Rationally Neglected Stocks

Oleg Chuprinin* (University of New South Wales)

Arseny Gorbenko (Monash University)

Chang-Mo Kang (Hanyang University)

Abstract

There are large cross-sectional differences in the probability and magnitude of mispricing among stocks. Mispricing is traditionally attributed to stock-specific frictions. We show that mispricing can be explained in a rational equilibrium where investors allocate investigative resources to stocks to maximize their expected profits from arbitrage. Stocks with smaller dollar profit potential are allocated less attention and their percentage mispricing is higher on average. For such stocks, information discovery by investors is slow and the mispricing is corrected mostly through mandatory disclosures by firms. Using measures of institutional attention and trading discreteness, we confirm this mechanism empirically. The attention allocation channel explains cross-sectional pattern of mispricing better than any classic arbitrage frictions.

Key words: Mispricing, Information discovery, Attention, Institutional trading

JEL codes: G11, G14

* Please address your correspondence to Oleg Chuprinin: Room 349, School of Banking and Finance, UNSW Business School, 2052 Sydney Australia; tel.: (+61) 2 9385 5856; fax: (+61) 2 9385 5856; email: o.chuprinin@unsw.edu.au.

Introduction

Investors' attention is a limited resource, but one that is essential to ensure speedy information discovery and fair pricing of financial assets. Recent empirical studies have demonstrated a direct effect of investor attention on asset prices (e.g., Hirshleifer, Lim, and Teoh (2009); Ben-Rephael, Da, and Israelsen (2017); Peress and Schmidt (2018)). These studies focus on exogenous variation in attention and investigate price distortions resulting from a lack of attention to firms' fundamentals. However, since attention is a resource and investors are value-maximizing agents, it is natural to ask how investors choose to allocate their attention to different investment opportunities and what are the implications of this choice for the cross-section of stock prices.

Mispricing of financial assets is traditionally attributed to market frictions, such as transaction costs, illiquidity, short sale constraints, and investor irrationality. Under this view, cross-sectional differences in the magnitude of mispricing follow directly from the corresponding differences in the severity of frictions (e.g., stocks that are costlier to trade are more likely to be mispriced). In this paper, we propose and test an alternative mechanism, in which variation in mispricing results not from differences in frictions but from investors' incentives to expend investigative resources to discover information. We pursue two goals. First, we present a simple model that provides a rational explanation for the cross-sectional variation in attention allocated to stocks.[1] This explanation does not rely on any external stylistic preferences for particular stocks. Instead, these preferences arise endogenously as investors maximize their expected gains. Second, we show how this variation affects mispricing magnitudes in the cross-section of stocks. We confirm our theoretical predictions in a series of tests.

[1] We use the term "attention" as an informal moniker for all investigative resources available to an investor. For example, consider a portfolio manager who needs to decide how to allocate his analysts' workload to research different companies. For simplicity, we refer to this situation as the "attention allocation problem", and it applies mostly to institutional money managers—investors who engage in active information discovery.

Our agenda can be summarized as follows: (i) explaining why prices of some stocks are slow to reflect changes in fundamentals, (ii) understanding why investors update some of their positions regularly while leaving others stale,[2] and (iii) relating these phenomena in a rational resource-allocation framework.

The mechanism we study can be illustrated with the following idealized example. By buying shares in a firm, the investor acquires a claim to the firm's cash flows. The difference between the fair value of these cash flows and their market value is the maximum arbitrage profit available in that stock. For example, if a hedge fund acquires the whole supply of shares of a $100-million dollar firm at a 5% discount, it will realize an arbitrage profit of $5 million (ignoring price impact). By allocating more investigative resources to a stock, the investor increases the probability of detecting mispricing in that stock and therefore increases his expected arbitrage profit in that stock. In equilibrium, the investor would allocate these resources in such a way that the marginal expected profit is the same for all stocks.

The main implication of this mechanism is that we should not expect stocks to be equally mispriced on average. The same marginal expected profit is not synonymous with the same level of mispricing. Rather, mispricing—i.e. the absolute gap between the stock price and the fundamental value—should be higher for smaller stocks in equilibrium. This effect has nothing to do with liquidity or costs of investigation, but is a direct outcome of the rational dollar-profit maximization problem. This effect is not unique to micro stocks but operates continuously over the whole spectrum of sizes.[3]

We develop a simple model with two stocks and one investor. At time 0, the investor decides now much attention to allocate to each stock. His total level of attention is constrained. The investor is aware

[2] Consider the following statistics (see Section 4 for details). In a sample of institutional trades from ANcerno, trading frequency varies substantially across stocks for the same account. The average time between trades of the same stock by the same account in a year is 24 days, but the variation is large: from P10 of 1 day to P90 of 60 days. We observe a similarly large variation in the turnover of mutual fund positions. Some portfolio positions are updated regularly while others are permanently stale. For example, for the quartile of stocks with the lowest (highest) position turnover, in 59.2% (3.6%) of cases the position is unchanged from the previous quarter.

[3] This line of reasoning echoes the argument of Berk and van Binsbergen (2015), who stress the importance of equilibrium effects in asset management. Among other things, they show that an investor's skill is best measured by the dollar value extracted from capital markets rather than a percentage measure, such as alpha. Similarly, in our setting, measuring mispricing in a traditional way, i.e. as a percentage deviation of the price from fundamentals or a percentage return to a trading strategy, produces a strong size bias: smaller investment opportunities are more mispriced on average.

that greater attention to a stock improves the signal precision and allows him to measure fundamentals more accurately and increase his expected arbitrage profit in the stock.[4] At time 1, the investor receives the signals for both stocks, takes positions, and the stock prices adjust according to demand. The more precise the signal in a stock, the greater the informed demand, and the closer the price moves towards the fundamental value. The gaps that remain after the investor trades the stocks are the "equilibrium mispricing". These levels of mispricing are equilibrium in the sense that the investor takes them into account when maximizing his expected gains by allocating attention at time 0. The last event in our model occurs at time 2, when both companies make public disclosures and announce their fair values. At this time, both stocks adjust all the way to the corresponding fundamental values and the mispricing gaps close completely. This step is not essential to the theory, but it justifies a convenient empirical measurement of the equilibrium percentage mispricing. Specifically, the mispricing before the disclosure can be measured as an absolute stock return observed at the disclosure.

The model delivers two major predictions: one describing stock mispricing and the other highlighting the role of capital constraints. The first prediction—that prices of smaller stocks deviate more from fundamental value—might appear trivial. However, the traditional explanations of this phenomenon are based on arbitrage frictions (e.g., Hong, Lim, and Stein (2000); Amihud (2002); Fama and French (2008), among others). In contrast, we rationalize this effect through optimal resource allocation by professional investors. In this channel, size is the *first-order* driver of mispricing and is not a mere proxy for frictions that correlate with size. In other words, even if we allow arbitrage frictions to be the same for all stocks, we should still observe that percentage mispricing is a monotonically decreasing function of size.

We confirm the first-order effect of size empirically by comparing how mispricing varies by size vs. other stock characteristics commonly associated with slow price discovery. We consider the following classes of measures: illiquidity (Amihud illiquidity and bid-ask spread), lending market frictions (stock

[4] "Mispricing" here refers to the gap between the price and the fundamental value of the stock. This is the standard theoretical definition of mispricing, shared by many classical papers in information economics (e.g., Grossman and Stiglitz (1980); De Long et al. (1990); Froot, Scharfstein, and Stein (1992)).

borrowing fee and lending supply), retail investor participation (retail ownership and trading activity), and information environment (media coverage, tangibility of information, and financial report complexity).

In the first series of tests, we focus on the stock price reactions to earnings announcements. For each such event, we take the absolute abnormal stock return at the announcement to proxy for the gap between the price and the fundamentals *before* the earnings news was released by the firm.[5] As expected, each of the following is associated with a bigger price movement at the announcement: smaller size, high illiquidity and spread, high borrowing fee and low lending supply, high retail investor participation, low media coverage and information tangibility, and high information complexity.

Yet, size dominates the other characteristics in both statistical and economic significance. The *t*-statistics for size are twice as high as those for spread and retail trading—the other strong predictors. If size is included jointly with other variables, it dominates to the point where some of these variables lose significance entirely. Using double-sorting analysis, we find that the effect of size remains strong after pre-sorting by other characteristics, but the opposite is generally not true. In sum, *no explicitly measured friction* explains mispricing as well as size does. Importantly, the effect of size is monotonic: it operates over the entire spectrum of stocks and is not specific to the smallest quintile. The attention-allocation model predicts this smooth pattern exactly, whereas friction-based explanations suggest that mispricing should be concentrated in special stocks.

Next, we examine proxies for institutional attention: the news search frequency in the Bloomberg Terminal (see Ben-Rephael, Da, and Israelsen (2017)) and the downloads of EDGAR filings. Stock size is the strongest predictor of institutional attention, more so than media coverage or even institutional ownership. Then, we analyze the discreteness and continuity of institutional trading in the ANcerno

[5] This approach aligns with the logic of the model and can also be justified as follows. We want to capture to what extent investors engage in information discovery on their own, i.e. before the firm releases information itself. If investors continuously seek information and trade on it, the stock price would adjust little during the announcement, since a significant portion of the information that could have been discovered had already been priced. In contrast, if investors lack incentives to discover information, their trading in the stock is sparse and largely uninformed, and the price corrects only when the firm itself releases material information to the public. Thus, a big price jump or drop at the announcement indicates that little information had been discovered by investors before the disclosure event.

database. We find that the sparseness of trading attributable to small size or lack of attention is associated with greater mispricing and stronger subsequent price corrections. Interestingly, trading inactivity that cannot be explained by size or inattention shows no such relationship.

The second prediction describes the role of capital constraints. We consider investors with different levels of capital availability (and call investors with less capital “small” for simplicity). Larger investors care more about the size of a potential investment opportunity, since this size represents a binding constraint on their profits. In contrast, small investors are bound by their own capital limits rather than by arbitrage capacities of investment opportunities. Accordingly, small funds are less discriminatory in their attention allocation and are relatively more content to follow and trade small stocks.

The main insight of this mechanism is that it endogenizes preference for small vs. large stocks as a function of capital availability. In particular, it can explain why specialization into small- vs. large-cap investment classes is so prevalent among asset managers. This separation arises endogenously in the investor’s maximization problem. This argument challenges, or at least augments, the view that “style” is just a reflection of the clientele-imposed mandate or the manager’s expertise in a given asset class.

We obtain evidence for this prediction by analyzing the Ancerno trading data and mutual fund holdings from Thomson Reuters. We find that the relation between trading frequency and stock size is significant only for accounts with a sufficiently high turnover in ANcerno. We then examine position stickiness in mutual fund portfolios, controlling for a variety of fund characteristics. Smaller stocks are more sticky on average: 54% of positions (the number of shares held) in the bottom-size quintile are not updated during a quarter, compared to 23% in the top-size quintile. However, this disparity decreases dramatically with fund TNA, and this effect is not explained by external style mandates or other controls. Overall, as predicted, smaller funds attend to different stocks in their portfolios more uniformly. Notably, alternative explanations for the size preference (e.g., illiquidity or execution costs) predict the opposite: funds with less capital should be relatively more discriminatory and less willing to update positions in smaller stocks due to higher proportional trading costs in these stocks.

This paper makes contributions to different strands of the finance literature. First, we add to the literature that provides rational explanations for apparently anomalous phenomena (Berk and Green (2004), Berk and van Binsbergen (2015)). We show that it is natural for stocks to have different levels of mispricing if the latter is measured in percentage terms. In other words, the equilibrium is not a situation where all stocks are on average equally mispriced.

Second, we highlight the first-order effect of size in the cross-section of stock prices. Theoretically, we show that there is a distinct channel linking company size to stock mispricing that is independent of frictions traditionally associated with slow price discovery. Empirically, we perform the first direct and comprehensive comparison between size and other stock characteristics (which capture different dimensions of limits to arbitrage) and show that size is the dominant driver of mispricing, whose effect is not subsumed by any explicitly measured friction. Size is perhaps the most ubiquitous variable in financial economics, included as a control in multiple studies, both corporate and asset pricing. Authors rarely specify what exactly size controls for in any given setting. We show, both theoretically and empirically, that size is not a mere proxy for other effects but plays its own unique role.

Third, we provide a rational explanation for style preferences in asset management. "Style" is usually interpreted as a fund mandate imposed exogenously and attributable to investors' tastes ("clientele effects"). Cronqvist, Siegel, and Yu (2015) shed light on the origins of one dimension of style—value vs. growth investing. We endogenize the other dimension—preference for small vs. large cap stocks—as a function of investment capital availability. We also provide empirical evidence in support of this channel.

This paper adds to the literature on investors' attention. Empirical studies show that investors' attention does influence stock prices (Hirshleifer et al. (2009); Da et al. (2011); Ben-Rephael et al. (2017); Peress and Schmidt (2018)). We do not consider attention shocks but describe a mechanism of endogenous attention allocation and its effects on mispricing magnitudes and price reactions to information disclosures by firms. A number of theoretical papers model attention allocation but adopt different perspectives. Peng (2005) examines the role of fundamental volatility. Andrei and Hasler (2015) and Kacperczyk, Van

Nieuwerburgh, and Veldkamp (2016) relate attention to the states of the economy. We focus on attention allocation in the cross-section of stocks and its relevance for the cross-sectional variation in price efficiency.

Finally, our results carry important implications for the literature on market anomalies (see Harvey, Liu, and Zhu (2016) for a review). In our setting, investors allocate attention to discover information that is not yet public. However, the same mechanism can be used to describe attention allocation to public news, provided that investors need to pay attention to such news in order to extract valuable signals (e.g., Engelberg, Reed, and Ringgenberg (2012)). Our conclusions suggest that an "anomaly" should not be measured in percentage returns but rather with *unextracted dollar arbitrage profits*. Specifically, an anomaly is stronger, in the sense that it is harder to reconcile with a rational mechanism, not if its percentage returns are higher but if its unextracted arbitrage profits are greater.

1. Theoretical framework

We consider a risk-neutral and two-period ($t = 0,1,2$) economy in which two perfectly divisible stocks are traded. We index these stock by subscript $j = \{L, S\}$. Without loss of generality, we normalize the risk-free rate as zero. Furthermore, since the number of outstanding shares is an arbitrary scaler, we will normalize the total supply of shares as one. The stock price is therefore synonymous with the market capitalization of the firm. Below we describe the financial market structure and the sequence of events.

At $t = 0$, the true (fundamental) values of the two firms, V_L and V_S, are not known to the market. Instead, all market participants observe the market capitalizations of the stocks, K_L and K_S, where $K_L > K_S$. Henceforth, we will refer to the market capitalization ratio $k = \frac{K_L}{K_S}$ as the "relative size" of the stocks. For either firm, its market capitalization deviates from its fundamental value by a factor of $1 + R_j$: $V_j = K_j(1 + R_j)$, where R_j corresponds to the degree of mispricing. This is the same as saying that the stock return between time 0 and the time when the full information becomes known to the market is R or, alternatively, R is the return an arbitrageur would realize if he were able to buy the stock at its time-0 price and sell it at fair value.[6] The agents are not aware by how much each stock is mispriced, i.e. they do not know R_j, but they know the distribution from which each R_j was pulled: $R_j \sim N\left(0, \frac{1}{\tau_0}\right)$. We refer to this distribution as the "prior" and to τ_0 as the common prior precision. In essence, R captures fundamental shocks. A mispricing occurs when a company is hit by a fundamental shock but its market value does not adjust. For it to adjust, either some investors need to investigate the stock and trade on the acquired information to eliminate/reduce the mispricing or the company itself should announce the news to the public.

We now describe the agents. There are two agents: (i) a professional fund manager, who can allocate his investigative resources (henceforth, "attention") across the two stocks to learn about their mispricing, and (ii) a noise trader, who does not receive information signals but still trades the stock. One can think of our theoretical fund manager as all actively managed funds that engage in information discovery to extract profits from the market and the noise trader as all other investors, who trade even in the absence of signals. It is possible that, in reality, the same investor sometimes belongs to the first set and sometimes to the second. At $t = 1$, before trading the stocks, the fund manager observes two noisy signals about the mispricing rates of the stocks: $r_j = R_j + \varepsilon_j$, where $\varepsilon_j \sim N\left(0, \frac{1}{\tau_j}\right)$ and $Cov(R_j, \varepsilon_j) = 0$. τ_j, which captures the fund manager's signal precision in stock j, is proportional to the attention allocated by the manager to the stock. Without loss of generality, we can assume that the coefficient of proportionality is 1, so τ_j is simply the amount of attention allocated to stock j. The manager can obtain a more precise signal on a stock by allocating more attention to that stock; however, the manager's aggregate attention, which we denote c, is a limited resource, and thus $\tau_L + \tau_S \leq c$.[7]

After obtaining the signals and updating his beliefs, the fund manager places a trading order $A_j(r_j)$ in stock j. Noisy traders, on the other hand, place a random trading order $u_j \sim N\left(0, \frac{1}{\tau_u}\right)$, and $Cov(R_j, u_j) =$

[6] The actual arbitrage return will be smaller, since the price would adjust to the arbitrageur's demand.

[7] We impose a technical assumption that c is higher than $(\sqrt{k} - 1)\tau_0$. This assumption ensures the existence of an internal equilibrium. Intuitively, the prior precision τ_0 cannot be too high; otherwise, attention allocation matters little and the investor is happy just sticking with his prior on one of the stocks instead of investigating the stock for additional information.

$Cov(\varepsilon_j, u_j) = 0$. The price of the stock then adjusts according to the aggregate demand. We will assume that the pricing function is linear in demand, i.e. $P_j = K_j\left(1 + \lambda_j(A_j + u_j)\right)$, where λ is the market depth parameter. This choice of the pricing function is motivated by the microstructure literature, notably by the Linear Bayesian Nash Equilibrium of Kyle (1985). The linear pricing function is also convenient computationally. However, what matters conceptually for our agenda are the following two properties of the price: (i) the greater the overall demand for the stock, the more the price adjusts, and (ii) the price does not adjust all the way to the fair value; otherwise, the fund manager would not realize any arbitrage profits. Property (i) ensures that the following logical connection holds: the more attention is allocated to the stock, the more precise the signal, and the closer the equilibrium price is to the fundamentals. Hence, information discovery by the fund manager translates to a higher price efficiency. Property (ii) is necessary to motivate the manager to engage in information discovery.

In the interests of generality, we will treat λ_j as a separate characteristic of a stock without attempting to endogenize it. If we assume that the Kyle model holds for each stock, then $\lambda_j = \frac{1}{2}\sqrt{\frac{\tau_j \tau_u}{\tau_0(\tau_j + \tau_0)}}$ in our setting. However, if a different microstructure mechanism is at work, then λ_j will be different. We do not wish to commit to the exact microstructure mechanism as long as the price obeys the two general properties mentioned earlier. These properties are practically reasonable and are sufficient to guarantee our main conclusions qualitatively, while the linear form of the pricing function allows us to obtain closed-form quantitative solutions. To isolate the effect of size, we will consider the case when the two stocks have the same liquidity, i.e. $\lambda_L = \lambda_S = \lambda$.

