Don't Sweat the Clear Stuff:

Mutual Fund Shorts and Research Intensity

Boone Bowles and Adam V. Reed*

July 2026

Abstract

Using mutual fund holdings linked to EDGAR requests, we study how fund families research long and short positions. Among families with observable research activity, short positions outperform comparable longs by roughly 100 basis points per quarter and receive about 10% more EDGAR requests. We develop a model of information acquisition with short-side frictions that rationalizes greater research intensity and higher returns, and a sequential-sampling extension predicts, paradoxically, that more-researched shorts can earn smaller absolute returns. The evidence is consistent with this inverse relation: attention concentrates on borderline shorts, while clear winners require less research.

JEL classification: D83, G14, G23

Keywords: mutual funds, information economics, short selling.

*Mays Business School, Texas A&M University (boone.bowles@tamu.edu), Kenan-Flagler Business School, University of North Carolina (adam_reed@unc.edu). The authors thank Dora Horstman, Yunzhi Hu, Matthew Ringgenberg, Jesse Davis, and conference and seminar participants at Texas A&M University, University of Cologne—Center for Financial Research, Southwestern Finance Association Annual Meeting (2024), Midwest Finance Association Annual Meeting (2024), Eastern Finance Association Annual Meeting (2024), 16th Annual Paris Hedge Fund Conference (2025), and 2nd Modern Finance Conference (2025). All errors are our own. We (the authors) have no conflicts of interest to disclose.

1 Introduction

Short selling plays a critical role in market efficiency while also posing distinct frictions and risks. Unlike buying, shorting is impeded by locate requirements, borrowing costs, and potentially unlimited downside risks. These frictions have inspired extensive literature examining whether and how short selling impacts price discovery and asset pricing more generally.[1] However, relatively little is known about how short-selling constraints shape institutional investors' decisions around information acquisition.

A particularly important class of institutional investors in this context is active mutual fund managers, who control large amounts of capital and often operate under strict portfolio constraints, including limitations on short selling. Yet for this investor class, we still know relatively little about whether the stocks that mutual fund families short subsequently underperform the stocks those same families hold long, or whether manager research differs across the long and short sides of their portfolios. This paper provides a within-family comparison of mutual fund short- and long-position returns and links that comparison to direct evidence on pre-trade research around those positions.

To answer these questions, we link two complementary data sets: (i) position-level holdings for U.S. equity mutual fund families, and (ii) the time-stamped request logs from the SEC's EDGAR filing system. Merging the two lets us observe—at the fund family-stock-quarter level—which public filings are requested by IP addresses linked to a mutual fund family before a position appears in their portfolio. This linkage gives us an unusually granular view of the information-production process inside a major class of institutional investors.

We document two significant facts. First, within the same mutual fund family, short positions substantially outperform long positions: shorts earn average abnormal returns of 125 basis points per quarter, compared with 30 basis points for long positions and 86 basis points for large long positions. Second, those same mutual fund families devote more observable pre-trade research to shorts, retrieving roughly 6% more EDGAR filings for prospective

[1] E.g., Jones and Lamont (2002), Boehmer et al. (2008), and Engelberg et al. (2018).

shorts than for prospective longs and about 10% more than for benchmark-adjusted active longs. Together, these facts raise a natural question: why would short positions be associated with both more research effort and larger subsequent abnormal returns?

To explain these patterns we introduce a tractable model, nested in the Grossman and Stiglitz (1980) framework, that puts a single twist on the classic setup: short positions carry extra, side-specific costs. An investor chooses how much noisy public information to acquire before deciding whether to take a fixed-size long or short position. The usual convex cost of precision applies to both longs and shorts, and both also face a fixed monitoring charge, but only shorting requires an additional locate fee and a size-dependent surcharge that captures fee-volatility and recall risk. These frictions raise the precision threshold the investor must clear before it is worthwhile to short.

Echoing the Grossman and Stiglitz (1980) intuition, the investor collects information until the expected benefit of additional information matches its marginal cost. Because the threshold is higher for shorts, the optimal precision, and hence research intensity, is greater on the short side. At the same time, shorts are executed only when the signal implies a larger mispricing. Taken together, the model provides a disciplined interpretation of Fact 1 (more observable research for shorts than longs) and Fact 2 (larger absolute returns for shorts than longs).

Importantly, this model yields additional testable predictions when we extend it to sequential information acquisition, following the logic formalized by Banerjee and Breon-Drish (2020). In this setting the investor samples signals until the expected mispricing is large enough to justify a short. When the first signal is extreme, the position resembles a clear winner and little additional research is required. When early signals are more equivocal, the investor purchases further precision and the trade becomes a borderline short. The framework therefore predicts an inverse relation inside the short portfolio: positions accompanied by more requests should realize smaller absolute abnormal returns, whereas longs—unburdened by the short-side cost threshold—should display no such pattern. This inversion is the model's distinctive mechanism: research intensity is not simply correlated with expected payoffs; it can also reflect how much uncertainty remains after initial screening.

Putting this prediction to the data yields evidence consistent with the clear-winner-versus-borderline-short logic. Among shorts, a one-standard-deviation *increase* in filing requests corresponds with roughly a 40-basis-point *decline* in absolute returns. For long positions, by comparison, the same increase in requests has no consistently discernible effect. This asymmetric slope—more specifically the inverse slope among shorts—is the pattern predicted by the sequential-sampling model and is difficult to anticipate without a formal model and from only a simple "more research, better performance" view. Together, the evidence supports the model's most nuanced economic insight.

Furthermore, complementary evidence underscores the broader economic logic of informed trading. Mutual fund families with more intensive EDGAR activity realize consistently higher benchmark-adjusted returns. This finding suggests that research is productive on average, supporting a rational-choice interpretation of information acquisition. Therefore, the negative slope observed among short positions need not indicate inefficient information acquisition. Instead, it is consistent with greater research effort being allocated to less obvious opportunities, which can still contribute positively to overall performance.

Additionally, we provide external support for the economic intuition of the model by showing that an independent measure of ex ante mispricing predicts future stock returns: stocks identified as overpriced subsequently earn low returns. This predictive relation supports the clear-winner interpretation of managers' differential research strategies. Although our empirical setting is mutual funds, the model's logic is broader: when informed agents screen opportunities before choosing how much more to learn, observed research intensity can be highest for marginal decisions, not necessarily for the trades with the largest eventual payoffs.[2]

[2]We interpret this evidence as model-consistent rather than as direct identification of the mechanism. The EDGAR logs identify public filing requests linked to mutual fund families, and the holdings data identify family-stock-quarter positions. We do not observe the internal signal sequence, individual decision-maker, or exact trade date. This limitation is important, but it does not eliminate the economic content of the test.

Our analysis makes several important contributions to the literature on short selling, information acquisition, and mutual fund behavior. First, we provide within-family evidence that mutual funds' shorts outperform their longs. We then link that performance gap to direct evidence on pre-trade research around those positions, addressing a critical gap left by earlier work that relies on aggregate proxies. Unlike studies utilizing aggregate search data (Da et al. (2011), Ben-Rephael et al. (2017)) or aggregated short-interest measures (Seneca (1967), Hong and Stein (2003), Asquith et al. (2005), Boehmer et al. (2008)), our data link family-level EDGAR activity to mutual fund positions, allowing us to better observe the relation between information gathering and subsequent portfolio decisions. Our focus on longs versus shorts within mutual fund families differs from earlier evidence that aggregates short positions across investors.

Second, our findings deepen the understanding of short-selling constraints by highlighting information acquisition as a margin through which short-selling frictions may affect investor behavior. Prior studies such as Engelberg et al. (2018) and An et al. (2021) document the broad effects of borrowing constraints and institutional limits; our work points to endogenous information acquisition, shaped by marginal costs and benefits, as a complementary channel through which these constraints influence institutional portfolios.

Third, we contribute to a small but growing literature on mutual fund shorting behavior. Building on Almazan et al. (2004), who document widespread self-imposed prohibitions, and Chen et al. (2013), who provide an early look at mutual fund shorting skill, we show that 48% of the mutual fund families in our study engage in short selling. We further show that those short positions are preceded by more observable research and deliver higher abnormal returns than corresponding long positions within the same mutual fund families. Our findings also clarify the relation between our evidence and An et al. (2021). They study long-short equity mutual funds at the fund-product level and show that total fund returns are shaped by cash holdings, market exposure, and flow management. We study a different margin:

The inverse slope and clear-winner proxy results are the patterns predicted by the model and are difficult to reconcile with a simple "more research, better performance" view.

position-level research and stock-level returns within mutual fund families that use short sales. Our claim is not that long-short mutual funds necessarily outperform as products; it is that the stocks mutual fund families choose to short are more intensively researched and subsequently underperform the stocks those same families hold long.

Fourth, we add new evidence to the empirical literature on institutional information acquisition. Our position-level approach complements recent EDGAR studies of hedge funds and other institutions (Chen et al. (2020a), Chen et al. (2020b), Crane et al. (2022), Gibbons et al. (2021), Gibbons (2023)) by linking EDGAR requests to both sides of mutual fund family portfolios. We also document a counterintuitive inverse relation: among shorts, more pre-trade research is associated with smaller absolute abnormal returns. This pattern, predicted by our model, challenges the simple view that more research should map monotonically into better trade performance and offers a framework for understanding how skilled investors allocate attention across clear and borderline opportunities.

Fifth, we extend the theoretical literature by explicitly incorporating sequential information acquisition into the classical Grossman-Stiglitz (1980) framework. By doing so, our model bridges insights from sequential search theory (Banerjee and Breon-Drish (2020)) and earlier foundational work on costly information acquisition (Hellwig (1980), Verrecchia (1982)). This theoretical extension yields tractable and novel predictions regarding managerial behavior under costly short selling, and we test those predictions using our mutual fund setting.

The remainder of the paper proceeds as follows. Section 2 details the data sources and construction of key variables. Section 3 presents our two core empirical facts about the short-versus-long differences in information acquisition and returns. Section 4 introduces and develops our theoretical framework, deriving predictions consistent with our empirical findings. Section 5 tests the model's novel prediction regarding the within-short relationship between information acquisition and returns. Section 6 provides additional tests and Section 7 concludes. A full description of the data and a detailed theoretical derivation are provided in the Online Appendix.

2 Data

This study draws on two primary data sources. First, we obtain quarterly mutual fund holdings from the Center for Research in Security Prices Survivor-Bias-Free U.S. Mutual Fund Database (*CRSP Mutual Funds*). For each mutual fund family i, stock j, and quarter t we record the number of shares reported (nbr_shares_{ijt}). A long position, $Long_{ijt}$, equals one when $nbr\_shares_{ijt} > 0$, while a short position, $Short_{ijt}$, equals one when $nbr\_shares_{ijt} < 0$. We also flag the first quarter in which a family initiates a position as $NewLong_{ijt}$ or $NewShort_{ijt}$, respectively. In addition, we construct benchmark-adjusted active weights at the family-stock-quarter level. Specifically, we first assign each underlying fund a representative benchmark portfolio based on that fund's Morningstar category, compute the fund-level benchmark-adjusted weight of stock j, and then aggregate these benchmark-adjusted weights to the family-stock-quarter level using fund assets under management as weights. We define $ActiveLong_{ijt} = 1$ when this aggregated benchmark-adjusted weight exceeds 33 basis points.[3]

Second, we measure research intensity with the U.S. SEC's EDGAR Log Files, which record every electronic request for a public filing. Each observation details the filing requested, the date and time of the request, and the requester's IP address.[4] Following the unmasking procedure detailed in Appendix A.1, we link IP addresses to mutual fund families. For every i–j–t we then count the number of filings family i requests about firm j during quarter t and label this $Requests_{ijt}$. Our baseline empirical proxy for information acquisition is $MFIA_{ijt} = ln(1 + Requests_{ijt})$. To capture pre-trade research around position formation, we also compute $Requests_{ijt,t-1}$, the two-quarter sum of requests, and use the

[3]Our results are robust to using cutoffs of both 25 and 50 basis points.

[4]For example, the filing request is for the 2013 annual report (10-K) for IBM. The request was made at 10:14 am on March 1, 2014. The request originated from the IP address 123.123.123.*abc*.

same transformation.[5]

Empirically, $MFIA$ captures requests for public SEC filings linked to mutual fund families. We therefore interpret it as a proxy for research intensity rather than as a complete measure of all information acquisition.

The quarterly holdings data do not reveal the exact trade date within the quarter, which matters most for interpreting return tests. The request-position tests are more direct: several specifications use lagged request windows, including requests measured before the quarter in which a new position is first observed. For the return-based analyses, Appendix A.4 reports conservative timing checks that use positions measured in quarter t and returns measured after quarter t.

We construct our sample beginning with all common stocks and all CRSP mutual fund families between 2010 and 2017 (30 quarters). We retain an i-j pair only if mutual fund family i reports a non-zero position in stock j at least once during the sample window. We further require that (i) the family takes *at least one* short position during the window, and (ii) its average quarterly EDGAR activity is ≥ 5 requests. Observations in which a family simultaneously holds both long *and* short positions in the same stock are dropped.

The final panel contains 2,575,774 family-stock-quarter observations. Roughly 49% indicate a long position, whereas just under 3% indicate a short position and nearly 2.5% are ActiveLongs. Aggregated to the family-quarter level, the median family submits 1,611 filing requests per quarter—the mean is 8,919—and at least one request appears in 99% of family-quarters. Our inclusion criteria reduce the fund count from the CRSP universe of 115 families to 55 families that actively short and have sufficient observable EDGAR activity. Thus, our tests focus on the subset of mutual fund families for which shorting behavior and EDGAR-linked research can both be observed. Summary statistics are reported in Table 1, and Figure 1 plots the time series of aggregate long, ActiveLong, short, new-long, and

[5]Additional stock-level variables and fund characteristics come from CRSP, Compustat, Thomson Reuters 13F, RavenPack, and the mispricing scores of Stambaugh and Yuan (2017). A complete variable dictionary appears in Appendix A.2.

new-short positions. Overall, the data afford a novel view of both sides of mutual fund family portfolios and of the information-gathering efforts that precede them.

[Table 1 about here.]

[Figure 1 about here.]

