kogan and Papanikolaou, 2013). In the investor-demand framework, movements in coefficients on observable characteristics map signals about expected payoffs and risk into portfolio weights (Koijen et al., 2023). The prominence of fundamental-based trading is therefore consistent with a risk-based channel: when characteristics proxy for cash-flow or discount-rate risk, fundamental-based demand should matter. Two tensions remain. First, risk-based models typically speak to a representative investor; our decomposition shows that the transmission of fundamentals is concentrated in small investors, while institutions contribute less on this margin. Second, we observe a nontrivial latent component precisely where characteristics weaken. Neither tension falsifies risk-based stories, but both suggest that risk alone is unlikely to account for the full cross-sectional pattern of motives and investor types

Heterogeneous-investor and mispricing explanations. Models with heterogeneous investors posit two interacting forces: unsophisticated traders who misinterpret fundamentals or trade for non-fundamental reasons, and rational traders who face limits to arbitrage and cannot fully offset mispricing (De Long et al., 1990; Shleifer and Vishny, 1997; Daniel et al., 2001; Barberis et al., 1998; Hong and Stein, 1999). Misinterpretation may reflect over-/underreaction to signals, selective or coarse use of fundamentals, or mistaken beliefs about signal precision; non-fundamental trading can stem from sentiment, liquidity, or financing frictions. With risk, funding, and benchmarking constraints, rational capital accommodates rather than eliminates these pressures, allowing return predictability to persist.

Our estimates align with this view. The fundamental-based component is large in aggregate and is transmitted primarily by small investors across anomaly categories, consistent with theories in which a subset of investors moves prices by acting strongly on — potentially imperfectly processed — fundamentals. The non-fundamental-based component is economically meaningful in specific anomalies and again tilts toward small investors, consistent with a non-fundamental-driven interpretation of trade. Institution-led forces appear mainly in the flow and industry channels and are secondary in level, as expected if delegated rebalancing and arbitrage frictions shape — but do not dominate — anomaly dynamics.

To link the evidence to theory more tightly, we integrate our nested-logit decomposition with the classic cash-flow/discount-rate (CF–DR) news breakdown. We focus on the dominant margins — two trading motives (fundamental-based demand for characteristics and non-fundamental-based latent demand) and three investor groups (small investors, mutual funds, and investment advisors) — which together account for 47% of the average anomaly's return variance. Following (Lochstoer and Tetlock, 2020), who show that competing anomaly theories imply distinct CF–DR footprints, we combine the CF–DR decomposition with our investor–motive breakdown to trace how each motive is transmitted by specific investor types through cash-flow versus discount-rate news. This integrated design preserves an interpretable investor lens while mapping our estimates onto theory-relevant channels, allowing direct tests of which combinations of motives and investors line up with CF or DR forces.

4.3. Decomposition of cash-flow and discount-rate news effects

We decompose innovations in log stock returns into cash-flow (CF) and discount-rate (DR) news,

$$\tilde{r}_{n,t} = r_{n,t} - E_t\left[r_{n,t}\right] \approx cf_{n,t} - dr_{n,t}, \tag{18}$$

using the log-linear present-value framework of Campbell (1991). Following Lochstoer and Tetlock (2020), we specify a VAR for expected returns and construct stock-level news as the sum of aggregate and market-adjusted components. As in Chen et al. (2013), we estimate stock-specific components with stock fixed effects to absorb persistent firm heterogeneity.

To align the CF/DR decomposition with our portfolio analysis, we aggregate stock-level log news to anomaly portfolios and convert them to simple returns. We follow Lochstoer and Tetlock (2020, Section IV "Portfolio Aggregation" in the Online Appendix) to obtain portfolio-level simple CF and DR news, $CF_{p,t}$ and $DR_{p,t}$, and the innovation to the simple portfolio return, $\tilde{R}_{p,t}$.