The time-1 prices persist until the true firm values V_j are disclosed to the market. On our timeline, this happens at $t = 2$. One can think of time 2 as the "disclosure time", e.g., the day when the firms release their filings which contain some fundamental information previously unknown to the market. At $t = 2$, the stock prices adjust to their respective V_j. The further away P_j was from V_j, the greater the adjustment. These adjustments are empirically observable and they allow us to infer the gap between P_j and V_j.

In what follows, we first consider the benchmark case, in which the fund manager can trade without constraints, and solve for his optimal trading strategy and attention allocation. Next, we introduce the capital constraint that restricts the set of feasible portfolios of the manager. Our main objective is to describe how the total attention c available to the manager is distributed between the stocks and the implications of this distribution for the stock prices P_j and the gap between P_j and V_j, which we dub "uneliminated mispricing". We distinguish between the percentage mispricing, $|P_j - V_j|/V_j$, and the dollar mispricing, $|P_j - V_j|$.

1.1 No capital constraints

We start off with the case when the fund manager is not bounded by capital constraints and can invest any amount in any stock. We solve the model using backward induction. Specifically, we first solve for the optimal trading strategy given the signal r_j, then solve for the fund manager's optimal attention allocation to stocks, and finally derive the corresponding stock prices P_j.

Consider the optimal trading strategy conditional on the signal r_j. Having received the signal, the fund manager Bayesian-updates his belief and forms the posterior $R_j|r_j \sim N\left(\xi_j r_j, \frac{1}{\tau_j+\tau_0}\right)$, where $\xi_j \equiv \frac{\tau_j}{\tau_j+\tau_0}$. His optimal demand conditional on r_j is the solution to the profit maximization problem:

$$\begin{aligned} A_j^* &= argmax_{A_j}\left[K_j\left(1+E\left(R_j|r_j\right)\right) - K_j\left(1+\lambda A_j\right)\right]A_j \\ &= \frac{\xi_j}{2\lambda} r_j. \end{aligned} \tag{1}$$

Plugging the optimal demand into the profit function, we can write the trading profit conditional on the signal r_j as:

$$\begin{aligned} \Pi_j|r_j &= \left[K_j\left(1+E\left(R_j|r_j\right)\right) - K_j\left(1+\lambda A_j^*\right)\right]A_j^* \\ &= K_j\left[\xi_j r_j - \frac{\xi_j}{2} r_j\right]\frac{\xi_j}{2\lambda} r_j \\ &= \frac{K_j \xi_j^2}{4\lambda} r_j^2. \end{aligned} \tag{2}$$

r_j is itself a random variable. It is pulled from a distribution whose variance parameter depends on the attention allocated to the stock. The fund manager will allocate attention so as to maximize his total expected profit from the two stocks. Since it is never optimal to leave any attention unallocated, we can solve for τ_L and note that $\tau_S = c - \tau_L$.

$$\max_{\tau_L \in (0,c)} \left(E(\Pi_L|r_L) + E(\Pi_S|r_S)\right) = E\left(\frac{K_L \xi_L^2}{4\lambda} r_L^2 + \frac{K_S \xi_S^2}{4\lambda} r_S^2\right) = \frac{K_L \xi_L + K_S \xi_S}{4\lambda\tau_0}. \quad (3)$$

The last equality follows immediately from the signal's distribution, $r_j = R_j + \varepsilon_j \sim N\left(0, \frac{1}{\tau_0 \xi_j}\right)$. Solving this optimization problem produces the optimal allocation (τ_L^*, τ_S^*) characterized by Proposition 1.

Proposition 1. The fund manager's optimal attention allocation (τ_L^*, τ_S^*) is given by

$$(\tau_L^*, \tau_S^*) = \left(\frac{\sqrt{k}c + (\sqrt{k}-1)\tau_0}{\sqrt{k}+1}, \frac{c - (\sqrt{k}-1)\tau_0}{\sqrt{k}+1}\right), \quad (4)$$

where $k = \frac{K_L}{K_S}$ is the relative size of firms.

Appendix A presents the formal proof of this proposition and lays out the properties of the optimal attention allocation, which we discuss here. First, the manager allocates more attention to the larger stock, i.e. $\tau_L^* > \tau_S^*$. Intuitively, for a given signal precision, the manager gains a higher expected profit by trading the larger stock and thus has a greater incentive to expend investigative resources to obtain accurate information on the degree of mispricing in this stock. Consistent with this intuition, the manager's attention gap between the two stocks increases with their market capitalization ratio k. Second, the attention allocation gap also increases in the prior precision τ_0. The intuition for this result is well demonstrated by the first-order condition of the fund manager's optimization problem in (3):

$$\frac{K_L}{K_S} = \frac{\xi'_S}{\xi'_L} \equiv \left(\frac{\tau_0 + \tau_L}{\tau_0 + c - \tau_L}\right)^2, \quad (5)$$

where $\xi'_j = \frac{\tau_0}{(\tau_0 + c - \tau_j)^2}$ is the first derivative of ξ_j with respect to τ_j. As the prior precision increases, the effect of attention on the manager's belief update diminishes, i.e. $\xi_S^{'}/\xi_L^{'}$ becomes lower for a given τ_L. Thus, at high levels of τ_0, the manager must allocate more attention to the larger stock to fully utilize its higher dollar profit potential relative to the smaller stock.

We now examine the equilibrium prices P_j. Taking the optimal attention allocation levels (τ_L^*, τ_S^*) and entering the corresponding optimal demands into the pricing function, we obtain: $P_j = K_j\left(1 + \frac{\xi_j^*}{2} r_j\right)$, where $\xi_j^* = \frac{\tau_j^*}{\tau_0 + \tau_j^*}$ is the relative signal precision at the optimal level of attention. We are interested in the percentage mispricing $\left|\frac{P_j - V_j}{V_j}\right|$. Its actual value depends on the realized signal r_j, hence our analysis will focus on the expected value. Technically, this expectation is taken conditional on the initial mispricing R (and thus depends on R), but we will omit this from out notation, since our conclusions hold for any value of R. The expected percentage mispricing is

$$\psi_j \equiv E\left(\left|\frac{P_j - V_j}{V_j}\right|\right) = \frac{E\left(\left|R_j - \frac{\xi_j^*}{2} r_j\right|\right)}{\left|1 + R_j\right|}. \quad (6)$$

This quantity measures the expected percentage mispricing that is equilibrium in the following sense: it is consistent with the optimal allocation of investigative resources by the rational investor. Higher ψ_j indicates a lower price efficiency of stock j, since its price *after* the investor makes his trade deviates more from the true firm value (on average). In other words, it indicates mispricing *that is not eliminated* by the rational actions of the investor. Given that the investor optimally allocates more attention to the larger stock ex ante $(\tau_L^* > \tau_S^*)$ and assuming that the stocks were equally mispriced at the beginning, we obtain the following result.

Proposition 2. At the optimal attention levels, the larger stock exhibits a lower expected percentage mispricing, i.e. $\psi_L < \psi_S$.

In simple terms, this proposition states that if the two stocks happen to experience the same fundamental shock, then more (less) of this shock will be reflected in the price of the larger (smaller) stock and less (more) of the mispricing will remain. The formal proof of Proposition 2 involves a series of steps and is presented in Appendix A. In particular, we show that ψ_j lies in $\left(0, \left|\frac{R_j}{1+R_j}\right|\right)$, implying that the initial mispricing R_j is (partially) adjusted by the (partially) informed demand. The price adjustment is positively associated with the relative signal precision ξ_j^*, which, in turn, increases in the manager's attention level τ_j^*. Since the manager chooses to allocate more attention to the larger stock, its price adjustment towards the true firm value during the disclosure is greater.

1.2 Capital constraints

We now consider the case when the fund manager is constrained in the amount of capital he can invest. This is a highly practical scenario, since all managers are constrained, at least in the short term, by the size of their portfolio and/or the amount of liquid capital they can dispense on new investments. This case differs from the baseline model. A constrained manager is aware that even if he obtains an accurate signal on the larger stock, his capital limit might prevent him from placing an optimal order in, and extracting the maximum value from, that stock. At the same time, the capital constraint is less likely to bind for the smaller stock, where the probability that the optimal order exceeds the constraint is smaller.[8]

Knowing this ex ante, a constrained manager would place less value on the information about the larger stock, since the capital constraint is a friction that drives a wedge between the quality of that information and the manager's ability to use that information to extract profits. Accordingly, in the constrained case, the manager's attention allocation is less skewed towards the larger stock, compared to the unconstrained case considered earlier. This mechanism predicts that the attention allocation gap between large and small stocks is smaller for managers of smaller portfolios than managers of larger portfolios. That is, if we treat fund size as exogenous, then managers of smaller funds would be relatively more content to focus on smaller stocks.

By endogenizing a manager's preference for size, this mechanism can account for the ubiquitous "style" differences among mutual funds. Somewhat simplistically, a fund is not small because it focuses on

[8] In an idealized theoretical setting with two stocks, it is hard to imagine that the fund manager would not be able to place optimal orders because of insufficient capital. However, in a world with many stocks, some small and some large, and frictions that prevent the manager from quickly liquidating his current positions to free up capital for new arbitrage opportunities, the capital constraint is a relevant consideration.

small stocks, but it chooses to focus on small stocks because it is small. When the fund increases in size, its preferences change (the result we demonstrate in the empirical section). We note that the capital constraint does not alter the conclusions for the cross-section of stocks (Proposition 2) but introduces a heterogeneity in the behavior of funds.

2. Mispricing and price adjustments

Our main theoretical prediction is that prices of smaller stock should, on average, deviate more from the stocks' fundamental values in percentage terms. Since fundamental value is not directly observable, our empirical tests will focus on the price adjustments at times when previously non-public material information is released to the market. This approach mirrors Proposition 2, which states that smaller stocks should experience a greater price adjustment when the market learns about the state of fundamentals. The adjustment can be either positive or negative, depending on whether the market under- or over-valued the stock before the news release. We are interested in the absolute magnitude of this adjustment, since our model does not differentiate between under- and over-valuations.[9]

We consider earnings announcements as major information events. Earnings announcements occur regularly and provide a rich sample of observations, both in the cross-section and time-series. The earnings information is material and the reports are standardized, alleviating concerns about different types of information being released by different firms or different modes of communication chosen by corporate managers.

In what follows, we investigate how the percentage stock price adjustment during earnings announcements varies with stock characteristics commonly featured in finance research. Our model formalizes the effect of size on mispricing (before the fundamental information is released to the market). However, it does so independently of other frictions, allowing for a possibility that these frictions also play

[9] In our model, positive and negative deviations of the price from fundamentals are treated symmetrically. Our empirical results are also largely symmetric between positive and negative news shocks. Thus, our agenda is not directly related to the size premium literature (e.g., Banz (1981); Fama and French (1992); Hameed, Lof, and Suominen (2017)). To generate a size premium, we would need to introduce asymmetric frictions for buying and selling. This extension is not essential to the economic mechanism we study and can be pursued in future work.

a role. Comparing the effect of size and that of other stock characteristics traditionally believed to affect price efficiency is an empirical task that we address in the sections below.

2.1 Sample and variables

Our sample consists of the common stocks in the CRSP universe and covers the period from 1990 to 2016. We apply the following standard filters: exclude stocks not listed on NYSE, AMEX, or NASDAQ, and microcap stocks, which have a price and a market capitalization of less than $1 and $10 million, respectively, (before any split adjustment) at the beginning of the observation month.[10]

Our analysis makes use of the following data: pricing and volume data from CRSP, accounting variables from Compustat, earnings announcements records from Compustat and I/B/E/S, stock lending supply and fees from Markit, media coverage and news content from RavenPack and Factiva, intraday trading from TAQ, and institutional holdings from Thomson Reuters 13F and Mutual Funds Holdings databases. In Section 3, we compute proxies for institutional attention using data on searches in the Bloomberg terminal and downloads from the SEC's EDGAR system. In Section 4, we construct measures of trading discreteness based on the institutional trading data from ANcerno.

We focus on a set of stock characteristics that have been linked to price efficiency in prior studies. These characteristics can be roughly categorized into the following classes: liquidity (e.g., Amihud (2002); Chordia, Roll, and Subrahmanyam (2008)), limits to arbitrage (e.g., Shleifer and Vishny (1997); Asquith, Pathak, and Ritter (2005); Saffi and Sigurdsson (2011)), retail investor participation (e.g., Kumar and Lee (2006); Barber, Odean, and Zhu (2009)), and information environment (e.g., Fang and Peress (2009); Engelberg and Parsons (2011); Tetlock (2011)).

Table 1 reports the summary statistics for the main variables of interest. *AmIlliq* and *Spread* proxy for illiquidity costs associated with trading in a stock. *AmIlliq* is the average daily Amihud (2002) illiquidity

[10] Following Gao and Ritter (2010), we adjust the trading volume from NASDAQ to avoid double counting of trades in the earlier period of the sample. We divide NASDAQ trading volume by 2.0 in the period before February 2001, by 1.8 in the period between February 2001 and December 2001, and by 1.6 in the period between January 2002 and December 2003. See Appendix B of Gao and Ritter (2010) for details.

measure for a stock in a quarter, computed as the absolute return divided by the daily dollar trading volume in the stock. We set the top 1% of *AmIlliq* values in each quarter to missing to minimize the influence of outliers.[11] *Spread* is the average daily relative spread for a stock in a quarter; it equals the difference between the closing ask and bid prices divided by the midpoint price. *AmIlliq* remains highly skewed even after trimming, but *Spread* is more evenly distributed with a mean and a median of 2.22% and 1.23%, respectively.

BorrowingFee and *Lendable* are measures of arbitrage (short-sale) constraints. These variables are reported by Markit (previously Data Explorers), a company that obtains securities lending data directly from custodians and prime brokers (see Saffi and Sigurdsson (2011) and Engelberg, Reed, and Ringgenberg (2018) for details). *BorrowingFee* is the average daily indicative borrowing fee for a stock in a quarter and *Lendable* is the fraction of outstanding shares available for borrowing at the beginning of the quarter. The average *BorrowingFee* and *Lendable* in our sample are 1.54% and 14.96%, respectively. Higher (lower) *BorrowingFee* (*Lendable*) generally indicates tighter short-sale constraints. *Lendable* has a high correlation with size (0.41), whereas *BorrowingFee* has a lower correlation (-0.18) likely due to high skewness in the distribution of fees.

RetailOwn and *RetailTrading* proxy for retail investors' market participation. *RetailOwn* equals one minus the proportion of institutional ownership in a stock at the beginning of the quarter as computed from Thomson 13F filings. *RetailTrading* captures the proportion of trades likely originated by retail investors and is calculated as the average daily ratio of the number of "small" trades to the total number of trades in the stock (computed from TAQ). The validity of this measure is confirmed until 2000 (see Barber, Odean, and Zhu (2009)), hence we only use it in the earlier periods of the sample.[12] Following Lee and Radhakrishna (2000) and Barber, Odean, and Zhu (2009), we define a trade as "small" if its value does not

[11] Amihud (2002), Acharya and Pedersen (2005), and Karolyi, Lee, and van Dijk (2012), among others, truncate or take the log of Amihud illiquidity to minimize the impact of outliers.
[12] Interpreting *RetailTrading* post 2000 is problematic due to an emerging trend of order splitting. Kyle, Obizhaeva, and Tuzun (2012) and Hu et al. (2018) cite decimalization and the rise of algorithmic trading as major drivers of increased order splitting, which starts in early 2000s.

exceed $5,000. The measures of retail investors' participation are highly correlated with stock size: the correlation coefficients are -0.63 for *RetailOwn* and -0.70 for *RetailTrading*, respectively.

The last group of variables captures different facets of the stock's information environment. *MediaCover* is the average daily number of news articles about a stock in a quarter in RavenPack.[13] We only count articles with a relevance score of 100 (this score indicates that the news story is specifically about the firm). *Tangibility* measures the predominance of quantitative, or tangible, information and is equal to the average ratio of the number of numeric characters to the total number of characters computed over all the articles about the firm in a given quarter in Factiva (see Chuprinin, Gaspar, and Massa (2018)). Finally, *FogIndex* and *10KLength* are measures of complexity and readability of a firm's annual report (see the detailed description of these variables in Li (2008)).[14] These measures are available annually, and we take their most recent versions at the beginning of each quarter.

[Insert Table 1 here]

Our main measure of price correction is the absolute value of the cumulative market-adjusted stock return over the first three trading days following the earnings announcement (including the event day). The three-day window is long enough so that the return is not contaminated by transitory price fluctuations. At the same time, this window is so short that we can conduct ours tests without taking a stance on the asset pricing model driving expected returns. Most of the three-day return can be safely attributed to the announcement reaction and not to the expected return component. Furthermore, we conduct our analysis separately for positive and negative surprises to show that the price adjustment is symmetric and cannot be interpreted as an anticipated compensation for some risk.[15]

---

[13] Recent studies using RavenPack data include Kolasinski, Reed, and Ringgenberg (2013); Dang, Moshirian, and Zhang (2015); Gao, Parsons, and Shen (2017); and Yu, Zhang, and Zhang (2017).
[14] We thank Feng Li for making this data available on his webpage: http://webuser.bus.umich.edu/feng/.
[15] One can argue that a longer-horizon return, e.g. over thirty trading days, is a better measure of the full price reaction to the earnings news, since it accounts for the post-earnings announcement drift. This drift is small in magnitude relative to the announcement return and is similar for positive and negative surprises, hence our results are almost identical whether we use the three-day or the thirty-day measure. The results for the thirty-day measure are reported in Tables B2-B4 of Appendix B.

A number of studies proxy for firm-specific information by idiosyncratic volatility or R^2 from the regression of stock returns on the market (e.g., Morck, Yeung, and Yu (2000); Jin and Myers (2006); Jiang, Xu, and Yao (2009)). These measures are not suitable for our purposes and do not align tightly with the predictions of the model. First, our objective is not to divide information into firm-specific and market components but to measure how much of the information indicative of fundamentals is not priced before the public disclosure. Second, idiosyncratic volatility and R^2 are generally problematic as measures of information, since they are also affected by firm-specific risk (Xu and Malkiel (2003); Jin and Myers (2006); Cao, Simin, and Zhao (2008)), noise in prices (Han and Lesmond (2011); Lee and Liu (2011); Morck, Yeung, and Yu (2013)), and numerous other issues (e.g., see Rajgopal and Venkatachalam (2011); Li, Rajgopal, and Venkatachalam (2014); Kelly (2014)).

To construct the sample of earnings announcements, we follow the procedure in DellaVigna and Pollet (2009) and Hirshleifer, Lim, and Teoh (2009). We retrieve the dates of quarterly earnings announcements from Compustat and I/B/E/S and discard observations where the difference in announcement dates between the databases is more than five calendar days. For the remaining observations with different dates we designate the earlier date as the event day.[16] If the earnings are released after 4:00pm EST we reassign the announcement date to the next trading day.

2.2 Regression analysis

In the first series of tests, we compare the effect of size on the price adjustment to the effect of other stock characteristics in the following specifications:

$$Abs\ Three\text{-}Day\ Return_{i,t} = \alpha + \beta ExplanatoryVariable_{i,t} + \delta_t + \varepsilon_{i,t} \quad \text{(7a)}$$

and $$Abs\ Three\text{-}Day\ Return_{i,t} = \alpha + \beta Size_{i,t} + \gamma CompetingVariable_{i,t} + \delta_t + \varepsilon_{i,t}. \quad \text{(7b)}$$

[16] DellaVigna and Pollet (2009) compare Compustat and I/B/E/S dates and conclude that from 1990 the earlier of the two dates is usually the actual announcement date.