3 Two Empirical Facts

This section establishes two stylized facts that motivate the remainder of the analysis. First, mutual fund families devote more observable research to their short positions than to their long positions. Second, those short positions, on average, subsequently earn better (i.e., more negative) risk-adjusted returns than corresponding long positions within the same mutual fund families.

3.1 Fact 1: Research Intensity and Position Type

We proxy for research intensity with the number of EDGAR filings requested by mutual fund family i for company j during quarter t. To test whether research intensity differs across position types we estimate

$$MFIA_{ijt} = \beta_1 Long_{ijt} + \beta_2 Short_{ijt} + \gamma_i + u_j + \omega_t + \varepsilon_{ijt}, \quad (1)$$

where $MFIA$ (defined earlier) includes either one quarter of requests or two quarters and the model includes mutual fund family, stock, and quarter fixed effects.

The coefficient estimates for *Long* and *Short* capture research intensity for long and short positions relative to non-holdings, within a mutual fund family. To test whether research intensity differs between shorts and longs, we test the linear combination $\beta_2 - \beta_1 = 0$. The results from estimating Equation 1 are shown in Panels A and B of Table 2.

[Table 2 about here.]

Three patterns stand out. First, the positive and statistically significant estimates for both *Long* and *Short* in Panel A indicate that mutual fund families devote more observable research to the stocks they hold—on either side of their portfolios—than to stocks they do not hold. This is true when requests and positions are measured simultaneously or when considering lagged requests. In economic terms, families make roughly 10% more requests for their long positions and roughly 20% more requests for their short positions, relative to non-holdings.

Second, the short-minus-long difference in Panel A remains about 6% and is statistically significant across the core specifications. Third, this pattern is stronger in the benchmark-adjusted comparison that replaces *Long* with *ActiveLong*. Panel B shows that the short-minus-active-long gap is around 10% across the core specifications.[6]

One concern is that the differential could be driven by large, legacy positions that require little incremental analysis. Panel C addresses this possibility by re-estimating Equation 1 with *NewLong* and *NewShort*—indicator variables for positions first observed during quarter t. Here, the differential is as large as 8%, and the effect is strongest in the specifications that use lagged request windows. Thus, before and around the formation of new positions, mutual fund families ramp up observable research on prospective shorts but not on prospective longs. Overall, the results in Table 2 establish our first empirical fact.

Fact 1: *Mutual fund families allocate disproportionately more observable research effort to the positions they hold short.*

[6] Appendix A.3 re-estimates these request-position regressions with stock-quarter and family-quarter fixed effects. The short-minus-long and short-minus-active-long gaps remain positive and similar in magnitude, which helps rule out common stock-quarter filing-supply shocks and family-quarter research intensity as the sole explanation.

3.2 Fact 2: Position Type and Returns

Next we ask whether shorts earn better returns than longs within mutual fund families. In other words, we compare the long sides of a family's portfolios with the short sides of the same family's portfolios. We estimate

$$Return_{ijt} = \alpha + \beta Short_{ijt} + \lambda Z_{jt} + \gamma_i + \varepsilon_{ijt}, \tag{2}$$

where $Return_{ijt}$ is either the excess return of stock j during quarter t ($ExRet$) or the stock's characteristic-adjusted return ($DGTW$).[7] Vector Z includes the market return and three risk factors: SMB, HML, and UMD.[8] The model also includes a fund family fixed effect.

We estimate the model across long-only, short-only, and pooled position samples, and we also report contemporaneous new-position results in the same exhibit. In the first two subsamples (only longs or only shorts) the intercept α captures the abnormal return. The third subsample includes both long and short positions, thus α captures the abnormal return for longs while $Short$ measures the additional return to shorts. We also include $ActiveLong$ to capture the abnormal return to meaningful long positions, and we include new positions. Table 3 shows the results.

[Table 3 about here.]

The evidence for longs is mixed: abnormal returns using the risk-factor model average 29 bps per quarter, but characteristic-adjusted returns average only 12 bps and are statistically insignificant (Column 2). Shorts are different: across all specifications they deliver sizable negative abnormal returns—between 103 and 128 bps per quarter—implying that shorts outperform longs by roughly 100 bps. Even accounting for ActiveLongs, which add 79 to 86 bps per quarter relative to ordinary long positions, shorts still outperform.

[7]Daniel et al. (1997).

[8]Fama and French (1993) and Carhart (1997).

To address timing concerns, we also estimate Equation 2 on new positions only.[9] Column 6 of Table 3 reports a similar pattern: new shorts deliver additional gains of up to 353 bps per quarter while new longs generate 50% less (258 bps per quarter). The histograms of characteristic-adjusted returns in Figure 2 visualize these findings and help establish our second empirical fact.

Fact 2: *Mutual funds' short positions consistently outperform their long positions.*[10]

[Figure 2 about here.]

4 Model

Our empirical analysis establishes two facts. First, mutual fund families undertake more observable research around short positions than around comparable long positions. Second, those shorts subsequently earn substantially larger abnormal returns than the longs. In this section we develop a theoretical framework that contains these two facts while clarifying why costly short positions can be associated with both greater research intensity and higher average payoffs. Accordingly, we introduce a concise analytical framework—a variant of the Grossman and Stiglitz (1980) costly-information model augmented with short-selling frictions—that rationalizes both empirical findings. Then, extending the baseline framework to incorporate sequential signal acquisition, the model generates additional cross-sectional predictions, which we test empirically in the sections that follow.

[9] New-position tests reduce the role of stale holdings, but because holdings are observed quarterly, they do not reveal the exact within-quarter trade date. Appendix A.4 reports a conservative future-return version that measures positions in quarter t and returns in quarter $t+1$. This test moves returns outside the position-formation quarter; short positions continue to earn negative next-quarter DGTW returns, while the new-position differential is weaker.

[10] This outperformance—over 100 bps per quarter ($\approx$ 4% per year)—remains economically meaningful even after accounting for plausible lending fees: 91% of stocks have lending fees under 1% per year (value-weighted the mean $\approx 0.25\%$), and excluding the 5% most expensive issues yields a mean fee of only 85 bps (D'avolio (2002); Blocher and Whaley (2015)).

4.1 Baseline Model

The model has two periods $t \in \{0, 1\}$ and a risky asset with equally likely binary fundamental values $V \in \{\nu_H, \nu_L\}$: the *value gap* is $\Delta \equiv \nu_H - \nu_L > 0$. Competitive noise traders pin down the pre-trade price at the unconditional mean, $P_0 = \frac{\nu_H + \nu_L}{2}$. Before trading, a risk-neutral investor may purchase a costly binary noisy signal ($x \in \{\nu_H, \nu_L\}$) that coincides with the true state with probability $P(x = V) = \frac{1}{2} + \frac{\pi}{2}$, where $\pi \in [0, 1]$ is the chosen precision. Precision is costly: $c(\pi) = \kappa\pi^2$, $\kappa > 0$. After observing the signal, the investor selects position side $s \in \{L \text{ (long)}, S \text{ (short)}, 0\}$ and position size q.

Opening any position incurs fixed monitoring costs $F > 0$, leading the investor to trade a minimal optimal block $q_{\min} > 0$ or abstain.[11] Long positions incur no additional costs ($C_L = F$) while short positions incur a locate fee $\phi > 0$—a one-off fee (sometimes embedded in the borrow rate, sometimes charged explicitly for hard-to-borrow stocks)—and a borrow surcharge $\eta q_{\min} > 0$. ($C_S = F + \phi + \eta q_{\min}$.)

At $t = 0$ the investor chooses precision π, observes the private signal x, updates beliefs via Bayes' rule, and chooses a trade side (s). At $t = 1$ the fundamental V is revealed and the investor's payoff is $\Pi \equiv q_{min}(V - P_0) - C_s - c(\pi)$.

Equilibrium is characterized by a single optimal research intensity, π^*; side-specific precision thresholds $\{\pi_L^{thr}, \pi_S^{thr}\}$ where $\pi_s^{thr} = \frac{2C_s}{q_{\min}\Delta}$; and an initial price P_0, such that (i) π^* maximizes expected profit ex ante; (ii) the investor can trade side $s \in \{L, S\}$ only if $\pi^* \geq \pi_s^{thr}$; and (iii) the uninformed market makers break even: $P_0 = \frac{(\nu_H + \nu_L)}{2}$. This model yields three propositions:[12]

Proposition 1. *Shorts require higher precision thresholds than longs: $\pi_S^{thr} > \pi_L^{thr}$.*

Intuition. Extra short-side fees ($C_S > C_L$) and the minimum-tranche requirement $q_{\min} > 0$ jointly raise the break-even hurdle. Put differently, a short bet must be backed by a more accurate signal before it is worthwhile.

[11] See Appendix A.5.1 for further details on the minimum tranche.

[12] Appendix A.5.1 provides proofs and closed-form expressions.

Proposition 2. *Conditional on a trade, the average precision for shorts is larger than for longs:* $E[\pi^* \mid s = S] > E[\pi^* \mid s = L]$.

Intuition. Shorts only occur in the high-precision equilibrium, whereas long positions are observed in both the high-precision equilibrium ($\pi^* \geq \pi_S^{thr}$) and the long-only equilibrium ($\pi^* < \pi_S^{thr}$).

Proposition 3. *Conditional on a trade, the average absolute abnormal return is larger for shorts than for longs:* $E\big[\,|\alpha|\,\big|\, s = S\big] \;>\; E\big[\,|\alpha|\,\big|\, s = L\big]$.

Intuition. A short is executed only if (i) precision satisfies $\pi^* \geq \pi_S^{thr}$ and (ii) the posterior value gap meets the stricter *value-gap threshold*: $\tau_s = \frac{C_s}{q_{\min}}$, with $\tau_S > \tau_L$. Both forces imply a larger perceived mispricing—and hence a larger realized abnormal return—than for a long.

Together Propositions 1 and 2 deliver the same ordering as our first empirical fact: research intensity is greater around short positions than around comparable long positions. Proposition 3 mirrors our second empirical fact by implying that, once a trade is made, shorts should earn systematically larger absolute abnormal returns than longs.

4.2 Extended Model: Sequential Signals

The baseline model explains why costly short positions can involve greater research intensity and larger absolute abnormal returns. These findings—with the underlying theory and intuition—invite a finer question: within a mutual fund family's short book (or long book), do trades requiring more research earn systematically different returns than those requiring less? At first glance the baseline logic implies a positive link between research effort and returns on either side of the book. But anecdotally, within the set of shorts some may be *clear winners* that seem overpriced after minimal analysis, whereas other trades, *borderline positions*, require extensive research.

To examine this within-side relation we extend the baseline model by allowing for *sequential signals*.[13] All primitives remain as in the baseline model except that now the investor

[13] See Banerjee and Breon-Drish (2020).

acquires information sequentially (i.e., one filing at a time) rather than in a single initial signal. Each signal is an independent observation of the fundamental with additive Gaussian noise: $x_j = V + \varepsilon_j, \;\; \varepsilon_j \sim \mathcal{N}(0, \sigma^2)$. Each draw has precision $\pi_0 = \frac{1}{\sigma^2}$ and costs $\kappa\pi_0^2$. By conjugacy, after N draws the total precision is $N\pi_0$ and the posterior mean is $m_N = \frac{1}{N}\sum_{j=1}^{N} x_j$.

After each draw the investor compares the *posterior dollar mispricing* $|m_N|$ with the side-specific threshold, $\tau_s = \frac{C_s}{q_{\min}}$. The investor can then (i) trade by taking side $s \in \{L, S\}$ at block size $q_{\min}$ if $|m_N| \geq \tau_s$; (ii) continue research by paying $\kappa\pi_0^2$ for another draw if the option value of more information is positive; or (iii) stop research and hold no position if the option value of another draw is negative.[14]

Equilibrium is defined by precision thresholds $\{\tau_L, \tau_S\}$ and initial price P_0 such that (i) the stopping rule described above is optimal given τ_s and P_0; (ii) each threshold τ_s satisfies $\tau_s = \frac{C_s}{q_{\min}}$, so the investor is indifferent at $|m_N| = \tau_s$; and (iii) uninformed market makers break even in expectation by setting $P_0 = \frac{(\nu_H + \nu_L)}{2}$.[15] This extended model delivers an important novel prediction:

Proposition 4. *Expected absolute abnormal returns of executed shorts decrease in the number of signals acquired.*

Intuition. More information draws reduce the posterior's variance such that, in expectation, the posterior crosses the threshold τ_S by a smaller margin. Consequently, on average, requiring more draws yields smaller posterior dollar mispricing and, hence, smaller average absolute returns.

Proposition 4 provides a sharp, novel prediction distinguishing our sequential-sampling model from simpler alternatives: a systematic negative relation between information acquisition and subsequent absolute returns *within short positions.* Intuitively, shorts requiring extensive research are those where investors acquire just enough conviction to trade, yielding

[14]See Appendix A.5.2 for further details.

[15]Appendix A.5.2 shows that these three conditions jointly deliver a unique equilibrium and derives closed-form expressions for the investor's continuation values and the option value of an additional signal. All proofs are also shown in the appendix.

smaller abnormal returns relative to *clear winners* identified after minimal analysis. We test this prediction empirically using our unique data in the next section.

In contrast, the model predicts a much weaker relation for longs. While longs share certain trading costs (F, $q_{\min}$), they do not face short-side frictions (e.g., locate fees, borrowing costs). Consequently, the trading threshold τ_L is lower, making it easier to clear even with weaker initial signals. Thus, longs requiring multiple signals before execution do not differ systematically in realized mispricing compared to those identified quickly, implying a flatter relationship between information acquisition and returns among longs.

5 Empirical Test of Proposition 4

In this section, we assess the key prediction of the sequential-sampling model: conditional on an observed short position, greater research intensity should be associated with smaller absolute abnormal returns. We proceed in two steps. First, we visualize the raw relation between research intensity (*Requests*) and abnormal returns for *ActiveLong* positions versus shorts—and then for *new* longs versus *new* shorts. Second, we estimate a panel regression that tests whether the slope between research intensity and absolute returns is negative for shorts.

Figure 3 plots abnormal returns against contemporaneous filing requests for *ActiveLong* positions (Panel A) and short positions (Panel B). For *ActiveLong* positions, the fitted line is approximately flat while, by contrast, the fitted line slopes slightly upward for shorts, indicating that positions accompanied by more EDGAR requests tend to earn less negative returns (i.e., worse returns). Under a simple "more research, better performance" view, this positive slope among shorts is counterintuitive; yet it is consistent with the model's prediction.