We then embed our investor–motive decomposition into this CF/DR framework. We replace $r_{n,t}$ in (18) with $r^{(k_i)}_{n,t}$, the log return on stock n induced by motive k for investor i, and apply the same procedure to obtain $\tilde{R}^{(k_i)}_{p,t}$ at the portfolio level. As in Section 2.3.2, we work with log transformations of return innovations and news. For portfolio log returns we have $\tilde{r}_{p,t} \approx \sum_k \sum_i \tilde{r}^{(k_i)}_{p,t} \approx cf_{p,t} - dr_{p,t}$, so the variance of log return innovations can be written as

$$Var(\tilde{r}_{p,t}) = \sum_k \sum_i Cov\left(\tilde{r}^{(k_i)}_{p,t}, cf_{p,t}\right) + \sum_k \sum_i Cov\left(\tilde{r}^{(k_i)}_{p,t}, -dr_{p,t}\right) + Cov\left(u_t, \tilde{r}_{p,t}\right), \tag{19}$$

where u_t is the residual that collects CF/DR approximation error and the log-to-simple-return mapping. Dividing both sides of (19) by $Var(\tilde{r}_{p,t})$ yields a three-way decomposition of variance into the effects of investor types, trading motives, and CF versus DR news. Full implementation details are in Appendix C.6.