Abs Three-Day Return$_{i,t}$ is the absolute return of stock *i* over the period [0,2] days around the earnings announcement in year-quarter *t* net of the CRSP value-weighted index return.[17] All independent variables are measured before the event to avoid any potential feedback effects with the announcement return. Specifically, *AmIlliq*, *Spread*, *BorrowingFee*, *RetailTrading*, and *MediaCover* are computed as the respective averages over the period of thirty trading days ([-32,-3]) prior to the event, while *Size*, *Lendable*, *RetailOwn*, *Tangibility*, *FogIndex,* and *10KLength* are computed as of 32 trading days prior to the event.[18] In all regressions we include year-quarter fixed effects and cluster standard errors by stock.

Panel A of Table 2 shows the univariate regressions. For both positive and negative surprises—i.e. events where the three-day return is positive or negative, respectively—we observe the following pattern: smaller stocks, less liquid stocks, more arbitrage-constrained stocks, and stocks with higher retail investor participation experience larger price corrections. Lower media coverage and greater information ambiguity, as proxied by information tangibility and the fog index, are also associated with larger corrections. These results are expected, since all frictions inhibit price discovery. However, size is clearly the strongest predictor of the correction magnitude with a *t*-statistic of -31.8 (-34.2) for positive (negative) surprises. To compare, other significant variables—*AmIlliq*, *Spread*, *BorrowingFee*, and *RetailTrading*—have *t*-statistics of 10.3 (4.5), 11.9 (13.4), 8.2 (10.8), and 15.4 (13.8), respectively.

In Panel B of Table 2, we conduct a pairwise comparison of stock size and other stock characteristics in the bivariate regressions. Size strongly dominates all other variables; in particular, its *t*-statistic is always higher (in absolute terms) and does not fall below 20 in most cases. Some competing

[17] We alert the reader to a slight discrepancy between the theoretical and the empirical version of the price adjustment variable. In the model, we measure percentage adjustment as $|P_j – V_j|/V_j$. In this ratio, the endogenous variable P_j appears only in the numerator, and the ratio has convenient statistical properties. In the regression, the absolute return at the announcement most closely corresponds to $|P_j – V_j|/P_j$. This distinction is minor since the difference between P_j and V_j is much smaller than the size itself. For example, it matters little whether we measure (105–100)/105 or (105–100)/100. Transforming the announcement return so it corresponds to $|P_j – V_j|/V_j$, as in the model, yields similar results.

[18] Here and henceforth, we use the log transformation of the 10-K report length following the literature convention (Li (2008); Miller (2010); Lee (2012)).

variables—namely, *AmIlliq*, *Spread*, *Lendable*, *RetailOwn*, and *MediaCover*—become insignificant or even change sign when included jointly with size.[19]

[Insert Table 2 here]

2.3 Non-parametric analysis

The regression tests above allow for an easy evaluation and comparison of statistical significance. Yet, the economic effects are obscured by differences in measurement units and likely non-linearities and are hard to interpret constructively. To address these issues, we perform the following non-parametric analysis. We sort stocks into five quintiles by a given characteristic and compute the average absolute price adjustment in each quintile. The difference between quintiles 1 and 5 captures the non-linear effect of a given variable on the pre-event mispricing and can be directly compared across all variables.[20]

Table 3 reports the results. The difference in *Abs Three-Day Return* between the extreme size quintiles is -2.99% (-2.48%) for positive (negative) surprises with a *t*-statistic of -38.0 (-44.7). *AmIlliq*, S*pread*, and *BorrowingFee* have lower economic magnitudes, with Q5-Q1 differences of 2.35% (1.77%), 2.27% (1.71%), and 2.14% (2.20%) for positive (negative) surprises, respectively, and consistently lower *t*-statistics. For most other variables, the Q5-Q1 gap does not exceed 1%.

[Insert Table 3 here]

To compare the relative contributions of different characteristics, we employ sequential double-sorting. First, we sort stocks into five bins by a given characteristic. Then, within each bin, we further sort stocks into five size quintiles and compute the average absolute announcement return in each quintile. The difference between quintiles 1 and 5 quantifies the effect of size conditional on a pre-sorting characteristic. Panels A1 and A2 of Table 4 report the results of this analysis. For comparison, Panels B1 and B2 report the results of the reverse procedure, where stocks are first sorted by size and then by another characteristic.

[19] *Size* also dominates in regressions where different competing variables are included jointly (see Table B1 of Appendix B). We cannot fit all the variables in one regression due to different periods of data availability.
[20] Note that this procedure is distinct from portfolio analysis, since we condition on the ex-post outcome of the earnings announcement surprise: positive or negative. We are not trying to predict returns but only quantify the magnitude of the price correction as a function of different stock characteristics.

Size always retains significance after controlling for confounding effects (Panels A1 and A2). In most cases, the difference in *Abs Three-Day Return* between the extreme *Size* quintiles (Q5-Q1) is around 2%-3%. This magnitude is similar to that in the single-sort analysis in Table 3, suggesting that most of the effect of *Size* is not subsumed by another stock characteristic. In contrast, only *BorrowingFee*, *Tangibility*, and *FogIndex* retain the sign and residual significance after pre-sorting by *Size* (Panels B1 and B2). Other variables tend to flip signs and/or lose significance after conditioning on *Size*.

[Insert Table 4 here]

Overall, the results in this section indicate that smaller stocks experience greater price corrections when companies release fundamental information. This effect is symmetric for positive and negative surprises and cannot be explained by theories grounded in risk preferences (e.g., that size is a proxy for unknown risk that requires compensation). Size comfortably dominates other stock characteristics commonly associated with pricing frictions in both the regression and non-parametric analysis. These results imply that there is something specific about size that inhibits information discovery until a point when the information is publicly disclosed. This effect is economically non-trivial and is consistent with a rational allocation of investigative resources in the cross-section of stocks, as predicted by the model.

3. The attention channel

So far, we provided evidence in support of our main cross-sectional prediction, namely that, absent public disclosures, prices of small stocks deviate more from their fair values. Even though the model delivers this prediction exactly, the exact channel requires further empirical support. There are several intermediate steps in the logic chain that require investigation.

One is the relation between attention allocated to the stock before the news release and the price adjustment during the release. This link is theoretically straightforward—less attention leads to a bigger price adjustment—but not empirically so, since real markets abound in confounding effects that do not feature in the model. We investigate this relation in Section 3.2 below.

Another is the result that size is related to the price adjustment through the trading decisions of investors: stocks that are traded less because they receive less attention subsequently experience larger price adjustments during earnings announcements. We investigate this channel in Section 4.

Finally, the very connection between stock size and attention paid to the stock by active fund managers (Proposition 1 in the model) is not self-evident. Indeed, we can expect institutional attention and size to be strongly correlated, but whether size is the primary driver of attention or just a correlate of some other variable, which has a first-order effect, is unclear.

3.1 Drivers of attention

We begin with the last item on the list: the relation between stock size and institutional attention. It is important to measure attention by institutional investors, since the optimizing agent in our model is likely to be an institutional money manager who commands investigative resources. Working proxies for institutional attention are inevitably crude and do not capture the whole range of investigative resources allocated to a stock. Hence we view the results in this section as only suggestive. The two commonly used measures of institutional attention are the search activity for stock-specific news in the Bloomberg terminal and the downloads of corporate filings from the SEC's EDGAR online system.

Bloomberg terminal is a service widely used by investment professionals. Bloomberg records the number of times users actively search for news on a stock and the number of times they click on a news story. Based on these inputs, Bloomberg calculates a standardized news search score for each stock, assigning higher weights to more active searches. This score takes discrete values from 0 to 4, where 0 indicates that the daily search activity for the stock is in the lowest 80% of the distribution over the past month and 1 (2, 3, 4) indicates that the daily search activity is within 80-90% (within 90-94%, within 94-96%, above 96%) of the distribution.

Even though the score is measured relative to the distribution, it is reasonable to expect that if investors pay close attention to a stock then their search activity for that stock will spike when the company experiences a fundamental shock. In contrast, if investors do not pay attention to a stock, no search spikes should be evident. Bloomberg search score was first used as a measure of institutional attention by Ben-Rephael, Da, and Israelsen (2017), who provide a detailed description of the data.[21] We download daily news search scores for Russell 3000 stocks directly from the Bloomberg terminal for the whole period of data availability—from February 2010 to December 2016.[22]

Our goal is to examine how size compares against other stock characteristics that can conceivably drive attention. First, institutional investors might allocate less attention to stocks that are costlier to arbitrage. Second, the search score might simply reflect the level of media coverage in a stock rather than strategic resource allocation. Third, investors might devote more attention to stocks with a higher fundamental volatility or asset-specific shocks (Peng (2005)). Finally, major firm characteristics, such as the book-to-market ratio, leverage, and R&D expenses, are often used as first-pass screening variables in stock selection, and may also affect news search activity.

To compare the relative effects of different attention drivers, we run quarterly panel regressions of institutional attention on the stock characteristics as follows:

$$InstAtten_{i,t} = \alpha + \beta ExplanatoryVariable_{i,t} + \delta_t + \varepsilon_{i,t} \quad \text{(8a)}$$

$$\text{and } InstAtten_{i,t} = \alpha + \beta Size_{i,t} + \gamma CompetingVariable_{i,t} + \delta_t + \varepsilon_{i,t}, \quad \text{(8b)}$$

where $InstAtten_{i,t}$ is the average daily news search score from Bloomberg for stock *i* in year-quarter *t*. The set of explanatory variables includes *Size*, *AmIlliq*, *Spread*, *BorrowingFee*, *Lendable*, *RetailOwn*, *MediaCover*, and *FogIndex*, defined as in Table 1 (we lose variables that are not available after 2010). We also include a proxy for the fundamental volatility—*EarnVolatility*—calculated as the standard deviation of quarterly earnings reported by firm *i* over the past three years,[23] and standard firm characteristics: *BM*—the most recent book-to-market ratio of firm *i* available at the beginning of year-quarter *t* (calculated as in Fama and French (1993)), *Leverage*—the most recent ratio of total debt to total book value of assets of firm

[21] Other recent studies using the Bloomberg attention measure include Ben-Rephael et al. (2017) and Ungeheuer (2017).
[22] The data is missing for two short periods in our sample: from 6 December 2010 to 7 January 2011 and from 17 August 2011 to 2 November 2011.
[23] We require at least four quarters of non-missing observations in the past three years to calculate this measure. Our results are similar if we use the standardized unexpected earnings (SUE) instead of the earnings volatility.

i at the beginning of year-quarter *t*, and *R&D*—the most recent ratio of research and development expenses to gross sales of firm *i* at the beginning of year-quarter *t*.

Table 5 reports the results. As before, we first compare the effects of different variables in the univariate regressions (Panel A). *Size* is the strongest predictor of institutional attention with a *t*-statistic of 57.7. An interquartile-range increase in *Size* is associated with an increase in the search score of 0.11 (to compare, the average search score in our sample is 0.67). The effects of other stock characteristics are much weaker. For example, *RetailOwn* is the second strongest predictor with a *t*-statistic of -20.4. Notably, this variable is the complement of institutional ownership and should be related to institutional attention almost mechanically; however, its effect is still weaker than that of *Size*. Consistent with Peng (2005) and Kacperczyk, van Nieuwerburgh, and Veldkamp (2016), higher fundamental volatility also increases attention (*t*-statistic of 5.8).

Panel B reports the bivariate regressions, in which *Size* and other variables are compared pairwise. The effect of *Size* is virtually unaffected by the addition of a competing variable. The coefficients on *Size* are close to that in the univariate regression (0.2023-0.2364 vs. 0.2154), while their *t*-statistics do not fall below 40. In contrast, most of the competing variables flip sign or lose significance in the presence of size. The variables that retain the sign and significance, namely, *MediaCover*, *EarnVolatility*, and *Leverage*, experience a marked decline in coefficient magnitudes (from 0.0396 to 0.0144, from 0.2101 to 0.1164, and from 0.3298 to 0.1781, respectively).

[Insert Table 5 here]

In Table B5 of Appendix B we report the results from multivariate regressions and also consider an alternative proxy for attention, *EDGAR*, which measures how often users access the firm's financial filings through the SEC's EDGAR system.[24] In all specifications, *Size* remains the strongest driver of attention with stable economic magnitudes.

[24] Although EDGAR filings are available to everyone, institutions are more likely than individuals to use the EDGAR system (Ben-Rephael, Da, and Israelsen (2017)). We closely follow Loughran and McDonald (2017) and Ben-Rephael, Da, and Israelsen (2017) in filtering the EDGAR data. We exclude retrievals of index files and retrievals that result in error message, as well as requests of web crawlers and requests by single users accessing more than 50 filings on a

In summary, we find that size is the dominant driver of institutional attention in the cross-section of stocks. This result aligns with Proposition 1 of the model and supports the claim that the effect of size on price efficiency is likely first-order.

3.2 Attention and price adjustments

In this subsection, we investigate how the stock price reaction to an earnings announcement depends on the institutional attention to the stock *prior to* the announcement. The structure of the regression is the same as that in (7b), but the main independent variable is now $InstAtten_{i,t}$, computed as the average daily Bloomberg search score for stock *i* in the period [-32,-3] days relative to the earnings announcement in year-quarter *t*. Besides *InstAtten,* the set of explanatory variables includes *AmIlliq*, *Spread*, *BorrowingFee*, *Lendable*, *RetailOwn*, *MediaCover*, and *FogIndex*, defined as in Section 2 (we lose *RetailTrading* and *Tangibility* as they are not available after 2010).

We add one more explanatory variable in this analysis—*GoogleSearch*, computed as the most recent monthly value of the Google search index from Google Trends for stock *i* available 32 trading days before the event.[25] *GoogleSearch* has been suggested as a proxy for retail investors' attention (see Da, Engelberg, and Gao (2011)). Comparing the effects of retail and institutional attention is instructive, since we do not expect retail investors to be rational optimizers of investigative resources. Furthermore, including Bloomberg and Google-based attention measures in the same regression helps control for the common factors that draw attention of both individuals and institutions (e.g., media stories that are not picked up by our media coverage variable *MediaCover*).

given day. See Loughran and McDonald (2017) for a detailed description of the EDGAR database. The data on EDGAR queries is available from 2003.

[25] See https://trends.google.com/trends for details. We collect the data for 6,731 stocks (7,142 tickers) that satisfy the filtration criteria from Section 2 in the period from 2004 to 2016. Following Da, Engelberg, and Gao (2011), we search for the tickers rather than company names in Google Trends and exclude ambiguous abbreviations. We also exclude tickers consisting of one or two letters, as in Cziraki, Mondira, and Wu (2018). Google standardizes search index values between 0 and 100 and restricts the maximum number of downloaded search terms to five at a time. One needs to include a reference search term in each download for a cross-sectional comparison. We use "Russell 3000" as the reference term. Niessner (2015) discusses this procedure in detail.

Table 6 reports the results. The coefficient on *InstAtten* is negative and highly significant in all specifications. An increase in *InstAtten* of 1 point reduces the absolute three-day announcement return by 0.74%-0.95%.[26] The effect of institutional attention is symmetric for positive and negative surprises, which is not the case for most other explanatory variables. Even though *InstAtten* is a noisy proxy of committed investigative resources, its statistical relation to the price correction, as reflected in *t*-statistics, is still considerably stronger than that of any other variable.

We also note the following auxiliary results: i) controlling for *MediaCover* or *GoogleSearch* does not make *InstAtten* weaker, and ii) *MediaCover* is never significant while *GoogleSearch* is significant for positive surprises only. This suggests that the true price discovery depends mostly on the investigative efforts of institutional investors and is largely independent of the public information environment.

[Insert Table 6 here]

In summary, this section provides evidence on two key theoretical relationships: one between size and institutional attention and the other between institutional attention and mispricing. The missing link so far is the trading activity of professional investment managers—the subject of the next section.

4. Analysis of trading activity

An investor who pays little attention to a stock is less likely to detect a fundamental shock and update his prior. Absent reliable new information, the old prior represents the investor's best guess of the stock's fair value (provided that the fundamental shock is centered at the prior, as in the model). Thus, an inattentive investor would not rebalance his position in the stock for extended periods of time. Empirically, the pattern of position changes in neglected stocks should look discrete and the positions themselves should be "sticky" compared to stocks that receive more attention.[27] In this section, we investigate the discreteness/continuity of trading by institutional investors and its relation to mispricing.

[26] Using the thirty-day return yields similar results (see Table B6 of Appendix B).

[27] In theory, even when an investor receives information on a few stocks, he should rebalance his whole portfolio. However, since the required theoretical adjustments in informationally silent stocks are small, investors typically update specific positions while leaving the rest intact. We observe this pattern, in particular, in mutual fund portfolios, where select positions are updated more regularly than others.

4.1 Measuring trading discreteness

We base our analysis on the transaction data from ANcerno. This data covers daily trades of plan sponsors and investment managers and is described in detail in Puckett and Yan (2011) and Hu et al. (2018). According to Hu et al. (2018), the trading volume of ANcerno institutions accounts for about 12% of the total volume in CRSP (and about 15% of all institutional volume) in the period from 1999 to 2011.

Our sample ranges from 1999 to 2010. For this period, ANcerno provides an identification code for each institutional client account. We exclude accounts with less than eight quarters of data; this leaves us with 355 unique client IDs which we call "funds" for simplicity of exposition. Panel A of Table 7 shows that around 150 funds from our sample conduct at least one trade each quarter (*NumActiveFunds*). An average fund trades 420 unique stocks (*NumStocks*) and executes 14,002 trades (*NumTrades*) per quarter, which corresponds to approximately $1,445 million in volume (*TradingVolume*). These numbers are consistent with those reported by Puckett and Yan (2011), who document an average quarterly trading volume of $1,285 in their sample period.

The distributions of *NumTrades* and *TradingVolume* are skewed, with the respective medians of 1,438 trades and $134 million. This pattern highlights the heterogeneity of funds in our sample, which we exploit later in the analysis. The statistics for the dollar values of stocks bought (*StocksBought*) and sold (*StocksSold*) are similar, suggesting that ANcerno institutions are not special in the sense that they are not consistently growing or declining in size throughout the sample period.

[Insert Table 7 here]

For each combination of fund, stock, and calendar quarter, we compute two measures of trading discreteness. The first measure, *TimeBetweenTrades*, captures discreteness over time. It is equal to the average time (in calendar days) between trades of a fund in a given stock in a quarter. The second measure, *DistanceBetweenLots1*, captures discreteness of the trading orders and equals the average interval between unique trading lots (number of shares traded) by a fund in a stock in a quarter. To construct this measure, we calculate the average difference between the adjacent lots and standardize it by the number of shares outstanding at the end of the quarter. The low correlation (0.07) between *TimeBetweenTrades* and

DistanceBetweenLots1 suggests that these variables capture different facets of trading discreteness and provide complementary information.