[Figure 3 about here.]

To reduce the role of stale holdings, Figure 4 repeats the scatterplots focusing solely on new long and new short positions. The inverse relation for shorts is more prominent in this setting. For new longs, the raw fitted line is also somewhat negative, suggesting that the long-side relation may be similar to the short-side: positions accompanied by more EDGAR requests tend to earn lower returns.

[Figure 4 about here.]

To formally test these patterns we estimate

$$Return_{ijt} = \alpha + \beta_1 MFIA_{ijt} + \delta Short_{ijt} + \beta_2 Short_{ijt} X MFIA_{ijt} + \lambda Z_{jt} + \gamma_i + u_j + \epsilon_{ijt}, \quad (3)$$

which is similar to Equation 2 but with the inclusion of $MFIA$ and the interaction of $Short$ and $MFIA$. The model also includes mutual fund family and stock fixed effects.

The estimate for $MFIA$ captures the relation between research intensity and returns for ordinary long positions. The linear combination of $MFIA$ and the interaction term ($ShortXMFIA$) captures the short-side slope. In Panel B, $ActiveLong$ and $ActiveLongXMFIA$ allow the slope for benchmark-adjusted $ActiveLong$ positions to differ from the slope for ordinary long positions; the total $ActiveLong$ slope is $MFIA + ActiveLongXMFIA$. Table 4 shows the results.

[Table 4 about here.]

The short-side interaction remains economically important. In Table 4, the total short-side slope ($ShortXMFIA + MFIA$) ranges from about 0.17 to 0.29 across specifications and remains statistically significant. This implies that greater research intensity is associated with less negative—and therefore worse—short-side abnormal returns.

The long-side evidence is weaker and more nuanced. For ordinary long positions, the estimate for $MFIA$ is modest and specification-dependent. For $ActiveLong$ positions, the total slope ($MFIA + ActiveLongXMFIA$) is negative in all four specifications and statistically significant in three of them, but it is materially smaller than the corresponding short-side slope. Thus, the inverse research-return relation is strongest and most consistent for shorts.

We obtain related evidence focusing on new shorts and new longs (Table 5). This new-position analysis helps limit a pure monitoring interpretation: the pattern is not driven only by seasoned shorts that families revisit after those trades move against them. Here, the total short-side slope ($NewShortXMFIA + MFIA$) ranges from 0.42 to 0.62 and remains statistically significant, whereas the slope for new longs is positive. Overall, the negative relation between research intensity and performance is most clearly concentrated on the short side.

Because the data are quarterly, this new-position analysis helps limit, but cannot eliminate, timing and monitoring concerns. Appendix A.4 therefore repeats the specification using lagged $MFIA$ and future returns. In that stricter design, the $NewShortXMFIA$ interaction remains positive in the DGTW specifications, but the total short-side slope ($NewShortXMFIA + MFIA$) is imprecisely estimated. We interpret this as a demanding stress test: lagging requests excludes same-quarter research that may have occurred during position formation, so the appendix results are weaker than, but not inconsistent with, the baseline around-position evidence.

Taken together, these results show that more research is associated with smaller absolute abnormal returns for shorts in the baseline around-position specifications. The corresponding long-side relation is weaker and more specification-dependent. These findings are consistent with Proposition 4: conditional on an observed short position, more research activity is associated with smaller absolute abnormal returns.

[Table 5 about here.]

6 Alternative Explanation and Additional Evidence

The inverse research-return slope for executed shorts is intuitive following the sequential-sampling model, but there is a competing interpretation for the empirical finding. Perhaps extra EDGAR filings simply reflect more time spent on *public* information that has little incremental value, so the inverse relation reflects wasted effort rather than the clear-winner-versus-borderline logic of the model. We can examine this alternative empirically by asking whether heavier EDGAR research correlates with poorer overall performance at the family level. We can also examine whether observable EDGAR research is lower for potential shorts that look like "clear winners" ex ante, as the sequential-signal framework predicts.

6.1 Is public-document research wasteful?

Could the inverse relation between research and absolute returns among executed shorts reflect wasted effort? If so, fund families that generally do more public-document research should underperform. We test this by aggregating our sample to family-quarter observations and using the following regression model to explain family-level benchmark-adjusted returns:

$$Return_{it} = \alpha + \beta MFIA_{it} + \lambda X_{it} + \omega_t + \epsilon_{it}, \tag{4}$$

where $Return_{it}$ is the average benchmark-adjusted return across the actively managed funds of family i during quarter t and X_{it} includes various controls.[16] The variable of interest, $MFIA_{it}$, measures total requests for all stocks by mutual fund family i during quarter t.[17]

Table 6 shows the results using contemporaneous and lagged measures of EDGAR activity, and finds that the research-return slope is *positive*: fund families with more EDGAR research have better returns. The estimates suggest that if a family doubled its research

[16] Control variables include family size ($TNAM$), total fund flows over the last year ($Netflow12$), average portfolio concentration (HHI), average expense ratios ($ExpenseRatio$), and average portfolio turnover ($Turnover$).

[17] $MFIA_{it}$ aggregates $MFIA_{ijt}$ across all stocks for family i during quarter t.

efforts it would improve next-quarter's benchmark alpha by roughly 1.4 bp—small in percentage terms but $7 million for a median $50 billion family. Thus, EDGAR research appears productive on average, which weighs against the view that the inverse within-short slope simply reflects wasteful public-document research.

[Table 6 about here.]

6.2 Evidence with Clear-Winners

The evidence in Section 5 shows that for executed short positions, more research is related to lower absolute returns. These findings support the intuition behind Proposition 4: shorts requiring extensive research are *borderline shorts*, whereas those requiring less research are *clear winners*. This section provides further support for Proposition 4 by asking whether stocks that look like clear-winner shorts *ex ante* are the same stocks that require less research.

First, we show that overpriced stocks—identified *ex ante* using indicators based on the mispricing measure of Stambaugh and Yuan (2017)—earn low returns. Table 7 highlights this result, showing the results from testing the following regression model:

$$Return_{jt} = \alpha + \beta_1 Underpriced_{jt-1} + \beta_2 Overpriced_{jt-1} + \lambda Z_{jt} + \epsilon_{jt}, \tag{5}$$

where $Return_{jt}$ is either the excess return of stock j during quarter t ($ExRet$) or the characteristic-adjusted return ($DGTW$).[18]

[Table 7 about here.]

Second, we combine the *ex ante* overpricing indicator with an indicator for low uncertainty: we classify an overpriced stock with low disagreement as a clear-winner proxy for

[18] The variables of interest are *Underpriced* and *Overpriced*, which are indicators based on the mispricing measure of Stambaugh and Yuan (2017).

shorts.[19] Then for new short positions we estimate the following model:[20]

$$\begin{aligned} MFIA_{ijt} = {} & \beta_1 LowDisagree_{jt-1} + \beta_2 Overpriced_{jt-1} \\ & + \beta_3 LowDisagreeXOverpriced_{jt-1} + \gamma_i + u_j + \omega_t + \epsilon_{ijt}. \end{aligned} \tag{6}$$

For new short positions, the interaction of *LowDisagree* and *Overpriced* proxies for clear winners, since the stock has low uncertainty and the overpricing measure predicts negative future returns.[21] The model always includes mutual fund family and quarter fixed effects and includes stock fixed effects where noted. Table 8 shows the results.

[Table 8 about here.]

In the case of new long positions, there is a weak relation or no relation between research intensity and clear-winner proxies. This result is consistent with Proposition 4 and findings from Tables 4 and 5.

For new short positions, where *LowDisagreeXOverpriced* proxies for clear winners, the estimates are consistent with Proposition 4: mutual fund families do less observable research for clear-winner shorts. Column 5 shows that mutual fund families make 8% fewer requests for overpriced stocks with low disagreement compared to other stocks. When including stock-fixed effects, families make up to 12% fewer requests when stocks are overpriced and have low disagreement. Taken together, the evidence supports the sequential-sampling logic of Proposition 4: observable research is lower for shorts whose ex ante characteristics point to an obvious mispricing.

[19]To proxy for low uncertainty we define *LowDisagree* as equal to one for each stock-quarter observation with *Disagreement* in the lowest quintile.

[20]We estimate the analogous model for new long positions.

[21]Combining *LowDisagree* with *Underpriced* proxies for clear winners in the case of longs.

7 Conclusion

This paper links the holdings of U.S. equity mutual fund families with observations of their research behavior from the EDGAR Log Files and, in doing so, observes how mutual fund families allocate research effort for both long and short positions. Two empirical facts stand out: (i) mutual fund shorts outperform their longs by roughly 100 bps per quarter, a spread that remains economically large after plausible lending fees, and (ii) mutual fund families devote more observable pre-trade research to their shorts, especially relative to benchmark-adjusted active longs.

Then, using a parsimonious endogenous information model with short-specific frictions, we reconcile these facts while generating a novel insight about research, executed trades, and performance. The model predicts that short positions preceded by more research should deliver *smaller* absolute alpha. The data support this inverse slope between research and returns for shorts—and a weaker, more specification-dependent relation for longs—consistent with the model's clear-winner-versus-borderline-trade logic. Additional timing tests in the appendix use future returns and, for the new-position research-return specification, lagged requests to clarify the quarterly timing of the evidence.

By showing that information gathering—not just trade execution—is an important margin on which short-selling frictions operate, we sharpen the classic narrative about costly arbitrage. We also show that EDGAR research appears productive on average when investors ration attention and research efforts. Further research in this area should seek to better understand the factors that drive information acquisition decisions for other investor types and for other types of trades. Investigating the impact of information acquisition on other aspects of portfolio management and performance could also provide valuable insights for investors and financial professionals.

References

Almazan, A., Brown, K. C., Carlson, M., and Chapman, D. A. (2004). Why constrain your mutual fund manager? *Journal of Financial Economics*, 73(2):289–321.

An, L., Huang, S., Lou, D., and Shi, J. (2021). Why don't most mutual funds short sell? *Available at SSRN 3813790*.

Asquith, P., Pathak, P. A., and Ritter, J. R. (2005). Short interest, institutional ownership, and stock returns. *Journal of Financial Economics*, 78(2):243–276.

Banerjee, S. and Breon-Drish, B. (2020). Strategic trading and unobservable information acquisition. *Journal of Financial Economics*, 138(2):458–482.

Ben-Rephael, A., Da, Z., and Israelsen, R. D. (2017). It depends on where you search: Institutional investor attention and underreaction to news. *The Review of Financial Studies*, 30(9):3009–3047.

Blocher, J. and Whaley, R. E. (2015). Passive investing: The role of securities lending. *Working Paper Series*.

Boehmer, E., Jones, C. M., and Zhang, X. (2008). Which shorts are informed? *The Journal of Finance*, 63(2):491–527.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of finance*, 52(1):57–82.

Chen, H., Cohen, L., Gurun, U., Lou, D., and Malloy, C. (2020a). Iq from ip: Simplifying search in portfolio choice. *Journal of Financial Economics*, 138(1):118–137.

Chen, H., Desai, H., and Krishnamurthy, S. (2013). A first look at mutual funds that use short sales. *Journal of Financial and Quantitative Analysis*, 48(3):761–787.

Chen, Y., Kelly, B., and Wu, W. (2020b). Sophisticated investors and market efficiency: Evidence from a natural experiment. *Journal of Financial Economics*, 138(2):316–341.

Crane, A., Crotty, K., and Umar, T. (2022). Hedge funds and public information acquisition. *Management Science*.

Da, Z., Engelberg, J., and Gao, P. (2011). In search of attention. *The journal of finance*, 66(5):1461–1499.

Daniel, K., Grinblatt, M., Titman, S., and Wermers, R. (1997). Measuring mutual fund performance with characteristic-based benchmarks. *The Journal of finance*, 52(3):1035–1058.

D'avolio, G. (2002). The market for borrowing stock. *Journal of Financial Economics*, 66(2-3):271–306.

Engelberg, J. E., Reed, A. V., and Ringgenberg, M. C. (2018). Short-selling risk. *The Journal of Finance*, 73(2):755–786.

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

Gibbons, B. (2023). Public market information and venture capital investment. *Journal of Financial and Quantitative Analysis*, 58(2):746–776.

Gibbons, B., Iliev, P., and Kalodimos, J. (2021). Analyst information acquisition via edgar. *Management Science*, 67(2):769–793.

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

Hellwig, M. F. (1980). On the aggregation of information in competitive markets. *Journal of economic theory*, 22(3):477–498.

Hong, H. and Stein, J. C. (2003). Differences of opinion, short-sales constraints, and market crashes. *The Review of Financial Studies*, 16(2):487–525.

Jones, C. M. and Lamont, O. A. (2002). Short-sale constraints and stock returns. *Journal of Financial Economics*, 66(2-3):207–239.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, pages 1315–1335.

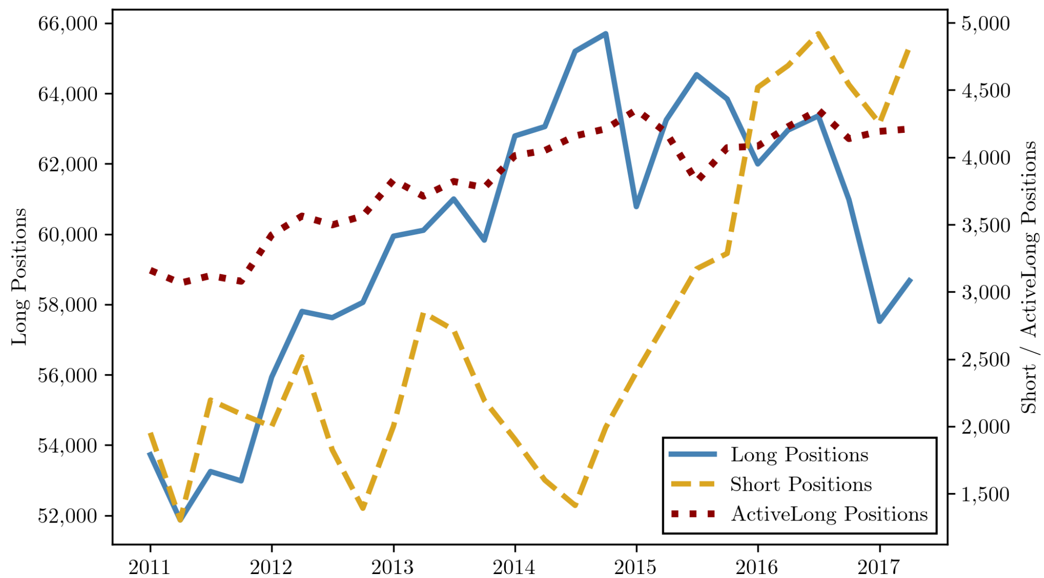
Lorden, G. (1970). On Excess Over the Boundary. *The Annals of Mathematical Statistics*, 41(2):520 – 527.