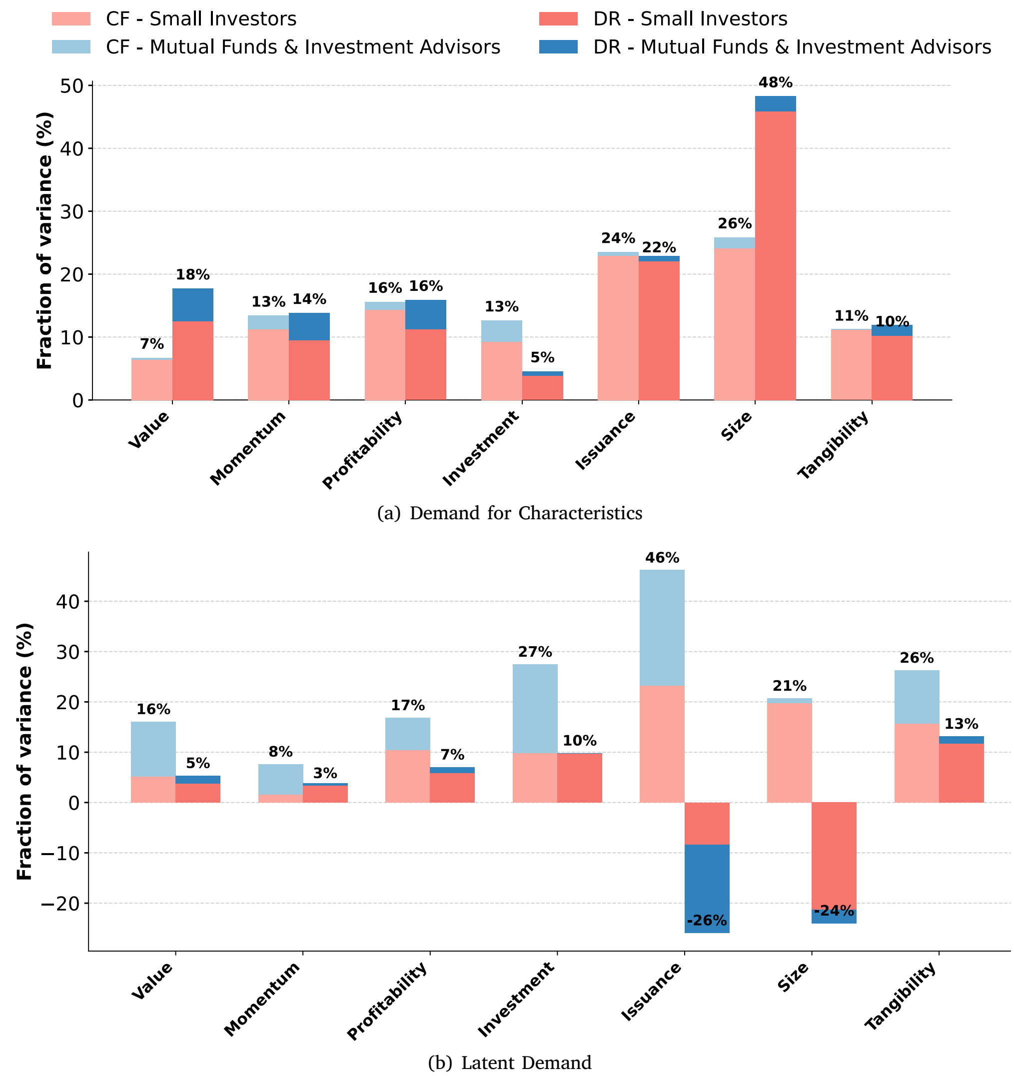


Fig. 3. Cash-Flow vs. Discount-Rate News by Investor Type and Trading Motive This figure reports a two-step variance decomposition of anomaly returns. Step 1 allocates total anomaly returns to trading motives and investor types. Step 2 splits the innovations in those motive–investor effects into cash-flow (CF) and discount-rate (DR) news. Panel (a) shows the fundamental-based component (demand for characteristics); Panel (b) shows the non-fundamental component (latent demand). Within each panel, we compare small investors to major institutions (mutual funds and investment advisors). Bars report the share of variance (by anomaly category) attributable to each motive×investor×news component; numbers above bars give the total motive effect for that investor type. Anomaly categories follow Section 2.3.2.

Fig. 3 summarizes how the investor-by-motive effects split into CF and DR news across anomaly categories. For each category, Panel (a) shows the fraction of variance in anomaly portfolio return innovations explained by CF and DR news associated with demand for characteristics, while Panel (b) reports the corresponding variance contributions from the CF and DR components associated with latent demand.[22]

Panel (a) shows that, when pooling small investors with mutual funds and investment advisors, both CF and DR components are important drivers of demand for stock characteristics, with magnitudes that are broadly similar. This near balance holds across most anomaly categories. The main exceptions are investment, where the CF channel is stronger than the DR channel (about 13% vs. 5%), and value and size, where the DR component is larger than the CF component (about 18% vs. 7% for value and 48% vs. 26% for size).

Panel (b) shows a clear tilt toward CF news. On average, comovement between latent demand and CF news accounts for most of the explained variance (around 13%), while DR news contributes a smaller share (around 4%). This CF dominance is especially pronounced for

[22] The CF/DR totals in this figure do not mechanically add up to the motive effects in Fig. 1 for two reasons. First, the CF/DR decomposition is defined for innovations in log returns, $\tilde{r}_{p,t}$, whereas Fig. 1 is based on realized returns; as a result, both the numerator and the denominator, $Var(r_p)$, differ across the two exercises. Second, following Lochstoer and Tetlock (2020), we drop the CF×DR cross-term in the variance identity and apply an additional log transformation to align our demand-based decomposition with their CF/DR framework. These necessary steps generate a residual wedge that we group into u_t, so gaps between CF+DR totals and Fig. 1 motive wedges are not surprising. In several anomaly categories, however, the CF/DR totals are quite close to the realized-return shares: for demand for characteristics, CF+DR accounts for 24% versus 25.5% in value and 46% versus 48% in issuance, while for latent demand it accounts for 37% versus 33.2% in investment and remains close to zero for size (−3% versus 3.8%). In categories where the gaps are larger, the CF/DR breakdown still preserves the main ranking of investor motives documented in Fig. 1, so we view it as a consistent, though not numerically identical, refinement of our baseline decomposition.

investment and size, where CF effects are large and DR effects are modest or even negative, and it also holds for value, profitability, tangibility, and, to a lesser extent, momentum. Overall, latent demand operates primarily through CF news rather than DR news across anomaly categories.