Panel B of Table 7 reports summary statistics for these measures. An average (median) time interval between consecutive trades of the same stock is 9.47 (4.71) days. Note that *TimeBetweenTrades* is computed conditional on a fund trading the stock at least twice in the quarter. The range of variation in this interval is substantial: from P10 of 0.4 days to P90 of 24.7 days. An average (median) standardized distance between lots is 0.0058% (0.0011%), while the P10-P90 range is 0.012%.

To conduct stock-level analysis, we compute the stock-specific versions of the discreteness measures by taking the weighted averages of the corresponding fund-level measures using weights proportional to the trading volumes of the funds in the stock. The resulting stock-level variables indicate how discretely or continuously ANcerno institutions trade the stock on average. For robustness tests, we complement these variables with two additional measures: *DistanceBetweenLots2* and *LotConcentration*. These measures do not require an account-level ID and can be computed over the fully anonymized sample, which extends to 2013 in our vintage of ANcerno. *DistanceBetweenLots2* measures the average interval between unique trading lots for a stock in a quarter across all funds that traded the stock. *LotConcentration* measures the concentration of unique trading lots and is equal to the Herfindahl index which sums up the squared fractions of different unique trading lots in a stock in a quarter. The fraction of each trading lot is defined as the number of trading lots with a given share volume divided by the total number of trades executed by all funds in the stock in that quarter. Panel C of Table 7 reports the summary statistics for all stock-level variables. The correlation between *DistanceBetweenLots1* and *DistanceBetweenLots2* is 0.64, suggesting that we do not lose too much information about lot discreteness in the full sample.

4.2 Relation to price adjustments

First, we verify that stock size and attention have a strong effect on trading discreteness. Table 8 reports the results from quarterly regressions of the discreteness measures on *Size*, *InstAtten,* and *EDGAR*.[28] There is a strong negative relation (*t*-statistics exceeding 10 in absolute magnitude) between all measures of trading discreteness and size/attention.

[Insert Table 8 here]

Next, we examine how trading discreteness measured before an earnings announcement relates to the stock price correction at the announcement. Regressing the absolute three-day announcement return on the discreteness measures reveals no clear effects: the coefficients are unstable and generally lack significance. It is perhaps not surprising to find this non-result. There are at least two competing mechanisms that predict the opposite effects of trading discreteness on prices. First is the attention channel: positions in neglected stocks are not updated timely, hence these stocks are more likely to stay mispriced and should experience greater price corrections at the disclosure. The offsetting channel is that investors are less likely to update their positions in stocks that *they know are not mispriced*. This channel predicts a negative relation between trading discreteness and the price correction magnitude.

To disentangle these effects, we decompose trading discreteness into two components: one driven by stock size/attention and the residual. Specifically, we regress trading discreteness on size (at the beginning of the quarter) or attention (during the quarter) and extract the predicted and the residual values from the regression. Then, we regress the absolute three-day announcement return separately on these two components (*Abs Three-Day Return*$_{i,t}$ is defined as in (7a) and (7b)):

$$Abs\ Three\text{-}Day\ Return_{i,t} = \alpha + \beta Predicted(Trading\ Discreteness)_{i,t} + \delta_t + \varepsilon_{i,t} \quad (9a)$$

and $$Abs\ Three\text{-}Day\ Return_{i,t} = \alpha + \beta Residual(Trading\ Discreteness)_{i,t} + \delta_t + \varepsilon_{i,t}. \quad (9b)$$

[28] In Table B7 of Appendix B we explore other determinants of trading discreteness. In all regressions, we standardize coefficient magnitudes by dividing raw *TimeBetweenTrades* by 100 and multiplying raw *DistanceBetweenLots1* and *DistanceBetweenLots2* by 100.

Table 9 reports the results. Different panels correspond to different measures of trading discreteness. In all cases, size-driven trading discreteness is strongly positively related to the absolute announcement return (*t*-statistics above 20). This result holds symmetrically for positive and negative surprises. The economic magnitude of this effect is similar to that in Table 3: the difference in *Abs Three-Day Return* between the extreme quintiles of the size-driven trading discreteness is 2.21%-2.72%. The residual component, in contrast, is negatively related to the absolute announcement return.

The results are largely similar for the attention-driven trading discreteness: higher attention-driven discreteness predicts a stronger price correction at the announcement. However, there is one instance where this effect is insignificant, namely, if we focus on negative surprises and proxy for attention with the download activity from EDGAR. A possible explanation for this weak effect is that the EDGAR activity is a particularly noisy proxy for attention in a stock that is overvalued due to high demand. Institutional investors who already hold the stock prefer to subscribe to automatic updates from their data providers instead of downloading reports from a public repository such as EDGAR.

[Insert Table 9 here]

Overall, the evidence suggests that a lack of continuity in institutional trading that can be attributed to the effects of size and attention is associated with mispricing that is less likely to be eliminated early and is more likely to persist until the disclosure, where it causes a major price correction. These results hold robustly for different measures of trading discreteness.

5. Fund size heterogeneity

In this section, we exploit the heterogeneity in attention allocation incentives arising due to the variation in sizes of investment portfolios of different funds. As we noted in Section 1.2, it is harder for a small fund to extract large dollar profits from an arbitrage opportunity, even if the fund is fully aware of its existence. In a short term, the fund's dollar gains are limited by the amount of capital the fund presently controls. Thus, the benefit of having accurate information on larger stocks declines with fund size.

Ideally, we would like to examine how the allocation of investigative resources across stocks varies with portfolio size. Unfortunately, we are not aware of a reliable measure of fund-specific attention allocation. However, noting the tight connection between attention and trading discreteness, we can investigate variation in trading discreteness (or position stickiness) over fund-stock pairs.

Our first test utilizes ANcerno data. Specifically, we use the fund-stock trading discreteness measures we introduced in Section 4. To understand how the effect of stock size on trading discreteness varies with fund size we estimate the following regression:

$$Trading\ Discreteness_{j,i,t} = \alpha + \beta StockSize_{i,t} + \gamma FundSize_{j,t}$$
$$+\theta StockSize_{i,t} \times FundSize_{j,t} + \delta_t + \varepsilon_{j,i,t}, \quad (10)$$

where *Trading Discreteness*$_{j,i,t}$ is either *TimeBetweenTrades* or *DistanceBetweenLots1*, *StockSize*$_{i,t}$ is the natural logarithm of the market capitalization of stock *i* at the beginning of year-quarter *t*, and *FundSize*$_{j,t}$ is the proxy for the dollar size of fund *j* in the year of quarter *t*. The standard errors in this regression are double-clustered by stock and fund. We consider two proxies for fund size: *FundVolume*, calculated as the overall trading volume of fund *j* in the year, and *BigFund*—a dummy variable equal to one if fund *j* has an above-median *FundVolume* in the year in the ANcerno sample, and zero otherwise.[29]

The results are reported in columns (2)-(3) and (5)-(6) of Table 10. For comparison, columns (1) and (4) report the regressions run at the fund-group and stock level, where *BigFund* indicates the fund group—small (0) or large (1); these specifications smooth the transition from stock to fund-stock level analysis and reveal the consistency of the empirical patterns. The interaction coefficient is negative and is both statistically and economically significant for both measures of discreteness. For the large-fund group, the coefficient on *StockSize* is equal to -0.7552 (=0.8214-1.5766), indicating that large funds trade smaller stocks more discretely. In contrast, the positive coefficient on stock size (0.8214) indicates that small funds trade smaller stocks less discretely (i.e. more frequently). The interaction coefficients in the fund-stock

[29] The exact portfolio composition and the actual portfolio size are not reported by ANcerno. Hence, in this test, our measures of fund size are based on the fund's total trading volume.

level regressions are negative and highly significant, confirming that larger funds trade smaller stocks relatively more discretely: by 83% and 174%, respectively, based on columns (3) and (6).

It is worth pointing out that the observed effects cannot be explained by trading discreteness driven by liquidity considerations. If large funds split their small-stock orders into finer lots to reduce price impact, we would observe a relatively more continuous trading of these funds in these stocks, contrary to the actual findings.

[Insert Table 10 here]

Overall, the frequency of trading is more sensitive to the size of an investment opportunity if the fund commands more capital.[30] The ANcerno data allows us to compute trading discreteness; however, its main limitation is that we cannot observe any account characteristics apart from trading. We therefore complement our investigation with the analysis of mutual fund holdings (from Thomson Reuters). The data on mutual fund holdings does not contain trades, and trading discreteness cannot be directly computed. Instead, we construct its binary analogue—variable *PositionUnchanged*—which equals one if a fund has not changed its share position in a stock since the previous quarter, and zero otherwise. We sometimes refer to this measure as "position stickiness". Even though direct intra-quarter trades cannot be observed, we deem it unlikely that a fund would trade a stock in such a way that the number of shares held at the beginning and the end of the quarter is exactly the same. Accordingly, we assume that if the share position is the same at the beginning and the end of the quarter, then the fund did not trade the stock during the quarter, and the position can be assumed sticky.

Panel A of Table 11 reports summary statistics on position stickiness by the quintiles of stock size formed at the beginning of each year. As expected, stickiness monotonically decreases with stock size: on average, funds do not change 54% (23%) of their positions in the smallest (largest) stocks. Panel B shows statistics for each Morningstar fund equity style—Large Cap, Mid Cap, and Small Cap. Two conclusions

[30] In Table B8 of Appendix B we show that this effect is robust to controlling for other stock characteristics and their interactions with fund size. These controls do not materially affect the coefficient on *StockSize×FundSize*, which retains both statistical and economic significance in all specifications.

emerge from this panel. First, it is clear that the effect of stock size on position stickiness cannot be explained by the mandated fund style—we see a monotonic pattern in both the Large Cap and the Small Cap style. Second, consistent with our prediction, the effect of stock size on stickiness is more strongly pronounced in Large Cap than in Small Cap funds: for Large (Small) Cap funds, the difference in average stickiness between the smallest and largest stocks is 43% (27%).

[Insert Table 11 here]

We formally investigate the interplay between fund and stock size by re-estimating regression (10) with the following modifications. First, we use *PositionUnchanged* as the dependent variable and estimate a linear OLS regression. Second, we directly measure fund size with *FundTNA*—the natural logarithm of the total net assets (TNA) under management of fund *j* at the beginning of year-quarter *t*. To facilitate the interpretation of economic effects, we also consider the binary transformation of TNA—variable *BigFund*, which equals one if the TNA of fund *j* is above the sample median in the observation quarter, and zero otherwise. In select specifications, we control for fund age, expense ratio, and risk. *FundAge* is the number of years since the fund inception, *ExpenseRatio* is the most recent end-of-fiscal-year expense ratio of the fund, and *ReturnVolatility* is the standard deviation of the fund's monthly returns in the previous year. These data points are retrieved from the Morningstar U.S. Open-End Fund Database. Finally, we add fund style fixed effects in all regressions; the style is given either by the "Equity Style" (Morningstar Equity Style Box) or "Category" (Morningstar Investment Category) classification.

Table 12 reports the results. Similar to Table 10, all interaction coefficients are negative and reliably significant, indicating that the effect of stock size on position stickiness is greater in larger funds. The economic effects are more modest than for the trading discreteness measures, possibly due to the crudeness of the stickiness variable, which only identifies fully intact positions and does not differentiate between small and big adjustments.[31] The sensitivity of position stickiness to stock size increases by 47% (=-1.7937/-3.8084) in the above-median fund-TNA group relative to the below-median group.

[31] It is possible to quantify the degree of the change in position from quarter to quarter. However, in this case, it is hard to know whether the position was adjusted in one trade or gradually over time. Moreover, it is unclear how to

These results are consistent with the rational attention allocation equilibrium but are hard to explain with direct frictions, such as costs of rebalancing a position. Indeed, if rebalancing costs determined the frequency of position adjustments, then smaller funds would be relatively less keen to rebalance small stocks, because the rebalancing costs would be relatively more punishing, as they would constitute a bigger fraction of the fund's portfolio.

[Insert Table 12 here]

6. Conclusion

We present a simple model describing the allocation of investigative resources by rational investment managers and study the implications of this allocation for cross-sectional stock mispricing. Managers allocate more attention to stocks that promise greater dollar-valued arbitrage opportunities. This allocation mechanism produces a distinct pattern of mispricing, where smaller stocks are, on average, more mispriced in percentage terms. This mispricing is not eliminated by arbitrage action, since managers do not receive up-to-date signals on these stocks. We find strong empirical evidence consistent with the model's predictions. The magnitude of price corrections observed during corporate disclosures is an inverse monotonic function of stock size, and the effect of size is stronger than that of any other stock characteristic traditionally linked to price efficiency. This effect varies smoothly across the spectrum of size and is not specific to small or special stocks. Size is also the first-order driver of institutional attention. Stocks allocated less attention form sticky portfolio positions. Trading in these stocks is more discrete and the price discovery is slow. We also predict and confirm empirically that investors commanding less capital allocate their attention more evenly and are more likely to transact in smaller stocks. In our model, the preference for investment style—small vs. large cap stocks—arises endogenously as an outcome of rational resource allocation and capital constraints. Funds endowed with smaller portfolios are relatively more content to specialize in smaller stocks. In summary, we show that stock size is the first-order driver of mispricing,

standardize the position change, as different denominators produce measures with vastly different properties. In the interest of robustness and measurement sharpness, we opt out of this analysis.

more so than standard frictions, and that investors' preferences for different size classes can arise endogenously in the attention allocation problem.

References

Acharya, V. V., and L. H. Pedersen. 2005. Asset pricing with illiquidity risk. *Journal of Financial Economics* 77: 375-410.

Amihud, Y. 2002. Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets* 5:31-56.

Amihud, Y., and R. Goyenko. 2013. Mutual fund's R^2 as predictor of performance. *Review of Financial Studies* 26:667-694.

Andrei, D., and M. Hasler. 2015. Investor attention and stock market volatility. *Review of Financial Studies* 28:33-72.

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

Banz, R. W. 1981. The relationship between return and market value of common stocks. *Journal of Financial Economics* 9:3-18.

Barber, B. M., Odean, T., and N. Zhu. 2009. Do retail trades move markets? *Review of Financial Studies* 22:151-186.

Ben-Rephael, A., Carlin, B. I., Da, Z., and R. D. Israelsen. 2017. Demand for information and asset pricing. Working paper, Idiana University, University of Notre Dame, and Michigan State University.

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

Berk, J. B., and J. H. van Binsbergen. 2015. Measuring skill in the mutual fund industry. *Journal of Financial Economics* 118:1-20.

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

Bernard, V. L., and J. K. Thomas. 1989. Post-earnings-announcement drift: Delayed price response or risk premium? *Journal of Accounting Research* 27:1-36.

Cao, C., Simin, T., and J. Zhao. 2008. Can growth options explain the trend in idiosyncratic risk? *Review of Financial Studies* 21:2599-2633.

Chan, W. S. 2003. Stock price reaction to news and no-news: Drift and reversal after headlines. *Journal of Financial Economics* 70:223-260.

Chordia, T., Roll, R., and A. Subrahmanyam. 2008. Liquidity and market efficiency. *Journal of Financial Economics* 87:249-268.

Chuprinin, O., Gaspar, S., and M. Massa. 2018. Adjusting to the information environment: News tangibility and mutual fund performance. *Management Science* forthcoming.

Cohen, L., Frazzini, A., and C. Malloy. 2008. The small world of investing: Board connections and mutual fund returns. *Journal of Political Economy* 116:951-979.

Cremers, K. J. M., and A. Petaijisto. 2009. How active is your fund manager? A new measure that predicts performance. *Review of Financial Studies* 22:3329-3365.

Cronqvist, H., Siegel, A., and F. Yu. 2015. Value versus growth investing: Why do different investors have different styles? *Journal of Financial Economics* 117:333-349.

Cziraki, P., Mondira, J., and T. Wu. 2018. Asymmetric attention and stock returns. Working paper, University of Toronto and Verde Asset Management.

Da, Z., Engelberg, J., and P. Gao. 2011. In search of attention. *Journal of Finance* 66:1461-1499.

Dang, T. L., Moshirian, F., and B. Zhang. 2015. Commonality in news around the world. *Journal of Financial Economics* 116:82-110.

DellaVigna, S., and J. Pollet. 2009. Investor inattention and Friday earnings announcements. *Journal of Finance* 64:709-749.

De Long, J. B., Shleifer, A., Summers, L. H., and R. J. Waldman. 1990. Positive feedback investment strategies and destabilizing rational specualtion. *Journal of Finance* 45:379-395.

Desai, H., Ramesh., K., Thiagarajan, S. R., and B. V. Balachandran. 2002. An investigation of the informational role of short interest in the Nasdaq market. *Journal of Finance* 57:2263-2287.

Edelen, R. M., Ince, O. S., and G. B. Kadlec. 2016. Institutional investors and stock return anomalies. *Journal of Financial Economics* 119:472-488.

Engelberg, J. E., and C. A. Parsons. 2011. The causal impact of media in financial markets. *Journal of Finance* 66:67-97.

Engelberg, J. E., McLean, R. D., and J. Pontiff. 2018. Anomalies and news. *Journal of Finance* 73:1971-2001.

Engelberg, J. E., Reed, A. V., and M. C. Ringgenberg. 2012. How are shorts informed? Short sellers, news, and information processing. *Journal of Financial Economics* 105:260-278.

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

Fama, E. F., Fisher, L., Jensen, M. C., and R. Roll. 1969. The adjustment of stock prices to new information. *International Economic Review* 10:1-21.

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

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

Fama, E. F., and K. R. French. 2008. Dissecting anomalies. *Journal of Finance* 63:1653-1678.

Fang, L., and J. Peress. 2009. Media coverage and the cross-section of stock returns. *Journal of Finance* 64:2023-2052.

Foucault, T., Pagano, M., and A. Röell. 2013. *Market liquidity: Theory, evidence, and policy*. New York: Oxford University Press.

Froot, K. A., Scharfstein, A. D., and J. C. Stein. 1992. Herd on the street: Informational inefficiencies in a market with short-term speculation. *Journal of Finance* 47:1461-1484.

Gao, P., Parsons, C. A., and J. Shen. 2017. Global relation between financial distress and equity returns. *Review of Financial Studies* 31:239-277.

Gao, X., and J. R. Ritter. 2010. The marketing of seasoned equity offerings. *Journal of Financial Economics* 97:33-52.

Grossman, S. J., and J. E. Stiglitz. 1980. On the impossibility of informationally efficient markets. *American Economic Review* 70:393-408.

Hameed, A., Lof, M., and M. Suominen. 2017. Slow trading and stock return predictability. Working paper, National University of Singapore and Aalto University.

Han, Y., and D. Lesmond. 2011. Liquidity biases and the pricing of cross-sectional idiosyncratic volatility. *Review of Financial Studies* 24:1590-1629.

Harvey, C. R., Liu, Y., and H. Zhu. 2016. ...and the cross-section of expected returns. *Review of Financial Studies* 29:5-68.

Hirshleifer, D., Lim, S., and S. T. Teoh. 2009. Driven to distraction: Extraneous events and underreaction to earnings news. *Journal of Finance* 64:2289-2325.

Hong, H., Lim, T., and J. C. Stein. 2000. Bad news travels slowly: Size, analyst coverage, and the cross-section of momentum strategies. *Journal of Finance* 55:265-295.

Hong, H., Kubik, J. D., and J. C. Stein. 2005. Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers. *Journal of Finance* 60:2801-2824.