Seneca, J. J. (1967). Short interest: bearish or bullish? *The Journal of Finance*, 22(1):67–70.

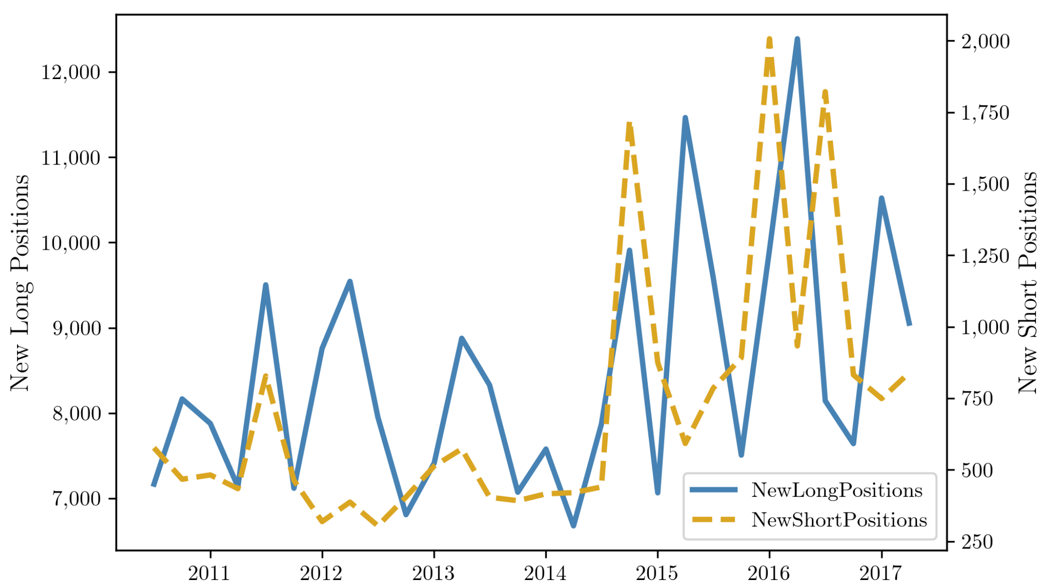
Sirri, E. R. and Tufano, P. (1998). Costly search and mutual fund flows. *The journal of finance*, 53(5):1589–1622.

Stambaugh, R. F. and Yuan, Y. (2017). Mispricing factors. *The review of financial studies*, 30(4):1270–1315.

Verrecchia, R. E. (1982). Information acquisition in a noisy rational expectations economy. *Econometrica: Journal of the Econometric Society*, pages 1415–1430.

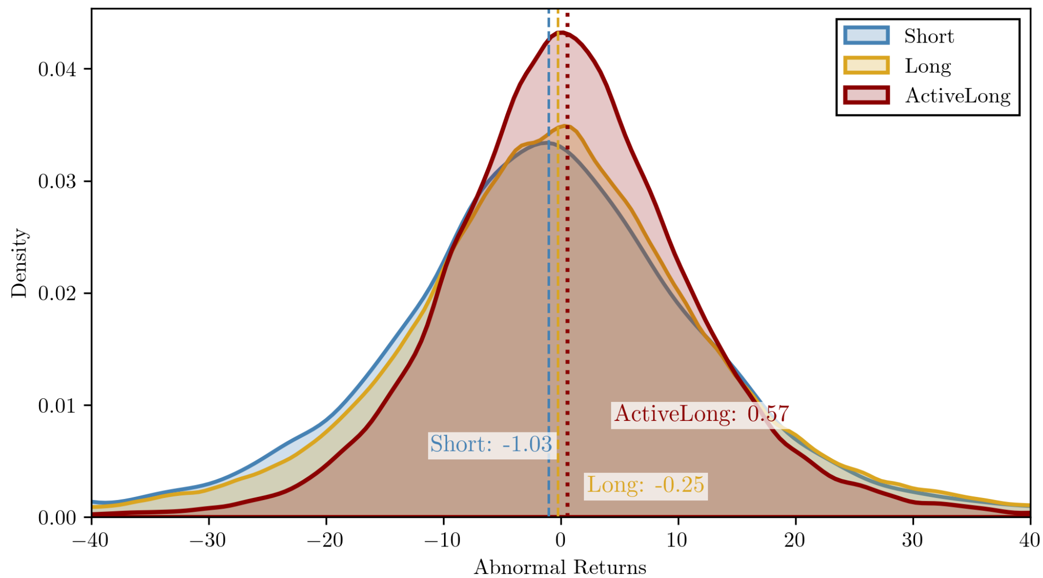


(A) All Long, ActiveLong, and Short Positions

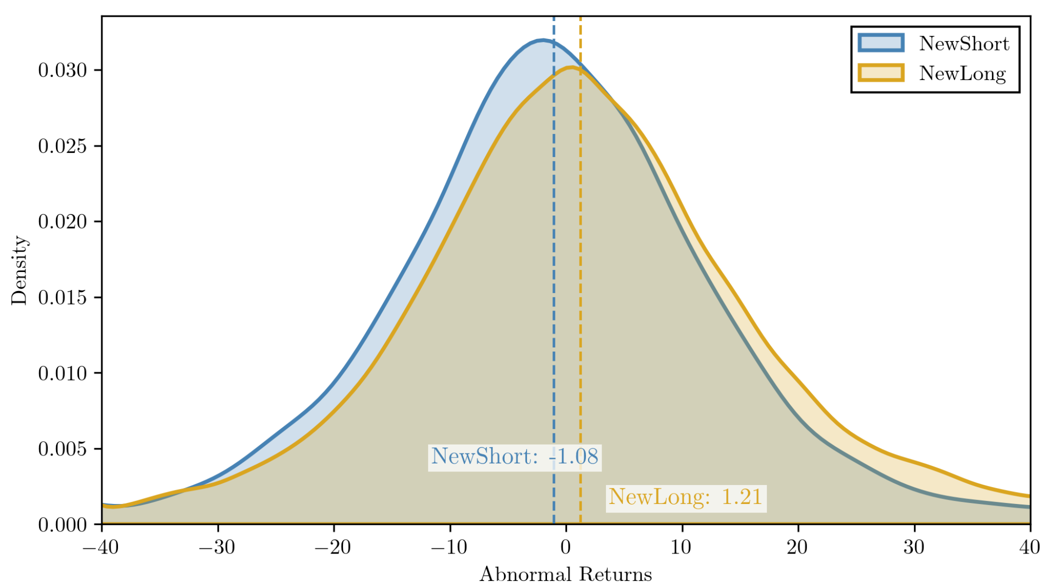


(B) New Long and New Short Positions

Figure 1: Time Series of Long, ActiveLong, and Short Positions
These time series plots show the changes in positions held by mutual funds over time. Panel A shows all long, ActiveLong, and short positions while Panel B shows the changes in new long and new short positions.

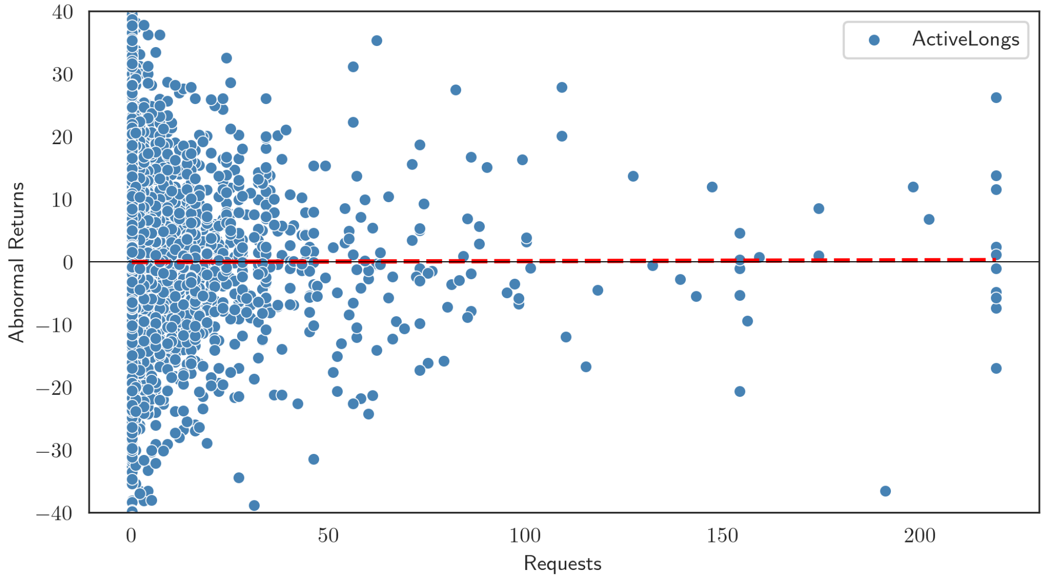


(A) All Long, ActiveLong, and Short Positions

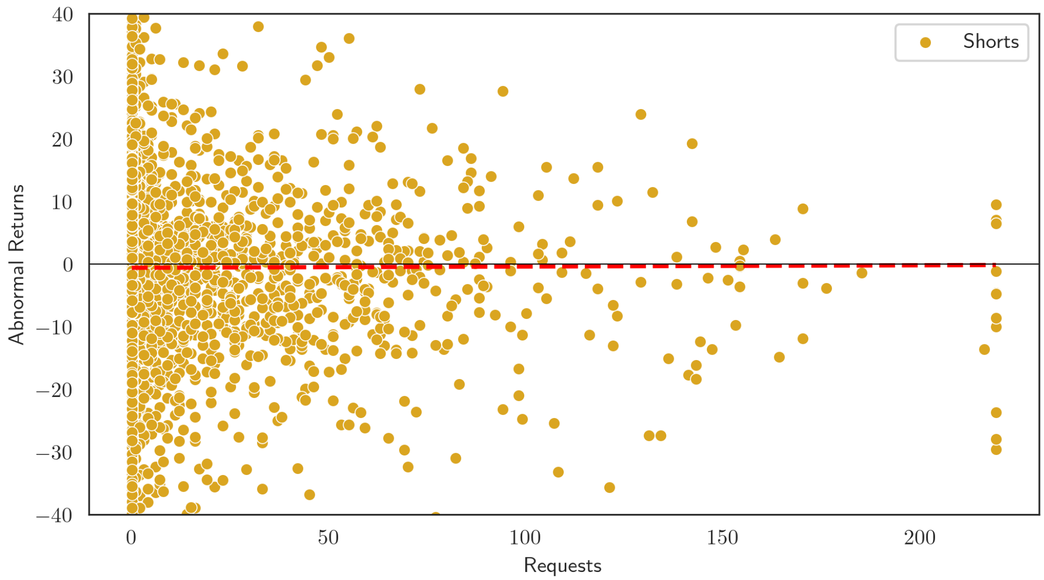


(B) New Long and New Short Positions

Figure 2: Histograms of Returns by Position
This figure shows the histogram of abnormal returns, as measured using equally-weighted DGTW returns, for long, ActiveLong, and short positions (Panel A) and for new long and new short positions (Panel B).

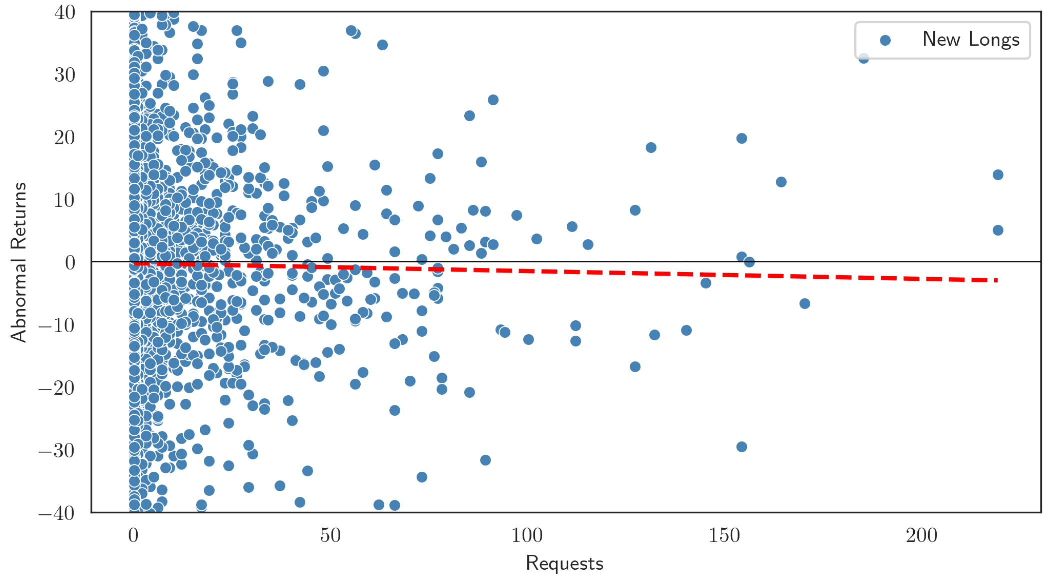


(A) All ActiveLong Positions

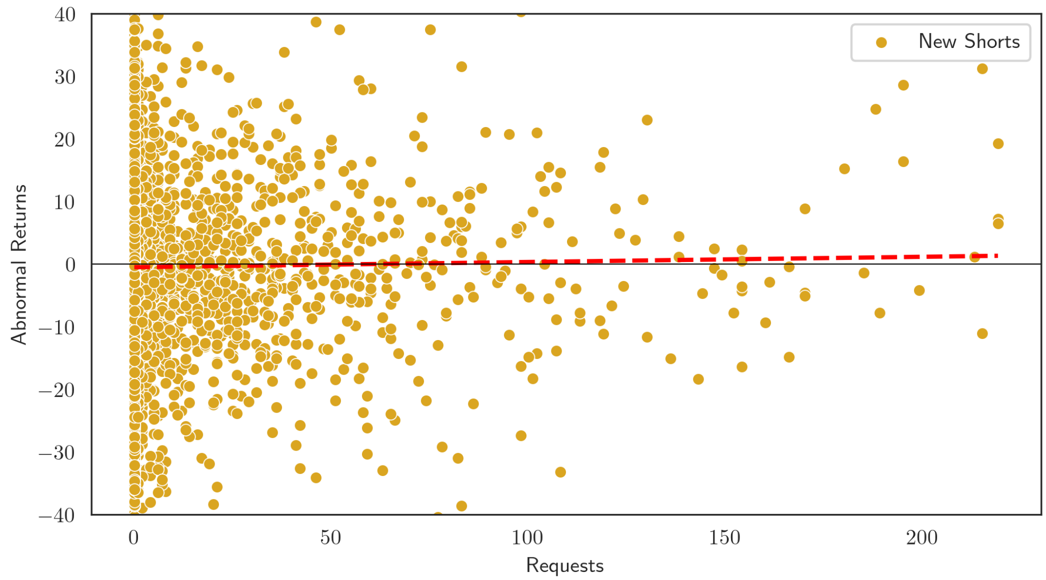


(B) All Short Positions

Figure 3: Scatter Plot of Requests and Abnormal Returns
This figure shows scatter plots of abnormal returns, as measured using equally-weighted DGTW returns, with requests for both ActiveLong positions (Panel A) and short positions (Panel B).