These findings connect our evidence to the theoretical foundations of demand-based asset pricing. Koijen et al. (2023) show that coefficients on stock characteristics encode investors' beliefs about expected cash-flow growth and risk exposures. The near 50/50 split between CF and DR news in Panel (a) implies that, for anomaly portfolios, variation associated with demand for characteristics loads about equally on news about cash flows and news about discount rates. Portfolio holdings data alone cannot pin down whether a given characteristic is primarily about expected profitability or risk, but the CF/DR decomposition indicates that both channels are quantitatively important and should enter any interpretation of characteristic-driven demand. At the same time, the more prominent role of DR news for value and size, and the more balanced role of cash flows and discount rates for momentum, caution against a uniform cash-flow-versus-discount-rate explanation of fundamental-based trading across anomalies.

This analysis also allows us to relate our results to the explanations of anomaly drivers in Lochstoer and Tetlock (2020). These authors find that 64%–80% of anomaly-return variance is driven by CF news. In our setting, CF news accounts for 58%–71% of the variation across anomaly categories, with an average of 62%, confirming that CF news dominates DR news in explaining anomaly returns. Based on similar magnitudes, Lochstoer and Tetlock (2020) argue that heterogeneous-investor theories such as Barberis et al. (1998) and Daniel et al. (2001), as well as the risk-based model of Kogan and Papanikolaou (2013), are particularly plausible because they generate anomalies primarily through CF news. Our evidence is consistent with this view, and our results on how investor–motive effects load on CF news further sharpen these conclusions.

First, small investors account for 26% of the CF-related variance in anomaly returns (14% from demand for characteristics in Panel (a) and 12% from latent demand in Panel (b)), nearly half of the total CF effect. This pattern aligns with heterogeneous-investor models such as Barberis et al. (1998) and Daniel et al. (2001), which attribute CF news primarily to unsophisticated investors whose beliefs are distorted by extrapolation and overconfidence. Second, the dominance of small investors in driving CF-news comovement with characteristic-driven demand holds across all anomaly categories and suggests that systematic biases in processing firm-specific information can simultaneously generate several of the major anomalies.

Our findings also offer a new perspective on noise-trading theories such as De Long et al. (1990), which posit that return variation is entirely driven by DR news. Lochstoer and Tetlock (2020) already find limited support for this view, as DR news account for only a small share of anomaly-return variance, and in our data this fraction is below one-third as well. By decomposing return variation by motives, we can sharpen this assessment. We proxy noise trading with latent demand — trading unrelated to observable fundamentals — and De Long et al. (1990) implies that its effects should load purely on DR news. Panel (b) allows us to test this implication directly: the share of variance attributed to latent demand that loads on DR news ranges from about 25% for value (5%/(5%+16%)) to about 33% for tangibility (13%/(26%+13%)). Thus, while pure noise-trading models receive some support given the DR loading of latent demand, they should explain at most one-third of the return variation induced by latent demand, with the remainder necessarily coming from CF news.

5. Conclusions

We present new evidence on what drives anomaly returns by decomposing return variation jointly by trading motives and investor types within a nested–logit demand system. Two forces emerge. Demand for characteristics (interpreted as fundamentals-based demand) dominates across anomaly categories and is transmitted primarily by small investors; latent (non-fundamentals-based) demand is the next most important component and likewise tilts toward small investors precisely where it matters. In contrast, institution-led channels — flow-induced trading and industry effects — are secondary in magnitude and concentrated in a few anomaly categories. This hierarchy is robust across market states and asset definitions.

From an econometric standpoint, substitution patterns from reduced-form flow-multiplier regressions validate our industry nesting. Finally, the small-investor residual is not simply "100% minus 13F": we exclude blockholders and short sellers, and retail-tilt diagnostics are consistent with a household-led channel, but we acknowledge remaining classification limits and encourage further research along this dimension.

These facts refine the mapping from evidence to theory. Risk-based explanations align with the prominence of fundamentals-based demand but speak less directly to the investor split and to the role of latent demand. Flow-based explanations capture institutional amplification in specific categories, yet flows are not the primary origin of anomaly variance on average. Heterogeneous-investor models with noise traders and constrained arbitrage organize the joint facts on motives and investor types: small investors disproportionately transmit fundamentals-based demand and, in selected anomalies, non-fundamentals-based demand, while institutions primarily mediate flows and industry tilts. Overall, anomalies are shaped primarily by investor demand for characteristics, with a meaningful non-fundamentals layer, calling for theories that combine investor heterogeneity, realistic substitution across asset groups (e.g., industries), and a disciplined role for institutional flows.