Hu, G., Jo, K. M., Wang, Y. A., and J. Xie. 2018. Institutional trading and Abel Noser data. *Journal of Corporate Finance* forthcoming.

Jiang, G. J., Xu, D., and T. Yao. 2009. The information content of idiosyncratic volatility. *Journal of Financial and Quantitative Analysis* 44:1-28.

Jin, L., and S. C. Myers. 2006. R^2 around the world: New theory and new tests. *Journal of Financial Economics* 79:257-292.

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

Karolyi, G. A., Lee, K.-H., and M. A. van Dijk. 2012. Understanding commonality in liquidity around the world. *Journal of Financial Economics* 105:82-112.

Kelly, P. J. 2014. Information efficiency and firm-specific return variation. *Quarterly Journal of Finance* 4:1-44.

Kim, O., and R. E. Verrecchia. 1991. Trading volume and price reactions to public announcements. *Journal of Accounting Research* 29:302-321.

Kolasinski, A. C., Reed, A. V., and M. T. Ringgenberg. 2013. A multiple lender approach to understanding supply and search in the equity lending market. *Journal of Finance* 68:559-595.

Kumar, A., and C. M. C. Lee. 2006. Retail investor sentiment and return comovements. *Journal of Finance* 61:2451-2486.

Kyle, A. S., Obizhaeva, A. A., and T. Tuzun. 2012. Trading game invariance in the TAQ dataset. Working paper, University of Maryland and Board of Governors of the Federal Reserve System.

Lee, B.-S. 1998. Permanent, temporary, and non-fundamental components of stock prices. *Journal of Financial and Quantitative Analysis* 33:1-32.

Lee, Y.-J. 2012. The effect of quarterly report readability on information efficiency of stock prices. *Contemporary Accounting Research* 29:1137-1170.

Lee, W. L., and M. H. Liu. 2011. Does more information in stock price lead to greater or smaller idiosyncratic return volatility? *Journal of Banking and Finance* 35:1563-1580.

Lee, C. M. C., and B. Radhakrishna. 2000. Inferring investor behavior: Evidence from TORQ data. *Journal of Financial Markets* 3:83-111.

Lehman, B. N. 1990. Fads, martingales, and market efficiency. *Quarterly Journal of Economics* 1:1-28.

Li, F. 2008. Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics* 45:221-247.

Li, B., Rajgopal, S., and M. Venkatachalam. 2014. R^2 and idiosyncratic risk are not interchangeable. *Accounting Review* 89:2261-2295.

Ljungqvist, A., and W. Qian. 2016. How constraining are limits to arbitrage? *Review of Financial Studies* 29:1975-2028.

Llorente, G., Michaely, R., Saar, G., and J. Wang. 2002. Dynamic volume-return relation of individual stocks. *Review of Financial Studies* 15:1005-1047.

Loughran, T., and B. McDonald. 2017. The use of EDGAR filings by investors. *Journal of Behavioral Finance* 18:231-248.

Miller, B. P. 2010. The effects of reporting complexity on small and large investor trading. *Accounting Review* 85:2107-2143.

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

Morck, R., Yeung, B., and W. Yu. 2013. R^2 and the economy. *Annual Review of Financial Economics* 5:143-166.

Niessner, M. 2015. Strategic disclosure timing and insider trading. Working paper, Yale University.

Peng, L. 2005. Learning with information capacity constraints. *Journal of Financial and Quantitative Analysis* 40:307-329.

Peress, J., and D. Schmidt. 2019. Glued to the TV: Distracted noise traders and stock market liquidity. *Journal of Finance* forthcoming.

Puckett, A., and X. Yan. 2011. The interim trading skills of institutional investors. *Journal of Finance* 66:601-633.

Rajgopal, S., and M. Venkatachalam. 2011. Financial reporting quality and idiosyncratic return volatility. *Journal of Accounting and Economics* 51:1-20.

Saffi, P. A. C., and K. Sigurdsson. 2011. Price efficiency and short selling. *Review of Financial Studies* 24:821-852.

Shleifer, A., and R. W. Vishny. 1997. The limits of arbitrage. *Journal of Finance* 52:35-55.

Stambaugh, R. F., and Y. Yuan. 2017. Mispricing factors. *Review of Financial Studies* 30:1270-1315.

Tetlock, P. C. 2007. Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance* 62:1139-1168.

Tetlock, P. C. 2010. Does public financial news resolve asymmetric information? *Review of Financial Studies* 23:3520-3557.

Ungeheuer, M. 2017. Stock returns and the cross-section of investor attention. Working paper, Aalto University.

Xu, Y., and B. G. Malkiel. 2003. Investigating the behavior of idiosyncratic volatility. *Journal of Business* 76:613-645.

Yu, J., Zhang, B., and L. Zhang. 2017. Who captures the power of the pen? *Review of Financial Studies* 31:43-96.

Table 1
Variables of Interest, Summary Statistics

This table reports summary statistics and the time range of availability for the main stock characteristics used in the analysis. The sample consists of all CRSP stocks with share codes 10 and 11, excluding penny and microcap stocks (see Section 2 for details). The unit of observation is stock-calendar quarter. *MCap* is the market capitalization of the stock at the beginning of a quarter; *AmIlliq* is the average daily Amihud (2002) illiquidity measure, calculated as the absolute daily return divided by the dollar trading volume on that day (top 1% of *AmIlliq* values are excluded); *Spread* is the average daily relative spread, calculated as the difference between bid and ask prices divided by the midpoint price; *BorrowingFee* is the average daily indicative borrowing fee for the stock in a quarter; *Lendable* is the number of shares of the stock available for borrowing divided by the number of shares outstanding at the beginning of a quarter; *RetailOwn* equals one minus the percentage of institutional ownership in the stock at the beginning of a quarter; *RetailTrading* is the average daily ratio of the number of small trades to the total number of trades in the stock in a quarter, where a small trade is defined as any trade up to $5,000 in value (see Lee and Radhakrishna (2000) and Barber, Odean, and Zhu (2009)). *MediaCover* is the average daily number of articles about the firm from RavenPack with the relevance score of 100; *Tangibility* is the measure of information tangibility calculated as the ratio of the number of numeric characters to the total number of characters in all Factiva articles about the firm in a quarter (see Chuprinin, Gaspar, and Massa (2018)); *FogIndex* (*10KLength*) is the Fog readability index (the number of words) of (in) the most recent 10-K report of the firm available at the beginning of a quarter (see Li (2008)). The last column reports the correlation coefficients of the variables of interest with the natural logarithm of the market capitalization (in $thousands).

Variable Name	Variable Description	Sample Period	Mean	P10	P25	Median	P75	P90	Correlation with *Size*
MCap	Market capitalization, $Thousands	1990-2016	2,469,789	27,175	64,836	225,470	956,012	3,645,589	
AmIlliq	Amihud (2002) illiquidity ($\times 10^6$)	1990-2016	1.510	0.001	0.003	0.044	0.579	3.533	-0.34
Spread	Ask-bid / midpoint price, %	1990-2016	2.22	0.08	0.26	1.23	3.03	5.51	-0.55
BorrowingFee	Stock borrowing fee, %	2002-2016	1.54	0.37	0.38	0.44	0.57	2.83	-0.18
Lendable	Lendable shares / shares outstanding, %	2002-2016	14.96	0.28	2.64	12.91	25.24	32.41	0.41
RetailOwn	1 – fraction of inst. ownership, %	1990-2016	59.60	16.27	35.75	64.22	85.94	95.44	-0.63
RetailTrading	Proportion of small trades from TAQ, %	1993-2000	40.42	15.17	24.70	37.83	54.43	72.75	-0.70
MediaCover	Number of articles (RavenPack)	2000-2016	0.87	0.00	0.10	0.35	0.90	1.83	0.16
Tangibility	News tangibility (Factiva), %	1998-2008	3.46	2.04	2.43	3.02	3.95	5.49	-0.24
FogIndex	Fog readability index of 10-K report	1995-2012	19.39	17.83	18.56	19.39	20.33	21.40	0.04
10KLength	10-K report length, number of words	1995-2012	31,173	10,153	16,352	24,948	38,086	58,298	0.24

Table 2

The Effect of Size and Size-Related Stock Characteristics on the Price Correction at Earnings Announcements

This table reports the results of the panel regressions of the stock price correction at earnings announcements on stock size and other explanatory variables. The dependent variable is *Abs Three-Day Return*$_{i,t}$—the absolute value of the market-adjusted cumulative stock return over the first three trading days (including the announcement day) following the earnings announcement of firm *i* in quarter *t*. We run the regressions separately for positive (Panels A1 and B1) and negative (Panels A2 and B2) surprises, i.e. events for which the three-day return is positive and negative, respectively. Panels A1 and A2 show the analysis for a single explanatory variable, while Panels B1 and B2 show the analysis where different competing explanatory variables are added to stock size. The main independent variable in each regression is *Size*$_{i,t}$, defined as the natural logarithm of the market capitalization (in $thousands) of firm *i* 32 trading days before its earnings announcement in quarter *t*. Other explanatory variables are defined as follows. *AmIlliq* is the average daily Amihud (2002) illiquidity measure of the stock in the period [-32,-3] relative to the earnings announcement ("event"), where day 0 is the day of the event (we scale all coefficients on *AmIlliq* by 10^3). *Spread* is the average daily relative spread of the stock (expressed in decimals) in the period [-32,-3] relative to the event. *BorrowingFee* is the average daily indicative borrowing fee (expressed in decimals) for the stock in the period [-32,-3] relative to the event and *Lendable* is the number of shares of the stock available for borrowing in the securities lending market divided by the number of shares outstanding 32 trading days before the event (expressed in decimals). *RetailOwn* equals one minus the fraction of institutional ownership in the stock at the end of the quarter preceding the event quarter (expressed in decimals) and *RetailTrading* is the average daily ratio of the number of small (not exceeding $5,000 in value) trades to the total number of trades in the period [-32,-3] relative to the event (expressed in decimals). *MediaCover* is the average daily number of relevant (with the relevance score of 100) articles about the firm from RavenPack in the period [-32,-3] relative to the event. *Tangibility* is the measure of information tangibility (expressed in decimals), equal to the average ratio of the number of numeric characters to the total number of characters in all Factiva articles about the firm in the quarter preceding the event. *FogIndex* is the Fog readability index for the firm's most recent 10-K report and *10KLength* is the natural logarithm of the number of words in the most recent 10-K report available 32 trading days before the event. The sample periods of different regressions coincide with the periods of availability of the relevant explanatory variables reported in Table 1. The inclusion of time (year-quarter) fixed effects is indicated at the bottom of each panel. *T*-statistics are reported in parentheses and are derived from standard errors clustered by stock. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Panel A1: Positive Surprises, Single Explanatory Variable

Regression Specification: $Abs\ Three\text{-}Day\ Return_{i,t} = \alpha + \beta ExplanatoryVariable_{i,t} + \delta_t + \varepsilon_{i,t}$

		Illiquidity		Arbitrage Constraints		Retail Participation		Information Environment/Transparency			
Expl. Var.	*Size*	*AmIlliq*	*Spread*	*BorrowingFee*	*Lendable*	*RetailOwn*	*RetailTrading*	*MediaCover*	*Tangibility*	*FogIndex*	*10KLength*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Coefficient	-0.0066***	2.2303***	0.3408***	0.1287***	-0.0175***	0.0040***	0.0479***	-0.0011***	-0.1261**	0.0003**	-0.0022***
(*t*-stat)	(-31.838)	(10.294)	(11.928)	(8.232)	(-3.433)	(2.637)	(15.447)	(-3.136)	(-2.551)	(2.449)	(-4.797)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs	119,865	119,061	116,222	63,859	64,292	119,085	28,354	80,949	14,892	65,036	64,573
Adjusted R^2	0.051	0.033	0.034	0.027	0.025	0.030	0.058	0.026	0.050	0.039	0.040

Panel A2: Negative Surprises, Single Explanatory Variable

Regression Specification: $Abs\ Three\text{-}Day\ Return_{i,t} = \alpha + \beta ExplanatoryVariable_{i,t} + \delta_t + \varepsilon_{i,t}$

		Illiquidity		Arbitrage Constraints		Retail Participation		Information Environment/Transparency			
Expl. Var.	*Size*	*AmIlliq*	*Spread*	*BorrowingFee*	*Lendable*	*RetailOwn*	*RetailTrading*	*MediaCover*	*Tangibility*	*FogIndex*	*10KLength*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Coefficient	-0.0055***	0.6775***	0.2141***	0.1135***	-0.0126***	0.0030**	0.0330***	-0.0010***	-0.2155***	0.0005***	-0.0006
(*t*-stat)	(-34.212)	(4.495)	(13.376)	(10.780)	(-3.203)	(2.230)	(13.770)	(-3.475)	(-4.876)	(3.425)	(-1.504)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs	119,124	118,427	115,746	63,400	63,679	118,374	28,877	79,550	12,674	64,483	64,034
Adjusted R^2	0.056	0.036	0.036	0.038	0.033	0.035	0.043	0.033	0.059	0.040	0.040

Table 2
Continued

Panel B1: Positive Surprises, Size plus a Competing Variable

Regression Specification: Abs Three-Day Return$_{i,t}$ = $\alpha + \beta Size_{i,t} + \gamma CompetingVariable_{i,t} + \delta_t + \varepsilon_{i,t}$

	Illiquidity		Arbitrage Constraints		Retail Participation		Information Environment/Transparency			
Competing Variable	*AmIlliq*	*Spread*	*BorrowingFee*	*Lendable*	*RetailOwn*	*RetailTrading*	*MediaCover*	*Tangibility*	*FogIndex*	*10KLength*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Coefficient *Size*	-0.0067***	-0.0064***	-0.0073***	-0.0081***	-0.0079***	-0.0035***	-0.0070***	-0.0082***	-0.0062***	-0.0062***
(*t*-stat)	(-30.944)	(-28.864)	(-24.670)	(-28.925)	(-35.133)	(-8.922)	(-25.634)	(-13.546)	(-24.738)	(-24.191)
Coefficient Compet. Var.	-0.0578	0.0164	0.0758***	0.0291***	-0.0208***	0.0260***	0.0007***	-0.3238***	0.0005***	0.0001
(*t*-stat)	(-0.269)	(0.652)	(4.845)	(5.845)	(-13.475)	(6.396)	(3.329)	(-6.239)	(3.711)	(0.173)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs	118,678	115,825	63,632	64,292	118,690	28,275	80,673	14,869	65,036	64,573
Adjusted R^2	0.052	0.049	0.050	0.050	0.054	0.061	0.045	0.074	0.059	0.059

Panel B2: Negative Surprises, Size plus a Competing Variable

Regression Specification: Abs Three-Day Return$_{i,t}$ = $\alpha + \beta Size_{i,t} + \gamma CompetingVariable_{i,t} + \delta_t + \varepsilon_{i,t}$

	Illiquidity		Arbitrage Constraints		Retail Participation		Information Environment/Transparency			
Competing Variable	*AmIlliq*	*Spread*	*BorrowingFee*	*Lendable*	*RetailOwn*	*RetailTrading*	*MediaCover*	*Tangibility*	*FogIndex*	*10KLength*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Coefficient *Size*	-0.0060***	-0.0058***	-0.0062***	-0.0069***	-0.0066***	-0.0036***	-0.0057***	-0.0046***	-0.0050***	-0.0051***
(*t*-stat)	(-35.128)	(-31.566)	(-29.515)	(-32.242)	(-36.186)	(-9.800)	(-28.332)	(-8.947)	(-24.526)	(-24.608)
Coefficient Compet. Var.	-1.3090***	-0.0798***	0.0734***	0.0301***	-0.0181***	0.0105***	0.0004***	-0.3380***	0.0006***	0.0014***
(*t*-stat)	(-7.983)	(-4.509)	(7.416)	(7.241)	(-12.845)	(3.093)	(3.331)	(-7.239)	(4.489)	(3.677)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs	118,004	115,310	63,170	63,679	117,938	28,785	79,257	12,657	64,483	64,034
Adjusted R^2	0.057	0.054	0.064	0.062	0.059	0.048	0.052	0.068	0.057	0.057

Table 3

Price Correction at Earnings Announcements, Single Sort by Stock Characteristics

This table shows how the stock price correction at earnings announcements varies across the quintiles of different stock characteristics (described in Table 2). Each quarter we assign a stock to one of the five groups (quintiles) according to the value of the sorting variable prior to the earnings announcement. Q1 (Q5) indicates the quintile with the lowest (highest) values. The table reports the mean of *Abs Three-Day Return* (defined in Table 2) in each quintile as well as the difference in means of *Abs Three-Day Return*, with the corresponding *t*-statistic, between the highest and the lowest quintile. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively. Panel A (Panel B) shows the analysis for positive (negative) surprises. The sample period for each sorting variable is reported in Table 1.

Panel A: Positive Surprises

Abs Three-Day Return by Quintiles of Stock Characteristics

		Illiquidity		Arbitrage Constraints		Retail Participation		Information Environment/Transparency			
Sorting Variable	*Size* (1)	*AmIlliq* (2)	*Spread* (3)	*BorrowingFee* (4)	*Lendable* (5)	*RetailOwn* (6)	*RetailTrading* (7)	*MediaCover* (8)	*Tangibility* (9)	*FogIndex* (10)	*10KLength* (11)
Q1	0.0751	0.0468	0.0509	0.0569	0.0694	0.0629	0.0496	0.0649	0.0601	0.0599	0.0651
Q2	0.0695	0.0589	0.0577	0.0586	0.0639	0.0607	0.0508	0.0710	0.0582	0.0632	0.0662
Q3	0.0642	0.0654	0.0634	0.0614	0.0612	0.0610	0.0547	0.0672	0.0561	0.0652	0.0655
Q4	0.0570	0.0698	0.0691	0.0645	0.0607	0.0617	0.0591	0.0671	0.0549	0.0659	0.0628
Q5	0.0452	0.0703	0.0735	0.0783	0.0644	0.0655	0.0719	0.0576	0.0548	0.0649	0.0595
Q5-Q1	-0.0299***	0.0235***	0.0227***	0.0214***	-0.0049***	0.0025***	0.0223***	-0.0073***	-0.0053***	0.0050***	-0.0057***
(*t*-stat)	(-38.021)	(30.573)	(27.626)	(16.901)	(-4.191)	(3.205)	(15.637)	(-7.409)	(-3.288)	(5.573)	(-6.294)

Panel B: Negative Surprises

Abs Three-Day Return by Quintiles of Stock Characteristics

		Illiquidity		Arbitrage Constraints		Retail Participation		Information Environment/Transparency			
Sorting Variable	*Size* (1)	*AmIlliq* (2)	*Spread* (3)	*BorrowingFee* (4)	*Lendable* (5)	*RetailOwn* (6)	*RetailTrading* (7)	*MediaCover* (8)	*Tangibility* (9)	*FogIndex* (10)	*10KLength* (11)
Q1	0.0681	0.0451	0.0491	0.0540	0.0655	0.0610	0.0473	0.0620	0.0540	0.0562	0.0591
Q2	0.0672	0.0584	0.0566	0.0554	0.0629	0.0574	0.0502	0.0682	0.0526	0.0597	0.0629
Q3	0.0636	0.0645	0.0624	0.0598	0.0579	0.0577	0.0535	0.0663	0.0511	0.0615	0.0623
Q4	0.0555	0.0671	0.0669	0.0629	0.0577	0.0600	0.0558	0.0633	0.0496	0.0632	0.0603
Q5	0.0433	0.0627	0.0661	0.0760	0.0621	0.0623	0.0639	0.0562	0.0456	0.0618	0.0577
Q5-Q1	-0.0248***	0.0177***	0.0171***	0.0220***	-0.0034***	0.0013**	0.0166***	-0.0058***	-0.0084***	0.0056***	-0.0014*
(*t*-stat)	(-44.738)	(32.520)	(29.019)	(24.477)	(-4.074)	(2.139)	(14.712)	(-7.921)	(-5.649)	(7.035)	(-1.791)

Table 4

Price Correction at Earnings Announcements, Sequential Double Sort

Panel A of this table shows the effect of stock size on the earnings announcement price reaction conditional on sorting by different stock characteristics (described in Table 2). Each quarter we assign a stock to one of the five groups (quintiles) according to the value of a given characteristic prior to the earnings announcement. Q1 (Q5) indicates the quintile with the lowest (highest) values. Then, within each of these quintiles, we further sort stocks into five quintiles of size. The table reports the difference in means of *Abs Three-Day Return* (defined as in Table 2), with the corresponding *t*-statistic, between the highest and the lowest size quintile. Panel A1 (Panel A2) shows the results for positive (negative) surprises. Panel B reports the output of the reverse procedure: first, we sort stocks by size, and then, within each size quintile, we further sort stocks by a given characteristic. Panel B1 (Panel B2) shows the analysis for positive (negative) surprises. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively. The sample period for each sorting variable is reported in Table 1.