(A) New Long Positions



(B) New Short Positions

Figure 4: Scatter Plot of Requests and Abnormal Returns: New Positions
This figure shows scatter plots of abnormal returns, as measured using equally-weighted DGTW returns, with requests for both new long positions (Panel A) and new short positions (Panel B).

Table 1: Summary Statistics

This table provides summary statistics for our sample. The subscripts for each variable refer to mutual fund family i, stock j, and quarter t. The entire panel includes $2,575,774$ observations and eight variables are summarized over the entire panel. This table also provides summary statistics after collapsing the panel at the mutual fund family-by-quarter level (subscripts it) for which there are $1,159$ observations, and after collapsing the panel at the stock-by-quarter level (subscripts jt) for which there are $113,885$ observations for most of the variables. The period covers 30 quarters, from January 2010 through June 2017, and 115 different mutual fund families, of which 55 have at least one short position in our sample.

(1)	(2)	(3)	(4)	(5)
Variable	Mean	Std. Dev.	Median	N
Requests_{ijt}	2.46	12.20	0	2,575,774
$\mathbb{1}\{\text{Requests}_{ijt} \geq 1\}$	0.18	0.38	0	2,575,774
Long_{ijt}	0.49	0.50	0	2,575,774
Short_{ijt}	0.03	0.16	0	2,575,774
NewLong_{ijt}	0.04	0.20	0	2,575,774
NewShort_{ijt}	0.00	0.06	0	2,575,774
ActiveLong_{ijt}	0.02	0.15	0	2,575,774
$\text{BMAdjPortfolioWeight}_{ijt}$	0.03	0.14	0.00	2,575,774
Requests_{it}	8,919	21,525	1,611	1,159
$\mathbb{1}\{\text{Requests}_{it} \geq 1\}$	0.99	0.08	1	1,159
TNAM_{it} (*millions*)	156,152	356,936	50,740	1,159
Funds_{it}	71	61	57	1,159
Stocks_{it}	185	177	111	1,159
Turnover_{it}	56	33	52	1,150
HHI_{it}	79	79	51	1,159
$\text{Expense Ratio}_{it}$	1.15	0.32	1.20	1,150
Netflow12_{it}	0.002	0.019	0.001	1,153
MCAP_{jt} (*millions*)	5,600	21,341	824	113,885
Volume_{jt} (*thousands*)	1,209	2,788	317	113,885
ShortInt_{jt}	0.05	0.05	0.03	85,699
Mispricing_{jt}	50	13	50	63,737
Disagreement_{jt}	0.26	0.39	0.13	97,393
IdioVol_{jt}	0.02	0.01	0.02	85,039
News_{jt}	169	412	75	79,110
InstOwn_{jt} (*millions*)	95	241	26	62,078

Table 2: Research Intensity and Positions

This table shows the results from estimating Equation 1. The dependent variable, $MFIA_{ijt}$, is either one quarter of requests ($Requests_{ijt}$) or two quarters of requests ($Requests_{ijt,t-1}$). In Panel A, the independent variables indicate whether family i has a long or short position in stock j in quarter t. Panel B replaces *Long* with *ActiveLong* defined as 33 basis points above the benchmark portfolio weight. In Panel C, the independent variables indicate whether family i has a *new* long or *new* short position in stock j in quarter t. In Columns 3 and 6, $MFIA$ is lagged one quarter while in Columns 4 and 7, $MFIA$ is lagged two quarters. The model includes family, stock, and quarter fixed effects. Standard errors are clustered by family-stock and quarter and are shown in parentheses. Indicators ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
	MFIA = Requests_{ijt}			MFIA = $\text{Requests}_{ijt,t-1}$		
	MFIA	L1.MFIA	L2.MFIA	MFIA	L1.MFIA	L2.MFIA
Panel A. Long or Short Positions.						
Long	0.12***	0.11***	0.09***	0.17***	0.15***	0.13***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Short	0.18***	0.18***	0.16***	0.22***	0.21***	0.19***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
$H_0 : Short - Long = 0$	0.06**	0.07**	0.06**	0.05*	0.06*	0.06*
	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
Obs.	2,528,011	2,528,011	2,528,011	2,528,011	2,528,011	2,528,011
R^2	0.43	0.43	0.44	0.48	0.49	0.49
Panel B. ActiveLong or Short Positions.						
ActiveLong	0.07***	0.07***	0.06***	0.11***	0.10***	0.09***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Short	0.18***	0.17***	0.17***	0.20***	0.20***	0.20***
	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)
$H_0 : Short - ActiveLong = 0$	0.10***	0.11***	0.11***	0.09**	0.09**	0.10***
	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)
Obs.	2,573,701	2,573,701	2,573,701	2,573,701	2,573,701	2,573,701
R^2	0.44	0.45	0.45	0.49	0.49	0.50
Panel C. New Long or New Short Positions.						
NewLong	0.08***	0.01	-0.00	0.08***	0.01	-0.00
	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)
NewShort	0.09***	0.08***	0.07**	0.10***	0.09**	0.08**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
$H_0 : NewShort - NewLong = 0$	0.01	0.07**	0.07**	0.02	0.08**	0.08**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Obs.	2,575,774	2,575,774	2,575,774	2,575,774	2,575,774	2,575,774
R^2	0.44	0.45	0.45	0.49	0.49	0.50

Table 3: Positions and Returns

This table shows the results from estimating Equation 2 using long and short positions, using ActiveLong positions, and on new positions. In Panel A, the dependent variable is the excess return of stock j during quarter t ($ExRet$) and the independent variables (not shown) include quarterly risk-factors from the four-factor model (Fama and French (1993) and Carhart (1997)). In Panel B, the dependent variable is the characteristics-adjusted return ($DGTW$). Columns 2 and 3 report α as the average abnormal return of the stocks in the sample (either the sample of long positions or the sample of short positions). Columns 4 and 5 use the full sample of positions, thus α measures the abnormal return to stocks in long positions while $Short$ measures the additional return to stocks held in short positions and $ActiveLong$ measures the additional return to long positions in excess of 33 basis points above the benchmark portfolio weight. Column 6 only includes new positions, either newly long or short, thus α measures the abnormal return to new longs while $NewShort$ measures the additional return to new shorts. The model includes a mutual fund family fixed effect. Standard errors are clustered by stock and quarter and are shown in parentheses. Indicators ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively.

(1)	(2)	(3)	(4)	(5)	(6)
	Long Only	Short Only	All Positions		
Panel A. Independent Variable: Abnormal Return as ExRet					
MktReturn	1.02***	1.03***	1.02***	1.02***	1.06***
	(0.03)	(0.07)	(0.03)	(0.03)	(0.04)
α	0.28*	-0.86***	0.29*	0.25*	2.58***
	(0.14)	(0.29)	(0.14)	(0.14)	(0.18)
Short			-1.28***	-1.23***	
			(0.33)	(0.33)	
ActiveLong				0.86***	
				(0.30)	
NewShort					-3.53***
					(0.41)
Obs.	1,201,323	23,452	1,224,775	1,224,775	112,253
R^2	0.17	0.18	0.17	0.17	0.17
Panel B. Independent Variable: Abnormal Return as DGTW					
MktReturn	0.03*	0.08**	0.03*	0.03*	0.09***
	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)
α	0.12	-1.00***	0.12	0.08	2.28***
	(0.12)	(0.15)	(0.12)	(0.12)	(0.20)
Short			-1.07***	-1.03***	
			(0.23)	(0.23)	
ActiveLong				0.79***	
				(0.12)	
NewShort					-3.34***
					(0.35)
Obs.	967,808	19,617	987,425	987,425	85,444
R^2	0.00	0.01	0.00	0.00	0.01

Table 4: Long and Short Positions, Returns, and Research Intensity

This table shows the results from estimating Equation 3. In Columns 2 and 4 the independent variable is the excess return of stock j during quarter t (*ExRet*) and independent variables (not shown) include risk-factors (Fama and French (1993) and Carhart (1997)). Characteristics-adjusted returns (*DGTW*) are used in Columns 3 and 5. In Panel A, *Long* measures abnormal returns to long positions while *Short* measures additional returns to short positions. $MFIA$ measures the effect of research intensity on stock returns for long positions while *ShortXMFIA* measures the additional effect on short position returns. $ShortXMFIA + MFIA$ measures the total effect of research intensity on short position returns. Panel B includes *ActiveLong* and corresponding interactions with $MFIA$. The model includes mutual fund family and stock fixed effects. Standard errors are clustered by stock and quarter and are shown in parentheses. Indicators ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively. Observations and R^2 are approximately identical between panels and are shown once.

(1)	(2)	(3)	(4)	(5)
	MFIA = Requests_{ijt}		MFIA = $\text{Requests}_{ijt,t-1}$	
	ExRet	DGTW	ExRet	DGTW
Panel A. All positions				
Long	0.22	0.07	0.27*	0.11
	(0.15)	(0.10)	(0.15)	(0.10)
Short	-1.01***	-0.84***	-1.05***	-0.90***
	(0.25)	(0.19)	(0.25)	(0.19)
MFIA	0.05	0.08**	-0.04*	-0.00
	(0.03)	(0.03)	(0.02)	(0.02)
ShortXMFIA	0.22**	0.21**	0.21***	0.21***
	(0.09)	(0.09)	(0.07)	(0.06)
H_0 :ShortXMFIA + MFIA = 0	0.27***	0.28***	0.17**	0.21***
	(0.09)	(0.09)	(0.07)	(0.06)
Panel B. All positions including ActiveLong covariate.				
Long	0.19	0.04	0.24	0.08
	(0.15)	(0.10)	(0.15)	(0.10)
ActiveLong	0.60***	0.55***	0.60***	0.54***
	(0.09)	(0.07)	(0.09)	(0.07)
Short	-0.98***	-0.81***	-1.02***	-0.88***
	(0.25)	(0.19)	(0.25)	(0.19)
MFIA	0.06*	0.09**	-0.03	0.00
	(0.03)	(0.03)	(0.02)	(0.02)
ShortXMFIA	0.21**	0.20**	0.20***	0.21***
	(0.09)	(0.09)	(0.07)	(0.06)
ActiveLongXMFIA	-0.16***	-0.13***	-0.10**	-0.08**
	(0.05)	(0.05)	(0.04)	(0.04)
H_0 :ShortXMFIA + MFIA = 0	0.27***	0.29***	0.17**	0.21***
	(0.09)	(0.09)	(0.07)	(0.06)
H_0 :ActiveLongXMFIA + MFIA = 0	-0.10*	-0.04	-0.13***	-0.08**
	(0.05)	(0.04)	(0.04)	(0.04)
Obs.	1,224,678	987,374	1,224,678	987,374
R^2	0.22	0.07	0.22	0.07

Table 5: New Long and New Short Positions, Returns, and Research Intensity

This table shows the results from estimating Equation 3 on new positions. In Columns 2 and 4 the independent variable is the excess return of stock j during quarter t ($ExRet$) and independent variables (not shown) include risk-factors (Fama and French (1993) and Carhart (1997)). Characteristics-adjusted returns ($DGTW$) are used in Columns 3 and 5. In the results below, $NewLong$ measures the abnormal return to stocks in long positions while $NewShort$ measures the additional return to stocks held in short positions. Further, $MFIA$ measures the effect of research intensity on stocks held in long positions while the interaction term, $NewShortXMFIA$, measures the additional effect of research intensity on stocks in new short positions. Thus, $NewShortXMFIA + MFIA$ measures the total effect of research intensity on the returns of stocks in new short positions. The model includes mutual fund family and stock fixed effects. Standard errors are clustered by stock and quarter and are shown in parentheses. Indicators ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively.

(1)	(2)	(3)	(4)	(5)
	All Positions			
	MFIA = Requests$_{ijt}$		MFIA = Requests$_{ijt,t-1}$	
	ExRet	DGTW	ExRet	DGTW
NewLong	2.41***	2.10***	2.45***	2.15***
	(0.15)	(0.13)	(0.17)	(0.14)
NewShort	-2.96***	-2.81***	-3.02***	-2.90***
	(0.40)	(0.39)	(0.40)	(0.39)
MFIA	0.35***	0.39***	0.16*	0.17**
	(0.10)	(0.10)	(0.09)	(0.07)
NewShortXMFIA	0.25	0.23	0.25	0.27
	(0.23)	(0.20)	(0.18)	(0.16)
H_0 :NewShortXMFIA + MFIA = 0	0.61***	0.62***	0.42**	0.44**
	(0.22)	(0.21)	(0.17)	(0.17)
Obs.	111,706	85,133	111,706	85,133
R^2	0.25	0.10	0.25	0.10

Table 6: Research Intensity and Benchmark-Adjusted Returns

This table shows results from estimating Equation 4 on the quarterly sample of mutual fund family returns. The dependent variable is the average benchmark-adjusted return across the funds of family i during quarter t. The independent variable of interest, $MFIA_{it}$ is the number of requests made by family i during quarter t. The table uses $MFIA$ from the same quarter as when the returns are earned and in the previous four quarters. The other independent variables are described in Appendix A.2. The model includes quarter fixed effects. Standard errors are clustered by family and quarter and are shown in parentheses. Indicators ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively.