CRediT authorship contribution statement

Yizhang Li: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Stanislav Sokolinski:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Andrea Tamoni:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors have no conflicts of interest to disclose.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jfineco.2026.104257.

Data availability

Replication Package for "Which Investors Drive Anomaly Returns and How?" (Original data) (Mendeley Data)

References

Akbas, F., Armstrong, W.J., Sorescu, S., Subrahmanyam, A., 2015. Smart money, dumb money, and capital market anomalies. J. Financ. Econ. 118 (2), 355–382.

Asness, C.S., Frazzini, A., Pedersen, L.H., 2019. Quality minus junk. Rev. Account. Stud. 24 (1), 34–112.

Balasubramaniam, V., Campbell, J.Y., Ramadorai, T., Ranish, B., 2023. Who owns what? A factor model for direct stockholding. J. Financ. 78 (3), 1545–1591.

Bali, T.G., Cakici, N., Whitelaw, R.F., 2011. Maxing out: Stocks as lotteries and the cross-section of expected returns. J. Financ. Econ. 99 (2), 427–446.

Barber, B., Odean, T., 2008. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. Rev. Financ. Stud. 21 (2), 785–818.

Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. J. Financ. Econ. 49 (3), 307–343.

Ben-David, I., Franzoni, F., Moussawi, R., Sedunov, J., 2021. The granular nature of large institutional investors. Manag. Sci. 67 (11), 6629–6659.

Ben-David, I., Li, J., Rossi, A., Song, Y., 2022. Ratings-driven demand and systematic price fluctuations. Rev. Financ. Stud. 35 (6), 2790–2838.

Berk, J.B., Green, R.C., Naik, V., 1999. Optimal investment, growth options, and security returns. J. Financ. 54 (5), 1553–1607.

Betermier, S., Calvet, L.E., Knüpfer, S., Soerlie Kvaerner, J., 2024. Investor factors. J. Financ.

Betermier, S., Calvet, L.E., Sodini, P., 2017. Who are the value and growth investors? J. Financ. 72 (1), 5–46.

Bretscher, L., Schmid, L., Sen, I., Sharma, V., 2025. Institutional corporate bond pricing. Rev. Financ. Stud. hhaf067.

Campbell, J.Y., 1991. A variance decomposition for stock returns. Econ. J. 101 (405), 157–179.

Chaudhary, M., Fu, Z., Li, J., 2025. Corporate bond multipliers: Substitutes matter. Working Paper.

Chen, L., Da, Z., Zhao, X., 2013. What drives stock price movements? Rev. Financ. Stud. 26 (4), 841–876.

Choi, J., Tian, X., Wu, Y., Kargar, M., 2023. Investor demand, firm investment and capital misallocation. Working Paper.

Cohen, R.B., Gompers, P.A., Vuolteenaho, T., 2002. Who underreacts to cash-flow news? Evidence from trading between individuals and institutions. J. Financ. Econ. 66 (2), 409–462.

Cohen, R.B., Polk, C., Vuolteenaho, T., 2003. The value spread. J. Financ. 58 (2), 609–641.

Daniel, K.D., Hirshleifer, D., Subrahmanyam, A., 2001. Overconfidence, arbitrage, and equilibrium asset pricing. J. Financ. 56 (3), 921–965.

Davis, C.K., 2024. The elasticity of quantitative investment. Working Paper.

De Long, J.B., Shleifer, A., Summers, L.H., Waldmann, R.J., 1990. Noise trader risk in financial markets. J. Political Econ. 98 (4), 703–738.

DeVault, L., Sias, R., Starks, L., 2019. Sentiment metrics and investor demand. J. Financ. 74 (2), 985–1024.