Panel A1: Positive Surprises, Pre-Sort by a Competing Variable, Main Sort by Size

Q5-Q1 Difference in Means of Abs Three-Day Return

Pre-sort by Compet. Var.	Main sort by Size	Illiquidity		Arbitrage Constraints		Retail Participation		Information Environment/Transparency			
		AmIlliq	*Spread*	*BorrowingFee*	*Lendable*	*RetailOwn*	*RetailTrading*	*MediaCover*	*Tangibility*	*FogIndex*	*10KLength*
Quintile		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	Q5-Q1	-0.0203***	-0.0207***	-0.0358***	-0.0216***	-0.0226***	-0.0245***	-0.0246***	-0.0333***	-0.0250***	-0.0258***
	(*t*-stat)	(-19.606)	(-18.464)	(-17.652)	(-4.830)	(-16.609)	(-10.623)	(-6.796)	(-9.244)	(-12.739)	(-12.174)
2	Q5-Q1	-0.0204***	-0.0214***	-0.0300***	-0.0387***	-0.0314***	-0.0118***	-0.0237***	-0.0309***	-0.0256***	-0.0235***
	(*t*-stat)	(-15.787)	(-17.243)	(-19.257)	(-18.655)	(-23.672)	(-5.250)	(-11.832)	(-8.841)	(-12.256)	(-10.450)
3	Q5-Q1	-0.0224***	-0.0165***	-0.0278***	-0.0414***	-0.0369***	-0.0118***	-0.0319***	-0.0270***	-0.0289***	-0.0266***
	(*t*-stat)	(-15.963)	(-12.546)	(-15.546)	(-22.133)	(-25.002)	(-4.736)	(-17.058)	(-7.934)	(-12.436)	(-11.853)
4	Q5-Q1	-0.0239***	-0.0147***	-0.0285***	-0.0400***	-0.0377***	-0.0101***	-0.0352***	-0.0258***	-0.0303***	-0.0301***
	(*t*-stat)	(-15.081)	(-9.232)	(-13.662)	(-21.322)	(-24.811)	(-3.543)	(-19.154)	(-8.008)	(-13.857)	(-14.327)
5	Q5-Q1	-0.0287***	-0.0231***	-0.0184***	-0.0329***	-0.0310***	-0.0253***	-0.0384***	-0.0183***	-0.0327***	-0.0322***
	(*t*-stat)	(-10.277)	(-7.801)	(-3.894)	(-18.057)	(-11.211)	(-5.871)	(-22.676)	(-5.193)	(-13.856)	(-13.776)

Panel A2: Negative Surprises, Pre-Sort by a Competing Variable, Main Sort by Size

Q5-Q1 Difference in Means of Abs Three-Day Return

Pre-sort by Compet. Var.	Main sort by Size	Illiquidity		Arbitrage Constraints		Retail Participation		Information Environment/Transparency			
		AmIlliq	*Spread*	*BorrowingFee*	*Lendable*	*RetailOwn*	*RetailTrading*	*MediaCover*	*Tangibility*	*FogIndex*	*10KLength*
Quintile		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	Q5-Q1	-0.0199***	-0.0208***	-0.0295***	-0.0107***	-0.0237***	-0.0236***	-0.0150***	-0.0213***	-0.0198***	-0.0186***
	(*t*-stat)	(-18.748)	(-18.373)	(-16.832)	(-5.667)	(-17.805)	(-10.390)	(-9.015)	(-6.205)	(-12.476)	(-11.186)
2	Q5-Q1	-0.0251***	-0.0205***	-0.0274***	-0.0297***	-0.0262***	-0.0125***	-0.0179***	-0.0182***	-0.0188***	-0.0179***
	(*t*-stat)	(-19.068)	(-16.476)	(-20.558)	(-16.685)	(-21.483)	(-5.448)	(-10.690)	(-5.359)	(-10.715)	(-10.365)
3	Q5-Q1	-0.0256***	-0.0180***	-0.0271***	-0.0410***	-0.0303***	-0.0093***	-0.0291***	-0.0124***	-0.0209***	-0.0198***
	(*t*-stat)	(-18.428)	(-13.838)	(-14.192)	(-24.329)	(-24.391)	(-3.996)	(-17.172)	(-3.727)	(-12.143)	(-10.891)
4	Q5-Q1	-0.0230***	-0.0135***	-0.0187***	-0.0349***	-0.0283***	-0.0060**	-0.0274***	-0.0110***	-0.0216***	-0.0245***
	(*t*-stat)	(-15.797)	(-9.372)	(-10.906)	(-21.121)	(-22.573)	(-2.514)	(-16.873)	(-3.261)	(-12.258)	(-14.114)
5	Q5-Q1	-0.0194***	-0.0133***	-0.0068***	-0.0299***	-0.0189***	-0.0154***	-0.0354***	-0.0069**	-0.0279***	-0.0312***
	(*t*-stat)	(-13.764)	(-8.937)	(-3.123)	(-16.995)	(-13.985)	(-5.971)	(-24.083)	(-2.193)	(-15.585)	(-18.276)

Table 4
Continued

Panel B1: Positive Surprises, Pre-Sort by Size, Main Sort by a Competing Variable
Q5-Q1 Difference in Means of Abs Three-Day Return

Pre-sort by *Size* Quintile	Main sort by Compet. Var.	Illiquidity		Arbitrage Constraints		Retail Participation		Information Environment/Transparency			
		AmIlliq	*Spread*	*BorrowingFee*	*Lendable*	*RetailOwn*	*RetailTrading*	*MediaCover*	*Tangibility*	*FogIndex*	*10KLength*
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	Q5-Q1	-0.0061**	0.0093***	0.0073***	0.0223***	-0.0102***	0.0391***	0.0190***	-0.0193***	0.0079***	0.0078***
	(*t*-stat)	(-2.211)	(3.227)	(2.807)	(8.907)	(-5.130)	(9.440)	(8.010)	(-4.478)	(2.863)	(2.818)
2	Q5-Q1	-0.0158***	-0.0030*	0.0108***	0.0094***	-0.0109***	0.0099***	0.0166***	-0.0132***	0.0031	0.0010
	(*t*-stat)	(-10.435)	(-1.928)	(4.219)	(4.351)	(-6.821)	(3.200)	(8.192)	(-3.939)	(1.389)	(0.473)
3	Q5-Q1	-0.0116***	0.0007	0.0150***	-0.0025	-0.0091***	0.0019	0.0082***	-0.0131***	0.0053***	-0.0024
	(*t*-stat)	(-8.797)	(0.523)	(7.295)	(-1.356)	(-6.732)	(0.668)	(4.826)	(-3.814)	(2.882)	(-1.306)
4	Q5-Q1	-0.0047***	-0.0012	0.0111***	0.0014	-0.0145***	-0.0038	0.0076***	-0.0143***	0.0067***	0.0004
	(*t*-stat)	(-3.840)	(-0.896)	(6.242)	(0.930)	(-11.964)	(-1.330)	(4.862)	(-3.382)	(3.846)	(0.259)
5	Q5-Q1	0.0080***	0.0052***	0.0033**	0.0068***	-0.0127***	0.0095***	-0.0020	-0.0056**	0.0046***	-0.0013
	(*t*-stat)	(7.878)	(4.670)	(2.490)	(5.460)	(-12.843)	(5.093)	(-1.448)	(-2.183)	(3.111)	(-0.868)

Panel B2: Negative Surprises, Pre-Sort by Size, Main Sort by a Competing Variable
Q5-Q1 Difference in Means of Abs Three-Day Return

Pre-sort by *Size* Quintile	Main sort by Compet. Var.	Illiquidity		Arbitrage Constraints		Retail Participation		Information Environment/Transparency			
		AmIlliq	*Spread*	*BorrowingFee*	*Lendable*	*RetailOwn*	*RetailTrading*	*MediaCover*	*Tangibility*	*FogIndex*	*10KLength*
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	Q5-Q1	-0.0172***	-0.0015	0.0103***	0.0174***	-0.0046***	0.0239***	0.0137***	-0.0180***	0.0086***	0.0052***
	(*t*-stat)	(-11.769)	(-0.996)	(4.909)	(8.550)	(-3.191)	(9.192)	(6.956)	(-4.887)	(4.377)	(2.607)
2	Q5-Q1	-0.0228***	-0.0077***	0.0123***	0.0077***	-0.0093***	0.0045*	0.0109***	-0.0150***	0.0059***	0.0087***
	(*t*-stat)	(-15.908)	(-5.329)	(5.243)	(3.753)	(-6.556)	(1.710)	(5.739)	(-4.569)	(3.024)	(4.535)
3	Q5-Q1	-0.0152***	-0.0003	0.0160***	-0.0028	-0.0093***	-0.0031	0.0083***	-0.0129***	0.0060***	0.0021
	(*t*-stat)	(-11.078)	(-0.203)	(7.308)	(-1.489)	(-6.690)	(-1.153)	(4.704)	(-3.712)	(3.315)	(1.134)
4	Q5-Q1	-0.0082***	-0.0025*	0.0124***	0.0018	-0.0128***	-0.0068**	0.0060***	-0.0086**	0.0047***	0.0031*
	(*t*-stat)	(-6.580)	(-1.931)	(6.913)	(1.088)	(-9.965)	(-2.516)	(3.613)	(-2.329)	(2.823)	(1.807)
5	Q5-Q1	0.0073***	0.0051***	0.0055***	0.0059***	-0.0099***	0.0075***	-0.0023*	-0.0067**	0.0021	0.0000
	(*t*-stat)	(7.346)	(4.747)	(4.193)	(4.692)	(-9.404)	(3.768)	(-1.752)	(-2.328)	(1.517)	(0.012)

Table 5
Size vs. Other Stock Characteristics as Drivers of Institutional Attention

This table reports the results of quarterly panel regressions of institutional attention on stock size and other stock characteristics. The main dependent variable is *InstAtten*, defined as the average daily news search score for the firm in the Bloomberg terminal in a quarter (see Ben-Rephael, Da, and Israelsen (2017) for a description of the data). Panel A reports the regressions of *InstAtten* on a single stock characteristic, while Panel B reports the regressions of *InstAtten* on stock size plus other (competing) stock characteristics. The independent variables are defined as follows. *Size* is the natural logarithm of the market capitalization (in $thousands) of the firm at the beginning of a quarter. *AmIlliq* is the average daily Amihud (2002) illiquidity measure of the stock in a quarter (we scale all coefficients on *AmIlliq* by 10^3) and *Spread* is the average daily relative spread of the stock (expressed in decimals) in a quarter. *BorrowingFee* is the average daily indicative borrowing fee (expressed in decimals) for the stock in a quarter and *Lendable* is the number of shares of the stock available for borrowing in the securities lending market divided by the number of shares outstanding at the beginning of a quarter (expressed in decimals). *RetailOwn* equals one minus the fraction of institutional ownership in the stock at the beginning of a quarter (expressed in decimals). *MediaCover* is the average daily number of relevant (with the relevance score of 100) articles about the firm from RavenPack in a quarter. *FogIndex* is the Fog readability index for the firm's most recent 10-K report available at the beginning of a quarter. *EarnVolatility* is the standard deviation of the dollar earnings per share reported by the firm in the twelve quarters preceding the current quarter. *BM* is the firm's most recent book-to-market ratio available at the beginning of a quarter. *Leverage* is the firm's most recent ratio of total debt to total book value of assets (expressed in decimals) available at the beginning of a quarter. *R&D* is the firm's most recent ratio of research and development expenses to gross sales (expressed in decimals) available at the beginning of a quarter (we scale all coefficients on *R&D* by 10^3). The detailed description of these variables are provided in Section 3. The sample period in columns (1)-(7) and (9)-(12) of Panel A as well as in columns (1)-(6) and (8)-(11) of Panel B is from February 2010 to December 2016. The sample period in column (8) of Panel A and column (7) of Panel B is from February 2010 to December 2012. The inclusion of time (year-quarter) fixed effects is indicated at the bottom of each panel. *T*-statistics are reported in parentheses and are derived from standard errors clustered by stock. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Panel A: Single Explanatory Variable

Regression Specification: $InstAtten_{i,t} = \alpha + \beta ExplanatoryVariable_{i,t} + \delta_t + \varepsilon_{i,t}$

		Illiquidity		Arbitrage Constraints		Retail Participation	Information Environment/Transparency		Firm Characteristics			
Expl. Var.	*Size*	*AmIlliq*	*Spread*	*BorrowingFee*	*Lendable*	*RetailOwn*	*MediaCover*	*FogIndex*	*EarnVolatility*	*BM*	*Leverage*	*R&D*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Coefficient	0.2154***	-0.0212***	-21.7043***	-0.1542	0.9218***	-0.5764***	0.0396***	0.0108***	0.2101***	-0.0826***	0.3298***	-0.0435
(*t*-stat)	(57.654)	(-11.807)	(-13.733)	(-1.272)	(13.130)	(-20.383)	(4.474)	(6.409)	(5.792)	(-8.030)	(7.451)	(-0.268)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs	53,959	53,878	53,936	50,481	49,600	53,167	53,837	13,638	37,623	48,235	45,318	26,408
Adjusted R^2	0.521	0.042	0.095	0.037	0.069	0.101	0.158	0.044	0.068	0.055	0.060	0.044

Panel B: Size plus a Competing Variable

Regression Specification: $InstAtten_{i,t} = \alpha + \beta Size_{i,t} + \gamma CompetingVariable_{i,t} + \delta_t + \varepsilon_{i,t}$

		Illiquidity		Arbitrage Constraints		Retail Participation	Information Environment/Transparency		Firm Characteristics			
Competing Variable		*AmIlliq*	*Spread*	*BorrowingFee*	*Lendable*	*RetailOwn*	*MediaCover*	*FogIndex*	*EarnVolatility*	*BM*	*Leverage*	*R&D*
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Coefficient	*Size*	0.2175***	0.2298***	0.2229***	0.2304***	0.2208***	0.2023***	0.2210***	0.2364***	0.2218***	0.2182***	0.2189***
(*t*-stat)		(57.774)	(55.573)	(58.757)	(54.864)	(51.845)	(44.342)	(42.342)	(54.382)	(53.802)	(51.798)	(40.469)
Coefficient	Compet. Var.	0.0099***	8.9380***	1.2646***	-0.4991***	0.0904***	0.0144***	0.0016	0.1164***	0.0174**	0.1781***	0.5118***
(*t*-stat)		(8.649)	(10.204)	(11.820)	(-7.846)	(3.500)	(4.206)	(1.442)	(4.332)	(2.561)	(6.283)	(2.969)
Time FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs		53,878	53,936	50,481	49,600	53,167	53,837	13,638	37,623	48,235	45,318	26,408
Adjusted R^2		0.523	0.529	0.548	0.550	0.524	0.538	0.490	0.549	0.536	0.527	0.532

Table 6

The Effect of Institutional Attention vs. Other Stock Characteristics

This table reports the results of the panel regressions of the stock price correction at earnings announcements on institutional attention and other explanatory variables. The dependent variable is *Abs Three-Day Return*, defined as in Table 2. Panel A (Panel B) shows the analysis for positive (negative) surprises. The main independent variable is *InstAtten*, defined as the average daily news search score for the firm in the Bloomberg terminal in the period [-32,-3] days relative to the earnings announcement event. The competing explanatory variables include *AmIlliq*, *Spread*, *BorrowingFee*, *Lendable*, *RetailOwn*, *MediaCover*, *FogIndex* (defined as in Table 2), and *GoogleSearch*—a measure of retail investors' attention, equal to the most recent monthly index of Google searches for the firm available 32 trading days before the event (see Da, Engelberg, and Gao (2011) for details). The sample period in columns (1)-(7) is from February 2010 to December 2016 and the sample period in column (8) is from February 2010 to December 2012. The inclusion of time (year-quarter) fixed effects is indicated at the bottom of each panel. *T*-statistics are reported in parentheses and are derived from standard errors clustered by stock. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Panel A: Positive Surprises, Attention vs. Other Stock Characteristics

Regression Specification: $Abs\ Three\text{-}Day\ Return_{i,t} = \alpha + \beta InstAtten_{i,t} + \gamma CompetingVariable_{i,t} + \delta_t + \varepsilon_{i,t}$

		Illiquidity		Arbitrage Constraints		Retail Participation		Information	
Competing Variable		*AmIlliq*	*Spread*	*BorrowingFee*	*Lendable*	*RetailOwn*	*GoogleSearch*	*MediaCover*	*FogIndex*
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Coefficient	*InstAtten*	-0.0095***	-0.0085***	-0.0091***	-0.0088***	-0.0093***	-0.0076***	-0.0091***	-0.0078***
(*t*-stat)		(-10.695)	(-9.791)	(-10.249)	(-10.098)	(-10.693)	(-7.191)	(-10.465)	(-6.887)
Coefficient	Compet. Var.	2.0371	0.5859**	0.1150***	-0.0239***	0.0022	-0.0013***	-0.0002	-0.0003**
(*t*-stat)		(0.842)	(2.237)	(6.613)	(-3.396)	(0.711)	(-2.857)	(-1.311)	(-2.005)
Time FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs		31,801	32,113	32,066	31,927	31,673	26,452	32,113	7,561
Adjusted R^2		0.010	0.011	0.013	0.010	0.009	0.008	0.010	0.013

Panel B: Negative Surprises, Attention vs. Other Stock Characteristics

Regression Specification: $Abs\ Three\text{-}Day\ Return_{i,t} = \alpha + \beta InstAtten_{i,t} + \gamma CompetingVariable_{i,t} + \delta_t + \varepsilon_{i,t}$