(1)	(2)	(3)	(4)	(5)	(6)
	L0.MFIA	L1.MFIA	L2.MFIA	L3.MFIA	L4.MFIA
	Return	Return	Return	Return	Return
MFIA	0.016*	0.014*	0.013*	0.011*	0.008
	(0.008)	(0.008)	(0.007)	(0.006)	(0.005)
TNAM	0.003	0.004	0.004	0.004	0.005
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
NetFlow12	-0.259	-0.227	-0.188	-0.138	-0.071
	(1.150)	(1.121)	(1.124)	(1.121)	(1.144)
HHI	-0.071***	-0.070***	-0.070***	-0.069***	-0.068***
	(0.016)	(0.015)	(0.015)	(0.015)	(0.015)
ExpenseRatio	-0.090***	-0.089***	-0.089***	-0.088***	-0.087***
	(0.025)	(0.024)	(0.025)	(0.025)	(0.026)
Turnover	0.041	0.041	0.041	0.041	0.042
	(0.034)	(0.034)	(0.034)	(0.034)	(0.035)
Obs.	1,150	1,150	1,150	1,150	1,150
R^2	0.09	0.09	0.09	0.09	0.09
Quarter FE	Yes	Yes	Yes	Yes	Yes

Table 7: Mispricing and Returns

This table shows the results from estimating Equation 5 on the quarterly sample of stock returns. In Columns 2 through 4, the dependent variable is the excess return of stock j during quarter t ($ExRet$) and the independent variables (suppressed in the output) include quarterly risk-factors from the four-factor model (Fama and French (1993) and Carhart (1997)). In Columns 5 through 7, the dependent variable is the characteristics-adjusted return ($DGTW$) from (Daniel et al. (1997)). The indicator variable $L1.Underpriced$ equals one if stock j's mispricing score was in the bottom quintile in quarter $t-1$. The indicator variable $L1.Overpriced$ equals one if stock j's mispricing score was in the top quintile in quarter $t-1$. Standard errors are clustered by stock and quarter and are shown in parentheses. Indicators ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
	ExRet	ExRet	ExRet	DGTW	DGTW	DGTW
L1.Underpriced	0.22	0.56		0.34	0.61**	
	(0.34)	(0.41)		(0.24)	(0.27)	
L1.Overpriced	-1.34**		-1.39**	-1.10**		-1.19**
	(0.60)		(0.63)	(0.43)		(0.44)
Obs.	75,218	75,218	75,218	72,648	72,648	72,648
R^2	0.18	0.18	0.18	0.00	0.00	0.00

Table 8: Research Intensity and Clear-Winner Proxies

This table shows results from estimating Equation 6 on new positions. The dependent variable, $MFIA$, is the number of requests from family i about stock j in quarter t. The independent variables $LowDisagree$, $Underpriced$, $Overpriced$, and their interactions identify whether stocks are clear winners in a given quarter. The model always includes mutual fund family and quarter fixed effects and includes stock fixed effects where noted. Standard errors are clustered by family-stock and quarter and are shown in parentheses. Indicators ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
	New Long Only			New Short Only		
	MFIA	MFIA	MFIA	MFIA	MFIA	MFIA
LowDisagreeXUnderpriced	-0.04***	0.00	0.00			
	(0.01)	(0.01)	(0.01)			
LowDisagree			0.00			-0.00
			(0.01)			(0.03)
Underpriced			0.00			
			(0.01)			
Overpriced						-0.00
						(0.02)
LowDisagreeXOverpriced				-0.08**	-0.12**	-0.11*
				(0.04)	(0.05)	(0.06)
Obs.	110,887	110,770	110,770	13,600	13,125	13,125
R^2	0.45	0.48	0.48	0.63	0.66	0.66
Family FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Stock FE	No	Yes	Yes	No	Yes	Yes

Online Appendix

This appendix provides additional details to supplement the main text. There are five sections of this appendix. Appendix A.1 contains details about the EDGAR Log Files and the process of unmasking mutual fund IP addresses. Appendix A.2 contains a table detailing the variables used in this paper. Appendix A.3 reports rich fixed-effects robustness tests for the research-intensity regressions. Appendix A.4 reports conservative timing robustness tests. Appendix A.5 contains model details and extended proofs.

A.1 Unmasking IP Addresses in EDGAR Log Files

The EDGAR Log Files contain billions of observations of "requests" or "requests to view a filing." Each observation details the filing requested (accession number), the date and time of the request, and the requester (the IP address making the electronic request). A snapshot of the raw EDGAR Log Files is shown below:

IP Address	Date	Time	CIK	Accession Number
38.97.91.ecg	20170531	09:47:33	051143	000104746917001061
38.65.241.fhf	20170531	11:07:28	274191	000002741917000008
67.199.249.igg	20170624	12:27:02	320193	000032019317000009
216.223.41.aah	20170624	16:12:55	831259	000083125917000016

Given our focus in this paper on the requester, linking the masked IP addresses to identifiable investors (e.g., mutual fund families) is pivotal to our study. To unmask the IP addresses, we first notice the fourth octet in the examples above.[1] In place of the actual digits of the requesting IP address, the fourth octet is reported as a set of three letters. However, organizations typically register blocks of IP addresses, with the most common block fixing the first three octets and containing all 256 versions of the fourth octet.[2] In other words, only the first three octets are necessary to identify the organization that has registered that block of IP addresses.

Using this insight, we searched historical IP address registration records from 2010 through 2017 to identify the blocks of IP addresses registered to investment firms.[3] Then, using this hand-collected mapping between investment firms and IP addresses, we unmask the requesters in the EDGAR Log Files. As a result, the snapshot of raw data from above has been transformed into the following.

[1]For IP addresses, an *octet* is a group of eight bits, or the one to three digit numbers (from 0 to 255) separated by periods in the examples above.

[2]For example, all 256 IP addresses beginning with 38.97.91 will be registered to the same organization.

[3]IP registration records were acquired from MaxMind, https://www.maxmind.com/en/home.

Investment Firm	Date	Time	Ticker	Filing
Abrams Capital	20170531	09:47:33	IBM	10-K for 2016
Harbor Capital	20170531	11:07:28	TGT	10-K for 2016
Crabel Capital	20170624	12:27:02	AAPL	10-Q for Q2 2017
Ronin Capital	20170624	16:12:55	FCX	Earnings for Q2 2017

Furthermore, the three letters used to mask the fourth octet is static, not dynamic. This means, for example, that def replaces the digits 146 for every instance of 146. This allows us to identify unique IP addresses. In other words, though an unmasked mutual fund may make 50 requests one day, we can observe how many different IP addresses made those requests.

Finally, we have adjusted the data to remove likely bots. As mentioned, the raw EDGAR Log Files contain billions of requests with many thousands of requests per day coming from single IP addresses. It is unlikely that these thousands of requests per day represent a human actually clicking on documents in EDGAR. It is much more likely that they represent computer programs (bots) downloading large quantities of data at a time. Given these IP addresses do not fit with the spirit of our research, we remove them from the data. The removal process is as follows: we remove IP addresses that either (i) make over $1,000$ requests in a day or (ii) make requests for over 100 different CIKs (i.e., firms).

A.2 Variable Details

Variable	Description
Information Acquisition (MFIA)	Measured at mutual fund family-stock-quarter (ijt) level.
Requests_{ijt}	The sum of the number of requests made by mutual fund family i for filings about stock j in quarter t. The sum is winsorized at the 99th percentile and scaled by adding one and taking the natural log. This data is primarily derived from the EDGAR Log Files.
$\text{Requests}_{ijt,t-1}$	The sum of the number of requests made by mutual fund family i for filings about stock j in quarter t and in quarter $t-1$. The sum is winsorized at the 99th percentile and scaled by adding one and taking the natural log. This data is primarily derived from the EDGAR Log Files.
Position Variables	Measured at mutual fund family-stock-quarter (ijt) level.
Long_{ijt}	Indicator equal to 1 if any of the mutual funds within mutual fund family i hold shares of stock j in a long position at the end of quarter t. We exclude all observations where mutual fund family i holds long and short positions of stock j during the same quarter. This data is derived from the CRSP Mutual Funds database.
Short_{ijt}	Indicator equal to 1 if any of the mutual funds within mutual fund family i hold shares of stock j in a short position at the end of quarter t. We exclude all observations where mutual fund family i holds long and short positions of stock j during the same quarter. This data is derived from the CRSP Mutual Funds database.
NewLong_{ijt}	Indicator equal to 1 if $Long_{ijt} = 1$ and $Long_{ijt-1} = 0$.
NewShort_{ijt}	Indicator equal to 1 if $Short_{ijt} = 1$ and $Short_{ijt-1} = 0$.
$\text{BMAdjPortfolioWeight}_{ijt}$	Benchmark-adjusted family-stock portfolio weight, constructed by first assigning each underlying fund a representative benchmark portfolio based on its Morningstar category, computing the fund-level benchmark-adjusted weight in stock j, and then aggregating those fund-level benchmark-adjusted weights to the family-stock-quarter level using fund assets under management as weights. This variable is measured in percentage points of portfolio weight.
ActiveLong_{ijt}	Indicator equal to 1 if $BMAdjPortfolioWeight_{ijt} > 0.33$ (33 basis points) and 0 otherwise. This variable identifies benchmark-adjusted economically meaningful long positions.
Stock Characteristics	Measured at stock-quarter (jt) level.
ShortInt_{jt}	The number of shares held short divided by the number of shares outstanding. This ratio is winsorized at the 99th percentile and scaled by adding one and taking the natural log. This data is derived from the Compustat database.
IdioVol_{jt}	The standard deviation of daily abnormal returns for stock j during quarter t, winsorized at the 99th percentile. Abnormal returns are estimated using the five-factor model with momentum. This data is derived from the CRSP database as well as Kenneth French's website.
Disagreement_{jt}	The standard deviation of analyst expectations from the IBES database. The standard deviation is winsorized at the 99th percentile and scaled by adding one and taking the natural log.
News_{jt}	The total number of news articles and press releases for stock j during quarter t based on the Raven-Pack database. The sum of news articles is winsorized at the 99th percentile and scaled by adding one and taking the natural log.
MCAP_{jt}	The total market capitalization of firm j calculated as the share price multiplied by shares outstanding. The product is scaled by taking the natural log.
InstOwn_{jt}	The number of shares owned by institutions which file 13F reports, winsorized at the 99th percentile and scaled by adding one and taking the natural log. This data is derived from the Thomson Reuters 13F database.
Volume_{jt}	Share volume traded during quarter t, winsorized at the 99th percentile and scaled by taking the natural log. This data is derived from CRSP.
Mispricing_{jt}	Average mispricing score for quarter t. The score is from Stambaugh and Yuan (2017). For interpretation, we divide the score by 100 and subtract 0.50 to center on zero.

Underpriced_{jt}	Indicator equal to 1 if Mispricing_{jt} is in the bottom quintile of all stocks in quarter t.
Overpriced_{jt}	Indicator equal to 1 if Mispricing_{jt} is in the top quintile of all stocks in quarter t.
ExRet_{jt}	The excess return of stock j during quarter t. The excess return is calculated as the raw return less the risk-free rate. The excess return is multiplied by 100 for interpretability and is winsorized at the 1st and 99th percentiles.
DGTW_{jt}	The abnormal return of stock j during quarter t calculated as the difference between stock j's raw return less the appropriate equally-weighted DGTW benchmark return (Daniel et al. (1997)). The abnormal return is multiplied by 100 for interpretability and is winsorized at the 1st and 99th percentiles. This return is referred to as the "characteristic-adjusted return."
LowDisagree_{jt}	Indicator equal to 1 if Disagreement_{jt} is in the bottom quintile during quarter t.
$\text{LowDisagreeXUnderpriced}_{jt}$	Indicator equal to 1 if Underpriced_{jt} is 1 and stock j has low disagreement. Low disagreement is defined as Disagreement_{jt} in the bottom quintile during quarter t.
$\text{LowDisagreeXOverpriced}_{jt}$	Indicator equal to 1 if Overpriced_{jt} is 1 and stock j has low disagreement. Low disagreement is defined as Disagreement_{jt} in the bottom quintile during quarter t.

Family Characteristics	Measured at the family-quarter (it) level.
TNAM_{it}	Total net assets under management. Scaled by taking the natural log.
HHI_{it}	Herfindahl-Hirschman Index to measure portfolio concentration. Calculated as the sum of the squared portfolio share of each holding within a portfolio. Winsorized at the 99th percentile and divided by 100 for interpretability.
$\text{Expense Ratio}_{it}$	Total annual expenses and fees divided by year-end TNA. Winsorized at the 99th percentile.
Turnover_{it}	Minimum of aggregate purchases and sales of securities divided by average TNA over the calendar year. Winsorized at the 99th percentile and scaled using the natural logarithm.
Netflow12_{it}	The net growth in fund assets beyond reinvested dividends (Sirri and Tufano (1998)) over the past one year. Winsorized at the 1st and 99th percentiles.
Return_{it}	Measures the risk-adjusted returns for a given fund compounded over a six-month period. One risk-adjustment is to use "benchmark adjusted returns," where appropriate benchmark return is subtracted from the raw return of the fund. The second risk-adjustment is to account for the four-factor (FF4) model.

Macro Variables	Measured at the quarter (t) level.
MkReturn_{t}	The quarterly return on the market, from Kenneth French's database.
HML_{t}	The quarterly high-minus-low factor, from Kenneth French's database.
SMB_{t}	The quarterly small-minus-big factor, from Kenneth French's database.
UMD_{t}	The quarterly momentum factor, from Kenneth French's database.