Edelen, R.M., Ince, O.S., Kadlec, G.B., 2016. Institutional investors and stock return anomalies. J. Financ. Econ. 119 (3), 472–488.

Freyberger, J., Neuhierl, A., Weber, M., 2020. Dissecting characteristics nonparametrically. Rev. Financ. Stud. 33 (5), 2326–2377.

Fuchs, W., Fukuda, S., Neuhann, D., 2025. Demand-system asset pricing: Theoretical foundations. Working Paper.

Gabaix, X., 2012. Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance. Q. J. Econ. 127 (2), 645–700.

Gabaix, X., Koijen, R.S.J., Mainardi, F., Oh, S., Yogo, M., 2022. Asset demand of U.S. households. Working Paper.

Haddad, V., Huebner, P., Loualiche, E., 2024. How competitive is the stock market? theory, evidence from portfolios, and implications for the rise of passive investing. Working Paper.

Haddad, V., Muir, T., 2021. Do intermediaries matter for aggregate asset prices? J. Financ. 76 (6), 2719–2761.

Han, B., Kumar, A., 2013. Speculative retail trading and asset prices. J. Financ. Quant. Anal. 48 (2), 377–404.

Han, X., Roussanov, N.L., Ruan, H., 2022. Mutual fund risk shifting and risk anomalies. Working Paper.

Hanson, S.G., Sunderam, A., 2014. The growth and limits of arbitrage: Evidence from short interest. Rev. Financ. Stud. 27 (4), 1238–1286.

Hong, H., Stein, J.C., 1999. A unified theory of underreaction, momentum trading, and overreaction in asset markets. J. Financ. 54 (6), 2143–2184.

Huebner, P., 2024. The making of momentum: A demand-system perspective. Working Paper.

Kogan, L., Papanikolaou, D., 2013. Firm characteristics and stock returns: The role of investment-specific shocks. Rev. Financ. Stud. 26 (11), 2718–2759.

Koijen, R.S., Richmond, R.J., Yogo, M., 2023. Which investors matter for equity valuations and expected returns? Rev. Econ. Stud..

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

Koijen, R.S., Yogo, M., 2025. Exchange rates and asset prices in a global demand system.

Kolasinski, A.C., Reed, A.V., Ringgenberg, M.C., 2013. A multiple lender approach to understanding supply and search in the equity lending market. J. Financ. 68 (2), 559–595.

Laarits, T., Sammon, M., 2025. The retail habitat. Working Paper.

Lettau, M., Wachter, J.A., 2007. Why is long-horizon equity less risky? A duration-based explanation of the value premium. J. Financ. 62 (1), 55–92.

Lewellen, J., 2011. Institutional investors and the limits of arbitrage. J. Financ. Econ. 102 (1), 62–80.

Li, J., 2022. What drives the size and value factors? Rev. Asset Pricing Stud. 12 (4), 845–885.

Li, J., Lin, Z., 2025. Price multipliers are larger at more aggregate levels. Working Paper.

Lochstoer, L.A., Tetlock, P.C., 2020. What drives anomaly returns? J. Financ. 75 (3), 1417–1455.

Lou, D., 2012. A flow-based explanation for return predictability. Rev. Financ. Stud. 25 (12), 3457–3489.

Mainardi, F., 2021. A demand-based approach for short-selling. Working Paper.

McLean, R.D., Pontiff, J., Reilly, C., 2022. Taking sides on return predictability. Working Paper.

Santos, T., Veronesi, P., 2010. Habit formation, the cross section of stock returns and the cash-flow risk puzzle. J. Financ. Econ. 98 (2), 385–413.

Schwartz-Ziv, M., Volkova, E., 2024. Is blockholder diversity detrimental? Manag. Sci..

Shleifer, A., Vishny, R.W., 1997. The limits of arbitrage. J. Financ. 52 (1), 35–55.

Sloan, R.G., 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? Account. Rev. 289–315.

Vuolteenaho, T., 2002. What drives firm-level stock returns? J. Financ. 57 (1), 233–264.

Zhang, L., 2005. The value premium. J. Financ. 60 (1), 67–103.