		Illiquidity		Arbitrage Constraints		Retail Participation		Information	
Competing Variable		*AmIlliq*	*Spread*	*BorrowingFee*	*Lendable*	*RetailOwn*	*GoogleSearch*	*MediaCover*	*FogIndex*
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Coefficient	*InstAtten*	-0.0089***	-0.0087***	-0.0082***	-0.0083***	-0.0090***	-0.0074***	-0.0081***	-0.0076***
(*t*-stat)		(-10.746)	(-10.388)	(-10.161)	(-9.930)	(-10.728)	(-7.267)	(-9.329)	(-7.216)
Coefficient	Compet. Var.	-1.5872	-0.0037	0.0982***	-0.0142***	-0.0035	-0.0001	-0.0003	0.0000
(*t*-stat)		(-0.742)	(-0.046)	(8.897)	(-2.921)	(-1.488)	(-0.378)	(-1.564)	(0.068)
Time FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs		31,662	31,968	31,918	31,760	31,510	26,284	31,968	7,792
Adjusted R^2		0.021	0.021	0.028	0.021	0.021	0.019	0.021	0.018

Table 7
ANcerno Institutional Trades and Trading Discreteness

This table reports descriptive statistics for the accounts (funds) in the ANcerno database. Panel A reports quarterly statistics for the fund characteristics, Panel B reports statistics for the fund-level trading discreteness measures across stock-quarters, and Panel C reports quarterly statistics for the trading discreteness measures aggregated to the stock level. *NumActiveFunds* is the number of ANcerno funds that make at least one trade during a quarter. *NumStocks* is the number of unique stocks traded by the fund in a given quarter. *NumTrades* is the number of trades executed by the fund in a quarter. *TradingVolume* is the overall dollar trading volume of the fund in a quarter. *StocksBought* (*StocksSold*) is the overall dollar volume of stocks bought (sold) by the fund in a quarter. *TimeBetweenTrades* is the average time (in days) between trades of the fund in a given stock in a quarter and *DistanceBetweenLots1* is the average interval between unique trading lots of the fund for a given stock in a quarter. To calculate the latter variable, we first sort all the unique trading lots (measured in the number of shares traded) for the fund-stock in a quarter in an ascending sequence; then we calculate the difference between the neighboring lots, take the average of all such differences, and standardize it by the number of shares outstanding at the end of the quarter. In Panel C, *TimeBetweenTrades* and *DistanceBetweenLots1* are aggregated to the stock level as the value-weighted averages of the corresponding fund-stock measures, where the weights are proportional to the trading volume of the fund in the stock in a quarter. *DistanceBetweenLots2* measures the average interval between unique trading lots across funds and is computed as follows: first, we sort all the unique trading lots in a given stock in a quarter across all funds in an ascending sequence; then we calculate the difference between the neighboring lots, take the average of all such differences, and standardize it by the number of shares outstanding at the end of the quarter. *LotConcentration* measures the concentration of the unique trading lots across funds and is computed as the Herfindahl index of the proportions of the unique trading volumes. The proportion of each trading volume is computed as the number of trading lots with this volume in the stock in a quarter divided by the total number of trades executed by all funds in this stock in that quarter. The detailed description of all ANcerno-based variables are provided in Section 4. The sample period for *TimeBetweenTrades* and *DistanceBetweenLots1* in Panel B and Panel C is from January 1999 to December 2010. The sample period for *DistanceBetweenLots2* and *LotConcentration* in Panel C is from January 1999 to June 2013.

Panel A: Fund Characteristics						
Variable	Mean	P10	P25	Median	P75	P90
NumActiveFunds	150	82	109	155	199	224
NumStocks	420	64	133	260	497	963
NumTrades	14,002	219	602	1,438	3,374	18,588
TradingVolume ($Thousands)	1,445,290	15,456	42,963	133,878	513,251	2,673,340
StocksBought ($Thousands)	485,329	6,239	16,324	49,636	188,888	933,366
StocksSold ($Thousands)	487,263	6,672	18,222	56,085	204,012	1,000,680
Panel B: Trading Discreteness, Fund-Level						
Variable	Mean	P10	P25	Median	P75	P90
TimeBetweenTrades (Days)	9.47	0.37	1.47	4.71	12.00	24.67
DistanceBetweenLots1 (%)	0.0058	0.0001	0.0003	0.0011	0.0040	0.0120
Panel C: Trading Discreteness Aggregated at the Stock Level						
Variable	Mean	P10	P25	Median	P75	P90
TimeBetweenTrades (Days)	4.71	1.00	1.79	3.06	5.30	9.22
DistanceBetweenLots1 (%)	0.0179	0.0014	0.0030	0.0066	0.0152	0.0347
DistanceBetweenLots2 (%)	0.0083	0.0003	0.0007	0.0019	0.0054	0.0150
LotConcentration (%)	12.07	0.99	1.99	4.05	10.19	33.33

Table 8
Size and Attention as Predictors of Trading Discreteness

This table reports the results of quarterly panel regressions of trading discreteness on stock size and institutional attention. The trading discreteness measures are computed as in Panel C of Table 7; however, we divide *TimeBetweenTrades* by 100 and multiply *DistanceBetweenLots1* and *DistanceBetweenLots2* by 100 before running the regressions. The independent variables include *Size*, *InstAtten*, and *EDGAR*, defined as in Table 5. The sample periods are as follows: from January 1999 to December 2010 in columns (1) and (3), from January 2003 to December 2010 in columns (2) and (4), from January 1999 to June 2013 in columns (5) and (8), from January 2003 to June 2013 in columns (6) and (9), and from February 2010 to June 2013 in columns (7) and (10). The inclusion of time (year-quarter) fixed effects is indicated at the bottom of the table. *T*-statistics are reported in parentheses and are derived from standard errors clustered by stock. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Regression Specification: Trading Discreteness$_{i,t}$ = $\alpha + \beta$ExplanatoryVariable$_{i,t}$ + $\delta_t + \varepsilon_{i,t}$										
	Dep. Var.: *TimeBetweenTrades*		Dep. Var.: *DistanceBetweenLots1*		Dep. Var.: *DistanceBetweenLots2*			Dep. Var.: *LotConcentration*		
Expl. Var.	*Size*	*EDGAR*	*Size*	*EDGAR*	*Size*	*InstAtten*	*EDGAR*	*Size*	*InstAtten*	*EDGAR*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Coefficient	-0.0073***	-0.0061***	-0.0062***	-0.0031***	-0.0047***	-0.0022***	-0.0023***	-0.0542***	-0.0301***	-0.0332***
(*t*-stat)	(-59.975)	(-21.874)	(-47.207)	(-13.460)	(-39.865)	(-13.328)	(-15.754)	(-59.038)	(-12.715)	(-27.777)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs	169,160	106,889	164,660	104,903	194,761	21,163	132,603	206,033	21,350	138,285
Adjusted R^2	0.065	0.028	0.047	0.024	0.045	0.010	0.008	0.222	0.036	0.050

Table 9

Decomposition of the Trading Discreteness into the Size/Attention-Driven and the Residual Components

This table reports the results of panel regressions of *Abs Three-Day Return*, defined as in Table 2, on two orthogonal components of trading discreteness. The first component is trading discreteness predicted by stock size or institutional attention. It is equal to the predicted values from the regression of a given trading discreteness measure on stock size or institutional attention. The other component consists of the residuals from this regression. The analysis is run for the four measures of trading discreteness—*TimeBetweenTrades*, *DistanceBetweenLots1*, *DistanceBetweenLots2*, and *LotConcentration*, defined as in Panel C of Table 7 and calculated for the quarter preceding the event quarter. The results are reported in Panel A, Panel B, Panel C, and Panel D, respectively. The predictive variables are *Size*, *EDGAR*, and *InstAtten*, defined as in Tables 5 and 6. The sample periods are as follows: from January 1999 to December 2010 in columns (1)-(2) and (5)-(6) of Panel A and Panel B, from January 2003 to December 2010 in columns (3)-(4) and (7)-(8) of Panel A and Panel B, from January 1999 to June 2013 in columns (1)-(2) and (7)-(8) of Panel C and Panel D, from January 2003 to June 2013 in columns (3)-(4) and (9)-(10) of Panel C and Panel D, and from February 2010 to June 2013 in columns (5)-(6) and (11)-(12) of Panel C and Panel D. The inclusion of time (year-quarter) fixed effects is indicated at the bottom of each panel. *T*-statistics are reported in parentheses and are derived from standard errors clustered by stock. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Panel A: Decomposed Variable - Average Time between Trades within Fund

Regression Specifications: Abs Three-Day Return$_{i,t} = \alpha + \beta$*Predicted(TimeBetweenTrades)*$_{i,t} + \delta_t + \varepsilon_{i,t}$ *and Abs Three-Day Return*$_{i,t} = \alpha + \beta$*Residual(TimeBetweenTrades)*$_{i,t} + \delta_t + \varepsilon_{i,t}$

	Positive Surprises				*Negative Surprises*			
Predictor	*Size*		*EDGAR*		*Size*		*EDGAR*	
Components	Predicted (1)	Residual (2)	Predicted (3)	Residual (4)	Predicted (5)	Residual (6)	Predicted (7)	Residual (8)
Coefficient	0.8378***	-0.0667***	0.4268**	-0.0128	0.7143***	-0.0517***	0.1035	0.0040
(*t*-stat)	(23.145)	(-8.960)	(2.422)	(-1.533)	(22.863)	(-8.471)	(0.669)	(0.525)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs	53,073	53,073	35,005	35,005	51,780	51,780	34,784	34,784
Adjusted R^2	0.054	0.039	0.046	0.046	0.054	0.040	0.046	0.046

Panel B: Decomposed Variable - Average Distance between Unique Trading Lots within Fund

Regression Specifications: Abs Three-Day Return$_{i,t} = \alpha + \beta$*Predicted(DistanceBetweenLots1)*$_{i,t} + \delta_t + \varepsilon_{i,t}$ *and Abs Three-Day Return*$_{i,t} = \alpha + \beta$*Residual(DistanceBetweenLots1)*$_{i,t} + \delta_t + \varepsilon_{i,t}$

	Positive Surprises				*Negative Surprises*			
Predictor	*Size*		*EDGAR*		*Size*		*EDGAR*	
Components	Predicted (1)	Residual (2)	Predicted (3)	Residual (4)	Predicted (5)	Residual (6)	Predicted (7)	Residual (8)
Coefficient	1.0096***	-0.0052	0.3709**	0.0234**	0.8677***	-0.0067	0.0990	0.0302***
(*t*-stat)	(23.170)	(-0.943)	(2.413)	(2.104)	(23.030)	(-1.376)	(0.733)	(2.729)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs	52,731	52,731	34,903	34,903	51,431	51,431	34,659	34,659
Adjusted R^2	0.055	0.037	0.046	0.046	0.054	0.039	0.046	0.047

Table 9
Continued

Panel C: Decomposed Variable - Average Distance between Unique Trading Lots across Funds

Regression Specifications: $Abs\ Three\text{-}Day\ Return_{i,t} = \alpha + \beta Predicted(DistanceBetweenLots2)_{i,t} + \delta_t + \varepsilon_{i,t}$ *and* $Abs\ Three\text{-}Day\ Return_{i,t} = \alpha + \beta Residual(DistanceBetweenLots2)_{i,t} + \delta_t + \varepsilon_{i,t}$

	Positive Surprises						*Negative Surprises*					
Predictor	*Size*		*InstAtten*		*EDGAR*		*Size*		*InstAtten*		*EDGAR*	
Components	Predicted (1)	Residual (2)	Predicted (3)	Residual (4)	Predicted (5)	Residual (6)	Predicted (7)	Residual (8)	Predicted (9)	Residual (10)	Predicted (11)	Residual (12)
Coefficient	1.2720***	-0.0698***	3.2734***	0.5738**	0.8856***	0.0452*	1.0980***	-0.0434**	1.1899**	0.0131	0.1611	0.0583***
(*t*-stat)	(24.671)	(-3.601)	(5.651)	(2.286)	(3.266)	(1.878)	(24.197)	(-2.021)	(2.110)	(0.294)	(0.657)	(3.470)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs	62,095	62,095	8,639	8,639	43,843	43,843	60,935	60,935	8,320	8,320	43,733	43,733
Adjusted R^2	0.053	0.036	0.008	0.006	0.040	0.040	0.052	0.037	0.013	0.012	0.040	0.040

Panel D: Decomposed Variable - Herfindahl Concentration Index Based on Frequency of Unique Trading Lots across Funds

Regression Specifications: $Abs\ Three\text{-}Day\ Return_{i,t} = \alpha + \beta Predicted(LotConcentration)_{i,t} + \delta_t + \varepsilon_{i,t}$ *and* $Abs\ Three\text{-}Day\ Return_{i,t} = \alpha + \beta Residual(LotConcentration)_{i,t} + \delta_t + \varepsilon_{i,t}$

	Positive Surprises						*Negative Surprises*					
Predictor	*Size*		*InstAtten*		*EDGAR*		*Size*		*InstAtten*		*EDGAR*	
Components	Predicted (1)	Residual (2)	Predicted (3)	Residual (4)	Predicted (5)	Residual (6)	Predicted (7)	Residual (8)	Predicted (9)	Residual (10)	Predicted (11)	Residual (12)
Coefficient	0.1099***	-0.0474***	0.2383***	-0.0027	0.0670***	-0.0024	0.0932***	-0.0470***	0.0832**	-0.0278***	0.0088	-0.0048*
(*t*-stat)	(24.479)	(-15.442)	(5.696)	(-0.171)	(3.272)	(-0.655)	(23.718)	(-19.367)	(2.051)	(-4.616)	(0.474)	(-1.705)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs	63,005	63,005	8,667	8,667	44,186	44,186	61,887	61,887	8,346	8,346	44,147	44,147
Adjusted R^2	0.052	0.042	0.008	0.003	0.040	0.039	0.051	0.045	0.013	0.014	0.040	0.040

Table 10

Sensitivity of Trading Discreteness to Profit Potential by Fund Size

This table reports the results of quarterly panel regressions of trading discreteness on stock size and its interactions with fund size proxies. In columns (1) and (4), the unit of observation is fund group-stock-quarter. Every year we assign each fund to one of the two groups according to the value of *BigFund*—a dummy variable equal to one if the dollar trading volume of the fund in a year (*FundVolume*) is above the median in the sample for that year. The trading discreteness measures are *TimeBetweenTrades* and *DistanceBetweenLots1* computed for each group as described in Panel C of Table 7. In columns (2)-(3) and (5)-(6), the unit of observation is fund-stock-quarter and the trading discreteness measures—*TimeBetweenTrades* and *DistanceBetweenLots1*—are fund-stock-specific, as described in Panel B of Table 7. *StockSize* is the natural logarithm of the market capitalization (in $thousands) of the firm at the beginning of a quarter. We divide *TimeBetweenTrades* by 100 and multiply *DistanceBetweenLots1* by 100 before running the regressions; the coefficients on *StockSize* and its interactions are further scaled by 100. The sample period for this analysis is from January 1999 to December 2010. The inclusion of time (year-quarter) fixed effects is indicated at the bottom of the table. *T*-statistics are reported in parentheses and are derived from standard errors clustered by stock in columns (1) and (4) and double-clustered by fund and stock in columns (2)-(3) and (5)-(6). *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	Trading Discreteness Regressed on Stock and Fund Sizes Interacted					
	Dep. Var.: *TimeBetweenTrades*			Dep. Var.: *DistanceBetweenLots1*		
	Fund Group Regressions	Fund Regressions		Fund Group Regressions	Fund Regressions	
Expl. Var.	(1)	(2)	(3)	(4)	(5)	(6)
StockSize	0.8214***	7.2592***	1.2332***	-0.2354***	0.5997**	-0.1142***
	(35.860)	(7.480)	(12.217)	(-37.983)	(2.348)	(-12.504)
FundVolume		0.0383***			0.0064***	
		(5.049)			(3.277)	
StockSize ×		-0.3193***			-0.0394***	
		(-6.747)			(-3.187)	
BigFund	0.1717***		0.1311***	0.0619***		0.0331***
	(46.398)		(4.401)	(33.329)		(6.825)
StockSize × *BigFund*	-1.5766***		-1.0247***	-0.3854***		-0.1986***
	(-60.347)		(-5.251)	(-29.998)		(-6.602)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
NObs	265,586	2,336,761	2,336,761	258,625	2,240,722	2,240,722
Adjusted R^2	0.103	0.036	0.018	0.059	0.051	0.050

Table 11
Sticky Positions in Mutual Fund Portfolios

This table reports descriptive statistics for the measure of position stickiness in mutual fund portfolios, *PositionUnchanged*, equal to one if a fund has not changed its share holdings in the stock since the previous quarter, and zero otherwise. We sort stocks into five groups (quintiles) according to their market capitalization at the beginning of the year and report fund-stock-quarter statistics for each such quintile. Panel A reports statistics within the quintiles and Panel B reports statistics within the quintiles separately for the three fund style groups of the Morningstar equity style classification. The sample period is from 1990 to 2016.