A.3 Rich Fixed-Effects Research Tests

Table A1 reports a rich fixed-effects version of the research-intensity tests in Table 2. The baseline Table 2 specifications include family, stock, and quarter fixed effects. In this robustness test, we instead absorb stock-quarter and family-quarter fixed effects. Stock-quarter fixed effects compare different mutual fund families' requests for the same stock in the same quarter, absorbing stock-specific filing supply, news, and investor attention in that quarter. Family-quarter fixed effects compare the same family's requests across stocks in the same quarter, absorbing time-varying family-level research intensity.

The key short-side differences remain positive in this stricter specification. The short-minus-long difference remains about 6% for one-quarter $MFIA$ measures, and the short-minus-active-long difference remains about 9% to 10%. The two-quarter request measures are positive but somewhat less precise. For new positions, the lagged one-quarter $MFIA$ measures remain positive and statistically significant, while the two-quarter measures are weaker. Overall, the table supports the interpretation that the main Table 2 result is not driven solely by stock-quarter filing supply or by family-quarter research intensity.

[Table A1 about here.]

A.4 Strict Timing Robustness Tests

Our baseline empirical tests use quarterly holdings matched to quarterly EDGAR requests. This structure is useful for studying research around positions, but it cannot identify the exact within-quarter ordering of requests, trades, and returns. We therefore report two strict timing exercises. These tests are conservative because quarters are long and because lagging requests or moving returns to the next quarter removes same-quarter activity that may be part of position formation.

First, Table A2 repeats the positions-and-returns specifications from Table 3 using positions observed in quarter t and returns measured in quarter $t+1$. The next-quarter evidence is weaker than the baseline same-quarter evidence, as expected for a conservative timing exercise, but it aligns with the main short-side result because short positions continue to earn negative characteristic-adjusted returns and the pooled short coefficient remains negative in the DGTW specification. However, the new-position short differential is not statistically significant in the future-return specification. Thus, we interpret the strict timing version cautiously. It points to short-side return evidence in future DGTW returns, while the strongest new-position return spread remains concentrated around the quarter in which the position appears.

[Table A2 about here.]

Second, Table A3 provides the analogous timing robustness check for the new-position research-return specification in Table 5. It uses EDGAR requests measured before the position quarter and returns measured after the position quarter. In this strict design, lagged $MFIA$ is negatively associated with future returns for new longs, the $NewShortXMFIA$ interaction is positive in the DGTW specifications, and the total short-side slope (*NewShortXMFIA* + *MFIA*) is not statistically distinguishable from zero. We therefore view these results as a demanding stress test. Although the total short-side slope is imprecisely estimated, the positive DGTW interactions are directionally aligned with Table 5. The results do not reproduce the strong baseline estimates, but they are consistent with the idea that the baseline estimates should be interpreted as around-position evidence rather than clean trade-date identification.

[Table A3 about here.]

A.5 Model Details and Proofs

This section provides derivations underlying the theoretical framework developed in Section 4 of the main text. We proceed in three parts. Section A.5.1 presents the baseline one-shot information-acquisition model, proves equilibrium existence and uniqueness, and establishes Propositions 1–3. Section A.5.2 extends the environment to sequential-signal acquisition, derives the optimal stopping rule, and proves Proposition 4. Section A.5.3 provides details on the minimum tranche size ($q_{\min}$) assumption.

Throughout the appendix we retain the notation and parameterization introduced in the main text unless explicitly noted otherwise.

A.5.1 Baseline Model

Equilibrium Existence and Uniqueness

The risky asset has a binary fundamental value ($V \in \{\nu_H, \nu_L\}$) and $\Delta \equiv \nu_H - \nu_L > 0$ is the value gap. A risk-neutral investor may observe a signal of value by choosing precision $\pi \in [0, 1]$ and paying cost $c(\pi) = \kappa\pi^2$, $\kappa > 0$. Conditional on V, the signal $x \in \{\nu_H, \nu_L\}$ satisfies $P(x = V) = \frac{1}{2} + \frac{\pi}{2}$. The posterior mean therefore lies $\frac{\pi\Delta}{2}$ above or below the pre-trade price ($P_0 = \frac{\nu_H+\nu_L}{2}$), which is pinned down by competitive noise traders at the unconditional mean. After observing the signal, the investor selects position side $s \in \{L \text{ (long)}, S \text{ (short)}, 0\}$ and position size q.

Opening any position incurs fixed monitoring costs $F > 0$, and investors abstain from trading or trade a minimal block $q_{\min} > 0$. Trading costs for long positions total $C_L = F$ while for short positions $C_S = F + \phi + \eta q_{\min}$ where both $\phi > 0$ and $\eta q_{\min} > 0$.

Lemma A1. *For any chosen precision π, optimal trade decisions are:*

$$\textit{Go Long iff } x = \nu_H \textit{ and } \frac{\pi\Delta}{2} q_{\min} \geq C_L, \qquad \textit{Go Short iff } x = \nu_L \textit{ and } \frac{\pi\Delta}{2} q_{\min} \geq C_S.$$

Proof. Conditional expected profit from a long after a high signal is $\frac{\pi\Delta}{2}q_{\min} - C_L$. The investor trades only when this is non-negative; the symmetric argument gives the short cut-off. ■

Define the precision hurdles for long and short trades:

$$\pi_L^{thr} \equiv \frac{2C_L}{\Delta q_{\min}}, \qquad \pi_S^{thr} \equiv \frac{2C_S}{\Delta q_{\min}} > \pi_L^{thr}, \tag{A1}$$

and let the investor's *ex-ante* expected profit from optimally trading after observing x be:

$$\Gamma(\pi) \equiv \frac{1}{2}\Big[q_{\min}\frac{\pi\Delta}{2} - C_L\Big]_+ + \frac{1}{2}\Big[q_{\min}\frac{\pi\Delta}{2} - C_S\Big]_+ - c(\pi),$$

where $[z]_+ \equiv \max\{z, 0\}$. There are three relevant ranges for π:

- For $\pi \leq \pi_L^{thr} : \Gamma(\pi) = -c(\pi) < 0$.
- For $\pi \in (\pi_L^{thr}, \pi_S^{thr}]$: only long trades can be taken, so $\Gamma(\pi) = \frac{1}{2}(q_{\min}\frac{\pi\Delta}{2} - C_L) - c(\pi)$.
- For $\pi > \pi_S^{thr}$: both long and short trades can be taken, so $\Gamma(\pi) \equiv \frac{1}{2}(q_{\min}\frac{\pi\Delta}{2} - C_L) + \frac{1}{2}(q_{\min}\frac{\pi\Delta}{2} - C_S) - c(\pi)$.

Because $c(\pi)$ is strictly convex and the linear benefit pieces are continuous with kinks at both π_L^{thr} and π_S^{thr}, $\Gamma(\pi)$ is strictly concave on each region and continuous everywhere, so it attains a unique global maximizer $\pi^* \equiv \arg\max_{\pi\geq 0}\Gamma(\pi)$. If the maximal expected profit is non-positive, $\Gamma(\pi^*) \leq 0$, the investor acquires no information and never trades; otherwise she chooses the interior optimum $\pi^* > 0$ and follows the cut-off rule in Lemma A1. This proves **existence** and **uniqueness** of the equilibrium. ■

Proof of Proposition 1

Proposition 1 states that the precision threshold is higher for shorts than for longs: $\pi_S^{thr} > \pi_L^{thr}$.

Proof. Precision thresholds for longs and shorts are defined in Equation A1. By construction, $C_S = C_L + \phi + \eta q_{\min} > C_L$ where the locate fee and borrow surcharge are both greater than zero ($\phi, \eta q_{\min} > 0$). Therefore, $\pi_S^{thr} > \pi_L^{thr}$. ■

Proof of Proposition 2

Proposition 2 states that the average precision (or research intensity) for observed shorts is higher than for observed longs.

Proof. Recall the two precision thresholds ($\pi_L^{thr} = \frac{2C_L}{\Delta q_{\min}}$ and $\pi_S^{thr} = \frac{2C_S}{\Delta q_{\min}} > \pi_L^{thr}$), let the gross trading benefit be defined as $B(\pi) = \frac{1}{2}\Big[q_{\min}\frac{\pi\Delta}{2} - C_L\Big]_+ + \frac{1}{2}\Big[q_{\min}\frac{\pi\Delta}{2} - C_S\Big]_+$, and note that the slope of this benefit (the marginal benefit curve), with respect to chosen precision π, is a step-function with three flat regions:

- $B'(\pi) = 0$ for $\pi \leq \pi_L^{thr}$;
- $B'(\pi) = b_L = \frac{q_{\min}\Delta}{4}$ on $(\pi_L^{thr},\ \pi_S^{thr}]$ (only longs can cover fixed costs);
- $B'(\pi) = b_S = 2b_L = \frac{q_{\min}\Delta}{2}$ when $\pi > \pi_S^{thr}$ (both longs and shorts can cover fixed costs).

Also note that the investor's net objective, the ex-ante expected profit, is $\Gamma(\pi) = B(\pi) - \kappa\pi^2$ and the marginal cost is $2\kappa\pi$.

Matching the marginal-benefit step-function to the upward sloping marginal cost curve gives two candidate interior optima:

$$\pi^L(\kappa) = \frac{q_{\min}\Delta}{8\kappa}, \quad \pi^S(\kappa) = \frac{q_{\min}\Delta}{4\kappa};$$

which are feasible only if they lie on their intended segments. Define

$$\kappa_{high} = \frac{(q_{\min}\Delta)^2}{16C_L}, \quad \kappa_{low} = \frac{(q_{\min}\Delta)^2}{8C_S}$$

and note that $\kappa_{low} < \kappa_{high}$ whenever $C_S > 2C_L$. Thus, whenever short-selling costs are significantly higher than the costs of long positions, the investor's optimal precision falls into exactly one of the following cost bands.

- **High-cost band**. When $\kappa > \kappa_{high}$ the marginal cost curve is always higher than the marginal-benefit step function and, thus, the optimal precision is $\pi^* = 0$ and neither long or short trades are executed.[4]

- **Moderate-cost band**. When $\kappa_{low} < \kappa < \kappa_{high}$ the marginal cost curve intersects the marginal-benefits step-function only in the long-only region, thus, the optimal precision is $\pi^* = \pi^L(\kappa) \in (\pi_L^{thr}, \pi_S^{thr})$ and only long trades are executed.

- **Low-cost band**. When $\kappa < \kappa_{low}$ the marginal cost curve intersects the marginal-benefits step-function in the short region (or is always below b_S), thus, the optimal precision is $\pi^* = \pi^S(\kappa) \geq \pi_S^{thr}$ and both longs and shorts are executed.

Shorts are only observed in the low-cost band and are always backed by high precision ($\pi^* \geq \pi_S^{thr}$) while longs are observed in both the low-cost band and the moderate-cost band, thus some longs are backed by lower precision ($\pi_L^{thr} < \pi^* < \pi_S^{thr}$). Hence

$$E[\pi^* \mid s = S] \geq E[\pi^* \mid s = L],$$

with strict inequality when $C_S > 2C_L$ (so the moderate-cost band has positive width). ■

Proof of Proposition 3

Proposition 3 states that the average absolute abnormal return for observed shorts is higher than for observed longs.

Proof. After observing the binary signal $x \in \{\nu_H, \nu_L\}$ the investor's posterior mean is $E[V|x] = P(V = \nu_H|x)\nu_H + P(V = \nu_L|x)\nu_L$, so the perceived deviation from price is always $|E[V|x] - P_0| = \frac{\pi\Delta}{2}$. This is the size of the "gap" the investor believes exists between the fundamental value and the market price.

A trade of side $s \in \{L, S\}$ is executed only if the expected gross gain on the minimum block covers the fixed cost C_s; Lemma 1 therefore implies the return hurdle $|E[V|x] - P_0| \geq$

[4] At $\kappa = \kappa_{high}$ the investor is indifferent between $\pi = 0$ and π_L^{thr}; we take $\pi^* = 0$ without loss of generality.

$\frac{C_s}{q_{\min}} = \tau_s$. Because $\tau_S > \tau_L$, a short is opened only when the perceived mispricing is strictly larger than the minimum mispricing that triggers a long.

Shorts appear only in the low-cost band identified in Proposition 2; there the chosen precision satisfies $\pi^* \geq \pi_S^{thr}$, hence the mispricing at entry obeys $|E[V|x] - P_0|_{s=S} \geq \frac{\pi_S^{thr}\Delta}{2}$. Longs are observed (i) in the same low-cost band when $\pi^* \geq \pi_S^{thr}$, and (ii) in the moderate-cost band when $\pi_L^{thr} < \pi^* < \pi_S^{thr}$. In the second scenario, the mispricing lies strictly below $\frac{\pi_S^{thr}\Delta}{2}$. Hence the distribution of mispricings for longs is a mixture of large mispricings (from the low-cost band) and strictly smaller ones (from the moderate-cost band), whereas the short-side distribution contains only the large mispricings.

The realized abnormal return on a unit position is $\alpha = V - P_0$. Conditional on the investor's entry rule, its *absolute* expectation is monotone in the perceived mispricing. Because a correct-side payoff of $+\frac{\Delta}{2}$ occurs with probability $\frac{1+\pi}{2}$ while a wrong-side payoff $-\frac{\Delta}{2}$ occurs with probability $\frac{1-\pi}{2}$, larger perceived mispricing at entry translates into larger average $|\alpha|$. Therefore, $E[|\alpha||s = S] \geq E[|\alpha||s = L]$, with strict inequality when $C_S > 2C_L$ (so the moderate-cost band has positive width). ■

A.5.2 Extended Model: Sequential Signals

Equilibrium Existence and Uniqueness

Let the investor observe sequential Gaussian signals $x_j = V + \varepsilon_j,\ \varepsilon_j \sim \mathcal{N}(0, \sigma^2)$, each costing $\kappa\pi_0^2$ where $\pi_0 = \frac{1}{\sigma^2}$. After N draws the posterior mean and precision are $m_N = \frac{1}{N}\sum_{j=1}^{N} x_j$, and $\pi_N = N\pi_0$. Let $V_N(m)$ be the investor's *continuation value* after N draws and posterior mean m. At that node the investor chooses to either (i) **trade:** take side s on the minimum block $q_{\min}$ if $|m| \geq \tau_s$, earning $q_{\min}(|m| - \tau_s)$; (ii) **draw:** pay the cost $c = \kappa\pi_0^2$ and receive $E[V_{N+1}(m')]$; or (iii) **stop:** do nothing and receive value 0. Hence,

$$V_N(m) = \max\Big\{0,\ q_{\min}(|m| - \tau_L),\ q_{\min}(|m| - \tau_S),\ E[V_{N+1}(m')] - c\Big\}. \tag{A2}$$

Because the Gaussian update is linear and costs are constant, $V_N(\cdot)$ is even and strictly increasing in $|m|$ once $|m| \geq \tau_L$. This implies thresholds $0 < \hat{m}_N^L \leq \hat{m}_N^S$ such that the unique optimal action is **draw** if $|m| < \hat{m}_N^L$ and $E[V_{N+1}(m')] - c > 0$; **quit** if $|m| < \hat{m}_N^L$ and $E[V_{N+1}(m')] - c \leq 0$; **long** if $\hat{m}_N^L \leq m < \hat{m}_N^S$; and **short** if $|m| \geq \hat{m}_N^S$. Posterior variance is monotone decreasing in N, so the short boundary $h_N := \hat{m}_N^S$ is weakly increasing. Suppose $h_1 > \tau_S$. When $\tau_S = |m_1| < h_1$ the investor would strictly improve by shorting now rather than paying c to continue, contradicting optimality. Hence, $h_1 = \tau_S$ and, by monotonicity, $h_N = \tau_S$ for every N; the long boundary must equal τ_L by the same argument. Therefore, the optimal stopping rule is: (i) continue sampling while $|m_N| < \tau_L$ and the option value of another draw exceeds its cost; (ii) quit with no position if $|m_N| < \tau_L$ but the option value of another draw has fallen to zero; (iii) trade a long block as soon as $\tau_L \leq |m_N| < \tau_S$; and (iv) trade a short block as soon as $|m_N| \geq \tau_S$.

For this rule the investor is indifferent at the thresholds by definition of $\tau_s = \frac{C_s}{q_{\min}}$, and market makers break even at $P_0 = \frac{\nu_H + \nu_L}{2}$. Thus (τ_L, τ_S, P_0) constitutes an equilibrium. Now, assume a distinct pair $(\tilde{\tau}_L, \tilde{\tau}_S) \neq (\tau_L, \tau_S)$ also supports an equilibrium. If $\tilde{\tau}_S < \tau_S$, lowering the short boundary lets the investor save at least one signal cost while never reducing trade pay-offs, strictly increasing expected profit—a contradiction. If $\tilde{\tau}_S > \tau_S$ or the long boundary moves, a symmetric contradiction arises. Hence, the boundary pair (τ_L, τ_S) is unique. ■

Proof of Proposition 4

Proposition 4 states that the average absolute abnormal return for observed shorts decreases in the number of signals acquired.

Proof. Let the posterior dollar gap be $g_N \equiv |m_N - P_0|$ and its conditional standard deviation be $\sigma/\sqrt{N}$. A short block is traded the first time the gap reaches the fixed hurdle $g_N \geq \tau_S$ where $\tau_s = \frac{C_s}{q_{\min}}$. Denote this draw count by T_S.

Conditional on trading after exactly N signals ($T_S = N$), the gap must lie in $[\tau_S, \infty)$, with a truncated-normal tail. A standard formula gives the conditional mean of the overshoot:

$$E[g_N - \tau_S \mid T_S = N] = \left(\frac{\sigma}{\sqrt{N}}\right)\frac{\varphi(a_N)}{\Phi(-a_N)}, \quad a_N \equiv \frac{\tau_S\sqrt{N}}{\sigma}, \tag{A3}$$

where φ and Φ are the standard-normal pdf and cdf.

Because $\Phi(-a_N)$ grows and $\varphi(a_N)$ shrinks with a_n, the fraction is bounded above by $\frac{1}{\sqrt{2\pi}}$. Thus, $E[g_N \mid T_S = N] \leq \tau_S + \frac{\sigma}{\sqrt{2\pi N}}$ and the right side is strictly decreasing in N.

At settlement the one-block abnormal return is $\alpha = V - P_0$. Conditional on executing the short after precisely N filings, the posterior mean is m_N and the residual uncertainty in V is $\mathcal{N}(m_N, \frac{\sigma^2}{N})$. Therefore,[5]

$$E[|\alpha| \mid T_S = N] = \tau_S + \frac{\sigma}{\sqrt{2\pi N}} + \frac{\sigma\sqrt{2}}{\sqrt{\pi N}}, \tag{A4}$$

is the sum of three positive terms whose second and third pieces shrink at the rate $\frac{1}{\sqrt{N}}$. Hence, the whole expectation falls as N rises.

The details above focused on shorts. Long positions satisfy the same analytical bound as shorts (Equation A4) but with the lower hurdle $\tau_L < \tau_S$. The second and third terms of Equation A4—the overshoot of the posterior mean and the residual valuation noise—are identical on both sides and shrink like $\frac{1}{\sqrt{N}}$. What is different between longs and shorts is the baseline:

- **Shorts** must first clear the high dollar hurdle τ_S that reflects extra locate and borrow fees. Because τ_S materially exceeds τ_L, each additional signal cuts the *total* expected $|\alpha|$ by a perceptible amount, producing the downward slope highlighted in Proposition 4.

- **Longs** face the lower hurdle τ_L. Relative to the smaller baseline, the same $\frac{1}{\sqrt{N}}$ decline in the overshoot-and-noise terms is proportionally tiny, so the model predicts an almost

[5] See Lorden (1970).

flat relation between information acquired and subsequent absolute returns on the long side. ■

A.5.3 Minimum Tranche Size

Operational Practice. Large U.S. mutual-fund families rarely initiate a new position with one share. Opening a position—long or short—triggers portfolio-compliance checks, internal risk sign-off, back-office trade-ticket generation, and downstream monitoring of corporate events. Therefore, desks often impose an internal ticket floor: long orders are usually entered in round lots whose dollar value exceeds a preset minimum (often \$25–\$50 thousand for mid-cap or large-cap stocks), and prime brokers supply stock-loan locates in the same round-lot multiples on the short side. The model captures these institutional frictions with a fixed monitoring cost F and a non-divisible block $q_{\min}$. The surcharge $\eta q_{\min}$ for shorts mirrors the per-share borrow fee applied to that same lot.

Modeling Choice. Treating $q_{\min}$ as given is a simplification rather than a claim that managers could never trade different block sizes. In richer microstructure models investors do optimize quantity jointly with information (e.g., Kyle (1985)-type settings). Here, letting the analyst pick an arbitrarily small q would merely convert the fixed desk cost F into a per-share cost F/q, forcing us to append an additional first-order condition without altering the economics of information choice. Keeping $q_{\min}$ fixed therefore improves tractability while staying close to common desk practice.

Robustness. If one set $q_{\min} \to 0$ while keeping a strictly positive fixed cost F, the investor could open an infinitesimal stake but still incur the full monitoring expense—a degenerate corner that would require adding risk aversion or per-share execution costs to restore an interior optimum. Such extensions complicate notation yet leave the paper's qualitative insights intact: (i) higher short-side costs still raise the break-even precision hurdle; (ii) the κ-band logic that delivers Propositions 1–3 is unchanged; and (iii) in the sequential model, each additional draw would still shrink the posterior variance by $1/\sqrt{N}$ and hence reduce the expected overshoot term, preserving Proposition 4. We therefore retain the stylized—but industry-consistent—assumption of a fixed, non-optimized tranche size for analytical clarity.

Table A1: Research Intensity and Positions with Rich Fixed Effects

This table compares the key difference estimates from Table 2 with estimates from a richer fixed-effects specification. Baseline FE rows reproduce the differences in Table 2 using family, stock, and quarter fixed effects. Rich FE rows absorb stock-quarter and family-quarter fixed effects, comparing different families' requests for the same stock in the same quarter and the same family's requests across stocks in the same quarter. Columns 2–4 use one quarter of EDGAR requests and its one- and two-quarter lags. Columns 5–7 use the two-quarter request measure and its one- and two-quarter lags. Panel A compares shorts with longs, Panel B compares shorts with benchmark-adjusted active longs, and Panel C compares new shorts with new longs. Standard errors are clustered by family-stock and quarter and are shown in parentheses. Indicators ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
	MFIA = Requests_{ijt}			MFIA = $\text{Requests}_{ijt,t-1}$		
	MFIA	L1.MFIA	L2.MFIA	MFIA	L1.MFIA	L2.MFIA
Panel A. Short minus Long.						
Baseline FE	0.06**	0.07**	0.06**	0.05*	0.06*	0.06*
	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
Rich FE	0.06**	0.06**	0.06**	0.05	0.05	0.05
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Obs. (Baseline FE)	2,528,011	2,528,011	2,528,011	2,528,011	2,528,011	2,528,011
Obs. (Rich FE)	2,512,253	2,512,253	2,512,253	2,512,253	2,512,253	2,512,253
Panel B. Short minus ActiveLong.						
Baseline FE	0.10***	0.11***	0.11***	0.09**	0.09**	0.10***
	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)
Rich FE	0.09***	0.09***	0.10***	0.07*	0.07*	0.08*
	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)
Obs. (Baseline FE)	2,573,701	2,573,701	2,573,701	2,573,701	2,573,701	2,573,701
Obs. (Rich FE)	2,557,948	2,557,948	2,557,948	2,557,948	2,557,948	2,557,948
Panel C. NewShort minus NewLong.						
Baseline FE	0.01	0.07**	0.07**	0.02	0.08**	0.08**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Rich FE	0.03	0.06**	0.05**	0.04	0.06*	0.05
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Obs. (Baseline FE)	2,575,774	2,575,774	2,575,774	2,575,774	2,575,774	2,575,774
Obs. (Rich FE)	2,560,021	2,560,021	2,560,021	2,560,021	2,560,021	2,560,021

Table A2: Future Returns After Long and Short Positions

This table provides a timing robustness check for Table 3. It shows the results from estimating Equation 2 using positions measured in quarter t and returns measured in quarter $t+1$. The column structure follows Table 3: Columns 2 and 3 report long-only and short-only samples, Columns 4 and 5 use all positions, and Column 6 uses new positions. In Panel A, the dependent variable is next-quarter excess return ($F1.ExRet$). In Panel B, the dependent variable is next-quarter characteristic-adjusted return ($F1.DGTW$). The specifications include the quarterly risk factors; $MktReturn$ is reported, while SMB, HML, and UMD are omitted from the table. The model includes a mutual fund family fixed effect. Standard errors are clustered by stock and quarter and are shown in parentheses. Indicators ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively.

(1)	(2)	(3)	(4)	(5)	(6)
	Long Only	Short Only	All Positions		
			Long/Short	ActiveLong/Short	New Pos.
Panel A. Dependent Variable: Future Excess Return (F1.ExRet)					
MktReturn	1.02***	0.95***	1.02***	1.02***	1.07***
	(0.03)	(0.06)	(0.03)	(0.03)	(0.05)
α	-0.08	-0.20	-0.08	-0.08	-0.39*
	(0.15)	(0.26)	(0.15)	(0.15)	(0.21)
Short			-0.38	-0.38	
			(0.24)	(0.24)	
ActiveLong				0.09	
				(0.30)	
NewShort					-0.23
					(0.28)
Obs.	1,155,759	22,110	1,177,869	1,177,869	114,222
R^2	0.17	0.15	0.17	0.17	0.15
Panel B. Dependent Variable: Future Characteristic-Adjusted Return (F1.DGTW)					
MktReturn	0.03*	0.04	0.03*	0.03*	0.06**
	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)
α	-0.19	-0.50***	-0.19	-0.19	-0.44***
	(0.12)	(0.14)	(0.12)	(0.12)	(0.14)
Short			-0.32*	-0.32*	
			(0.18)	(0.18)	
ActiveLong				0.10	
				(0.13)	
NewShort					-0.04
					(0.18)
Obs.	932,613	18,590	951,203	951,203	82,953
R^2	0.00	0.00	0.00	0.00	0.00

Table A3: Future Returns and Lagged Research for New Positions

This table provides a timing robustness check for Table 5. It shows the results from estimating Equation 3 on new positions using a stricter timing design. Returns are measured in the quarter after the new position appears. Columns 2 and 3 use lagged one-quarter research intensity, $MFIA = Requests_{ijt-1}$, and Columns 4 and 5 use lagged two-quarter research intensity, $MFIA = Requests_{ijt-1,t-2}$. As in Table 5, $NewLong$ measures the future abnormal return to stocks in new long positions, $NewShort$ measures the additional future return to stocks in new short positions, $MFIA$ measures the relation between research intensity and future returns for new long positions, and $NewShortXMFIA$ measures the additional relation for new short positions. Thus, $NewShortXMFIA + MFIA$ measures the total short-side relation between lagged research intensity and future returns. The ExRet specifications include the quarterly risk factors; $MktReturn$ is reported, while SMB, HML, and UMD are omitted from the table. The model includes mutual fund family and stock fixed effects. Standard errors are clustered by stock and quarter and are shown in parentheses. Indicators ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively.

(1)	(2)	(3)	(4)	(5)
	Future Returns for New Positions			
	MFIA = Requests$_{ijt-1}$		MFIA = Requests$_{ijt-1,t-2}$	
	F1.ExRet	F1.DGTW	F1.ExRet	F1.DGTW
MktReturn	1.07***	0.05**	1.07***	0.05**
	(0.04)	(0.02)	(0.04)	(0.02)
NewLong	-0.25	-0.33***	-0.22	-0.31**
	(0.18)	(0.12)	(0.19)	(0.13)
NewShort	-0.45	-0.31	-0.49*	-0.34
	(0.27)	(0.23)	(0.28)	(0.25)
MFIA	-0.37***	-0.20**	-0.31***	-0.16*
	(0.10)	(0.09)	(0.10)	(0.08)
NewShortXMFIA	0.22	0.33*	0.21	0.28*
	(0.18)	(0.17)	(0.14)	(0.14)
H_0 :NewShortXMFIA + MFIA = 0	-0.15	0.13	-0.10	0.11
	(0.19)	(0.15)	(0.16)	(0.14)
Obs.	113,745	82,669	113,745	82,669
R^2	0.25	0.09	0.25	0.09
Family FE	Yes	Yes	Yes	Yes
Stock FE	Yes	Yes	Yes	Yes