Panel A: Statistics by Size Quintile

		Market Capitalization ($Thousands)		*PositionUnchanged*	
Quintile	NObs	Mean	Median	Mean	Median
1	704,711	75,763	70,135	0.54	1.00
2	1,585,610	220,931	204,718	0.40	0.00
3	2,736,788	536,080	482,610	0.32	0.00
4	4,346,681	1,438,264	1,270,836	0.26	0.00
5	9,394,471	26,852,553	9,308,152	0.23	0.00

Panel B: Statistics by Size Quintile within Fund Style

			Market Capitalization ($Thousands)		*PositionUnchanged*	
Fund Style	Quintile	NObs	Mean	Median	Mean	Median
	1	183,353	74,714	68,818	0.66	1.00
	2	322,891	220,799	202,670	0.55	1.00
Large	3	542,373	537,441	479,681	0.46	0.00
	4	1,234,658	1,513,702	1,338,070	0.35	0.00
	5	7,328,782	31,945,530	12,326,582	0.23	0.00
	1	64,594	73,015	69,261	0.55	1.00
	2	124,479	207,241	188,546	0.41	0.00
Mid	3	243,486	513,348	459,428	0.31	0.00
	4	725,938	1,559,206	1,390,063	0.23	0.00
	5	1,437,872	10,023,414	4,973,786	0.21	0.00
	1	456,764	76,572	70,682	0.49	0.00
	2	1,138,240	222,466	207,077	0.36	0.00
Small	3	1,950,929	538,539	487,315	0.28	0.00
	4	2,386,085	1,362,434	1,207,161	0.22	0.00
	5	627,817	5,943,317	3,130,280	0.22	0.00

Table 12

Sensitivity of Position Stickiness to Profit Potential by Fund Size

This table reports the results of linear-probability quarterly panel regressions of the mutual fund position stickiness on stock size and its interactions with fund size proxies. The dependent variable is *PositionUnchanged*, defined as in Table 11. The main independent variables are *StockSize*, defined as in Table 10, *FundTNA*—the natural logarithm of the total net assets of the fund at the beginning of a quarter, and *BigFund*—a dummy variable equal to one if the fund's TNA is higher than the median TNA in the sample in that quarter. Other independent variables include *FundAge*—the number of years since the fund first appeared in the sample, *ExpenseRatio*—the most recent end-of-fiscal-year expense ratio of the fund available in a quarter, and *ReturnVolatility*—the standard deviation of the fund's monthly returns in the previous year. The coefficients on *StockSize* and its interactions are scaled by 100. The sample period for this analysis is from January 1990 to December 2016. The inclusion of time (year-quarter), fund style, and fund investment category fixed effects is indicated at the bottom of the table. *T*-statistics are reported in parentheses and are derived from standard errors double-clustered by fund and stock. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	PositionUnchanged Regressed on Stock and Fund Sizes Interacted							
Expl. Var.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
StockSize	4.6515	4.7060	4.8853	4.6649	-3.8084***	-3.8619***	-3.8270***	-3.7843***
	(1.365)	(1.308)	(1.487)	(1.363)	(-6.217)	(-6.486)	(-6.161)	(-6.372)
FundTNA	0.0602**	0.0533*	0.0647**	0.0543*				
	(2.111)	(1.786)	(2.339)	(1.896)				
StockSize	-0.4828***	-0.4929***	-0.5025***	-0.4932***				
	(-2.854)	(-2.769)	(-3.064)	(-2.886)				
BigFund					0.2306***	0.2265***	0.2661***	0.2550***
					(2.601)	(2.686)	(3.068)	(3.089)
StockSize					-1.7937***	-1.8545***	-1.9969***	-2.0263***
					(-3.302)	(-3.530)	(-3.759)	(-3.949)
FundAge		0.0016		0.0014		0.0002		0.0001
		(1.620)		(1.403)		(0.247)		(0.133)
ExpenseR		-6.7510***		-7.2861***		-5.2891***		-5.9147***
		(-4.172)		(-4.825)		(-3.281)		(-3.938)
ReturnVol		-1.6760***		-1.6473***		-1.7381***		-1.6959***
		(-4.944)		(-5.162)		(-5.179)		(-5.351)
FE	Style +	Style +	Category +	Category +	Style +	Style +	Category +	Category +
NObs	18,742,862	14,669,441	19,200,190	15,006,074	18,742,862	14,669,441	19,200,190	15,006,074
Adjusted	0.052	0.059	0.046	0.055	0.051	0.057	0.045	0.053

Appendix A. Proof of propositions and properties

Proposition 1 The fund manager's optimal attention allocation (τ_L^*, τ_S^*) is given by

$$(\tau_L^*, \tau_S^*) = \left(\frac{\sqrt{k}c + \left(\sqrt{k} - 1\right)\tau_0}{\sqrt{k} + 1}, \frac{c - \left(\sqrt{k} - 1\right)\tau_0}{\sqrt{k} + 1}\right),$$

where $k = \frac{K_L}{K_S}$ is the relative size of firms.

Proof . We first solve the optimization problem in (3) without considering the positivity of τ_j and then verify that the solution meets this condition. The first-order condition of the optimization problem in (3) is

$$K_L \frac{\tau_0}{(\tau_0 + \tau_L)^2} - K_S \frac{\tau_0}{(\tau_0 + c - \tau_L)^2} = 0. \quad \text{(A1)}$$

Note that the second-order sufficient condition holds since

$$-2K_L \frac{\tau_0}{(\tau_0 + \tau_L)^3} - 2K_S \frac{\tau_0}{(\tau_0 + c - \tau_L)^3} < 0. \quad \text{(A2)}$$

The condition in (A1) can be rewritten as

$$k = \frac{K_L}{K_S} = \left(\frac{\tau_0 + \tau_L}{\tau_0 + c - \tau_L}\right)^2. \quad \text{(A3)}$$

Thus, the optimal attention allocation is given by $\tau_L^* = \frac{\sqrt{k}c + \left(\sqrt{k} - 1\right)\tau_0}{\sqrt{k} + 1}$ and $\tau_S^* = c - \tau_L^* = \frac{c - \left(\sqrt{k} - 1\right)\tau_0}{\sqrt{k} + 1}$. Finally, we check whether the optimal attention allocation is positive. Since $\sqrt{k} > 1$ and c and τ_0 are positive, the attention to the larger stock τ_L^* is positive for all parameters. The attention to the smaller stock τ_S^* is positive if and only if $c > \left(\sqrt{k} - 1\right)\tau_0$. We impose this condition as a parametric assumption in our model. We note that, if $c \leq \left(\sqrt{k} - 1\right)\tau_0$, the attention allocation to the smaller stock, τ_S^*, converges to zero, which corresponds to an infinite variance, $\frac{1}{\tau_S^*}$, of signal r_S.

Property 1.1 More attention is allocated to the larger stock: $\tau_L^* > \tau_S^*$.

Proof . From Proposition 1, $(\tau_L^*, \tau_S^*) = \left(\frac{\sqrt{k}c+(\sqrt{k}-1)\tau_0}{\sqrt{k}+1}, \frac{c-(\sqrt{k}-1)\tau_0}{\sqrt{k}+1}\right)$. Since $\sqrt{k} = \sqrt{\frac{K_L}{K_S}} > 1$ and $\tau_0 > 0$, it is straightforward that $\tau_L^* > \tau_S^*$.

Property 1.2 For a given aggregate level of attention c, the attention allocated to the larger stock increases with its relative size and the prior precision: $\frac{\partial \tau_L^*}{\partial k} > 0$ and $\frac{\partial \tau_L^*}{\partial \tau_0} > 0$.

Proof . By dividing the numerator and the denominator by $\sqrt{k}$, we can rewrite τ_L^* as

$$\tau_L^* = \frac{c + \left(1 - \frac{1}{\sqrt{k}}\right)\tau_0}{1 + \frac{1}{\sqrt{k}}}, \quad \text{(A4)}$$

which shows that τ_L^* increases in k, because the numerator increases and the denominator decreases in k. Finally, it is straightforward from (A4) that $\frac{\partial \tau_L^*}{\partial \tau_0} > 0$ since $k > 1$ and hence $1 - \frac{1}{\sqrt{k}} > 0$.

Proposition 2 At the optimal attention levels, the larger stock exhibits a lower expected percentage mispricing, i.e. $\psi_L < \psi_S$.

Proof . The numeration of ψ_j can be written as

$$E\left(\left|R_j - \frac{\xi_j^*}{2} r_j\right|\right) = \int_{-\infty}^{\frac{2R_j}{\xi_j^*}} \left(R_j - \frac{\xi_j^*}{2} r_j\right) f(r_j) dr_j + \int_{\frac{2R_j}{\xi_j^*}}^{\infty} \left(\frac{\xi_j^*}{2} r_j - R_j\right) f(r_j) dr_j, \quad \text{(A5)}$$

where is $f(r_j) = \frac{\sqrt{\tau_0 \xi_j^*}}{\sqrt{2\pi}} e^{\frac{-\tau_0 \xi_j^* (r_j - R_j)^2}{2}}$ is the probability density function of r_j. Integrating by parts,

$$\int r_j f(r_j) dr_j = \frac{\sqrt{\tau_0 \xi_j^*}}{\sqrt{2\pi}} e^{\frac{-\tau_0 \xi_j^* (r_j - R_j)^2}{2}} \left(-\frac{1}{\tau_0 \xi_j^*}\right) + \int R_j f(r_j) dr_j + C$$

$$= f(r_j)\left(-\frac{1}{\tau_0 \xi_j^*}\right) + \int R_j f(r_j) dr_j + C, \quad \text{(A6)}$$

where C is a free constant. Therefore,

$$\int_{-\infty}^{\frac{2R_j}{\xi_j^*}} \left(R_j - \frac{\xi_j^*}{2} r_j\right) f(r_j) dr_j = R_j F\left(\frac{2R_j}{\xi_j}\right) - \frac{\xi_j^*}{2}\left[-f\left(\frac{2R_j}{\xi_j}\right)\frac{1}{\tau_0 \xi_j^*} + R_j F\left(\frac{2R_j}{\xi_j}\right)\right] \quad \text{(A7)}$$

and

$$\int_{\frac{2R_j}{\xi_j^*}}^{\infty} \left(\frac{\xi_j^*}{2} r_j - R_j\right) f(r_j) dr_j = \frac{\xi_j^*}{2}\left[f\left(\frac{2R_j}{\xi_j}\right)\frac{1}{\tau_0 \xi_j^*} + R_j\left(1 - F\left(\frac{2R_j}{\xi_j}\right)\right)\right] - R_j\left(1 - F\left(\frac{2R_j}{\xi_j}\right)\right). \quad \text{(A8)}$$

We can thus rewrite $E\left(\left|R_j - \frac{\xi_j^*}{2} r_j\right|\right)$ as

$$E\left(\left|R_j - \frac{\xi_j^*}{2} r_j\right|\right) = R_j\left(2F\left(\frac{2R_j}{\xi_j^*}\right) - 1\right)\left(1 - \frac{\xi_j^*}{2}\right) + f\left(\frac{2R_j}{\xi_j^*}\right)\frac{1}{\tau_0}. \quad \text{(A9)}$$

Differentiating the above with respect to ξ_j^*, we obtain

$$-2R_j\left(1 - \frac{\xi_j^*}{2}\right) f\left(\frac{2R_j}{\xi_j^*}\right)\left(\frac{2R_j}{{\xi_j^*}^2}\right) - \frac{R_j}{2}\left(2F\left(\frac{2R_j}{\xi_j^*}\right) - 1\right) + \left(\xi_j^*\left(\frac{2R_j}{\xi_j^*} - R_j\right)\right) f\left(\frac{2R_j}{\xi_j^*}\right)\left(\frac{2R_j}{{\xi_j^*}^2}\right)$$

$$= -2R_j\left(1 - \frac{\xi_j^*}{2}\right) f\left(\frac{2R_j}{\xi_j^*}\right)\left(\frac{2R_j}{{\xi_j^*}^2}\right) - \frac{R_j}{2}\left(2F\left(\frac{2R_j}{\xi_j^*}\right) - 1\right) + 2R_j\left(1 - \frac{\xi_j^*}{2}\right) f\left(\frac{2R_j}{\xi_j^*}\right)\left(\frac{2R_j}{{\xi_j^*}^2}\right)$$

$$= \frac{R_j}{2}\left(1 - 2F\left(\frac{2R_j}{\xi_j^*}\right)\right). \quad \text{(A10)}$$

To verify the sign of $\frac{R_j}{2}\left(1 - 2F\left(\frac{2R_j}{\xi_j^*}\right)\right)$, we consider three cases: (i) $R_j > 0$, (ii) $R_j < 0$, and (iii) $R_j = 0$.

First, consider $R_j > 0$. Then, $\frac{2R_j}{\xi_j^*} > R_j$ since $\xi_j^* \in (0,1)$. Thus, $1 - 2F\left(\frac{2R_j}{\xi_j^*}\right) < 0$, since $F\left(\frac{2R_j}{\xi_j^*}\right) > F(R_j) = \frac{1}{2}$. Next, consider $R_j < 0$. Then, $\frac{2R_j}{\xi_j^*} < R_j$ and thus $1 - 2F\left(\frac{2R_j}{\xi_j^*}\right) > 0$. Note that in both cases $\frac{R_j}{2}\left(1 - 2F\left(\frac{2R_j}{\xi_j^*}\right)\right) < 0$. Finally, if $R_j = 0$, $\frac{R_j}{2}\left(1 - 2F\left(\frac{2R_j}{\xi_j^*}\right)\right) = 0$.

Since $\frac{R_j}{2}\left(1-2F\left(\frac{2R_j}{\xi_j^*}\right)\right)<0$ for all $R_j \neq 0$, $E\left(\left|R_j-\frac{\xi_j^*}{2}r_j\right|\right)$ and $\psi_j=\frac{E\left(\left|R_j-\frac{\xi_j^*}{2}r_j\right|\right)}{|1+R_j|}$ decreases in $\xi_\mathrm{j}^* \in (0,1)$ for all $R_j \neq 0$. Furthermore, since ξ_j^* increases in the optimal attention allocation τ_j^*, the expected mispricing ψ_j also decreases in τ_j^*.

Finally, by the properties of the optimal attention allocation, we have $\tau_L^* > \tau_S^*$, which implies that $\psi_L < \psi_S$ for all $R_j \neq 0$ at the optimal allocation levels.

Appendix B

Table B1

The Effect of Size and Size-Related Stock Characteristics on the Price Correction at Earnings Announcements, Additional Specifications

This table reports the additional specifications for the panel regression of the stock price correction at earnings announcements on stock size and other explanatory variables. The dependent variable is *Abs Three-Day Return*, defined as in Table 2. We run the regressions separately for positive (columns (1)-(5)) and negative (columns (6)-(10)) surprises, i.e. events for which the three-day return is positive and negative, respectively. The main independent variable is *Size* (defined as in Table 2); the competing explanatory variables include *AmIlliq*, *BorrowingFee*, *RetailOwn*, *RetailTrading*, *MediaCover*, and *FogIndex* (defined as in Table 2). The sample periods are as follows: from 1990 to 2016 in columns (1) and (6), from 1995 to 2012 in columns (2) and (7), from 2002 to 2016 in columns (3) and (8), from 2002 to 2012 in columns (4) and (9), and from 1995 to 2000 in columns (5) and (10). The inclusion of time (year-quarter) fixed effects is indicated at the bottom of the table. *T*-statistics are reported in parentheses and are derived from standard errors clustered by stock. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	Positive Surprises, Size plus Multiple Competing Variables					*Negative Surprises, Size plus Multiple Competing Variables*				
	Regression Specification: Abs Three-Day Return$_{i,t} = \alpha + \beta Size_{i,t} + \boldsymbol{\gamma CompetingVariables}_{i,t} + \delta_t + \varepsilon_{i,t}$									
Expl. Var.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Size	-0.0078***	-0.0075***	-0.0096***	-0.0093***	-0.0024***	-0.0069***	-0.0065***	-0.0083***	-0.0080***	-0.0026***
	(-33.905)	(-27.051)	(-29.140)	(-25.030)	(-4.425)	(-36.383)	(-27.458)	(-35.377)	(-26.762)	(-5.392)
AmIlliq	0.2785	0.6861	-2.7158***	-1.5832	0.0165	-1.0817***	-2.1186***	-3.5218***	-2.4109***	-1.2434**
	(1.282)	(1.131)	(-3.445)	(-1.406)	(0.018)	(-6.504)	(-6.237)	(-6.320)	(-2.953)	(-2.521)
BorrowingFee			0.1005***	0.1843***				0.0972***	0.1127***	
			(6.143)	(3.828)				(9.175)	(4.408)	
RetailOwn	-0.0206***	-0.0247***	-0.0266***	-0.0286***	-0.0231***	-0.0169***	-0.0189***	-0.0254***	-0.0252***	-0.0138***
	(-13.228)	(-12.274)	(-13.300)	(-11.527)	(-6.697)	(-11.835)	(-10.276)	(-14.419)	(-11.335)	(-4.378)
RetailTrading					0.0437***					0.0267***
					(8.089)					(6.067)
MediaCover			0.0012***	0.0011***				0.0008***	0.0008***	
			(4.006)	(3.021)				(5.459)	(3.304)	
FogIndex		0.0004***		0.0001	0.0019***		0.0006***		0.0002	0.0022***
		(3.334)		(0.396)	(4.336)		(4.471)		(1.216)	(5.611)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs	117,518	64,419	62,662	32,569	17,202	116,845	63,910	62,207	32,405	17,478
Adjusted R^2	0.055	0.064	0.057	0.081	0.072	0.060	0.062	0.075	0.079	0.061

Table B2

The Effect of Size and Size-Related Stock Characteristics on the Price Correction at Earnings Announcements, Thirty-Day Price Correction

This table reports the results of the panel regressions of the stock price correction at earnings announcements on stock size and other explanatory variables. The dependent variable is *Abs Thirty-Day Return*$_{i,t}$—the absolute value of the market-adjusted cumulative stock return over the first thirty trading days (including the announcement day) following the earnings announcement of firm *i* in quarter *t*. We run the regressions separately for positive (Panels A1, B1, and C1) and negative (Panels A2, B2, and C2) surprises, i.e. events for which the thirty-day return is positive and negative, respectively. Panels A1 and A2 show the analysis for a single explanatory variable; Panels B1 and B2 show the analysis where different competing explanatory variables are added to stock size; Panels C1 and C2 show the additional specifications where several competing explanatory variables are added to stock size. The main independent variable is *Size* (defined as in Table 2). The competing explanatory variables include *AmIlliq*, *Spread*, *BorrowingFee*, *Lendable*, *RetailOwn*, *RetailTrading*, *Tangibility*, *MediaCover*, *FogIndex*, and *10KLength* (defined as in Table 2). The sample periods of different regressions in Panels A1, A2, B1, and B2 coincide with the periods of availability of the relevant explanatory variables reported in Table 1. The sample periods in Panel C are as follows: from 1990 to 2016 in columns (1) and (6), from 1995 to 2012 in columns (2) and (7), from 2002 to 2016 in columns (3) and (8), from 2002 to 2012 in columns (4) and (9), and from 1995 to 2000 in columns (5) and (10). The inclusion of time (year-quarter) fixed effects is indicated at the bottom of the table. *T*-statistics are reported in parentheses and are derived from standard errors clustered by stock. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Panel A1: Positive Surprises, Single Explanatory Variable

Regression Specification: Abs Thirty-Day Return$_{i,t} = \alpha + \beta ExplanatoryVariable_{i,t} + \delta_t + \varepsilon_{i,t}$

		Illiquidity		Arbitrage Constraints		Retail Participation		Information Environment/Transparency			
Expl. Var.	*Size*	*AmIlliq*	*Spread*	*BorrowingFee*	*Lendable*	*RetailOwn*	*RetailTrading*	*MediaCover*	*Tangibility*	*FogIndex*	*10KLength*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Coefficient	-0.0179***	4.0334***	0.9773***	0.3469***	-0.1571***	0.0569***	0.1136***	-0.0022***	-0.1219	0.0010***	-0.0023***
(*t*-stat)	(-44.296)	(9.006)	(16.711)	(10.239)	(-16.303)	(17.840)	(16.556)	(-2.853)	(-1.214)	(3.674)	(-2.590)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs	119,106	118,435	115,432	64,713	65,196	118,406	26,678	82,525	15,224	64,138	63,688
Adjusted R^2	0.091	0.058	0.065	0.043	0.048	0.063	0.085	0.054	0.094	0.072	0.072

Panel A2: Negative Surprises, Single Explanatory Variable

Regression Specification: Abs Thirty-Day Return$_{i,t} = \alpha + \beta ExplanatoryVariable_{i,t} + \delta_t + \varepsilon_{i,t}$

		Illiquidity		Arbitrage Constraints		Retail Participation		Information Environment/Transparency			
Expl. Var.	*Size*	*AmIlliq*	*Spread*	*BorrowingFee*	*Lendable*	*RetailOwn*	*RetailTrading*	*MediaCover*	*Tangibility*	*FogIndex*	*10KLength*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Coefficient	-0.0110***	1.8781***	0.5188***	0.2800***	-0.1071***	0.0365***	0.0720***	-0.0017***	-0.3884***	0.0011***	0.0010
(*t*-stat)	(-43.998)	(7.293)	(16.313)	(13.991)	(-18.240)	(17.208)	(18.016)	(-3.557)	(-5.732)	(5.372)	(1.618)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs	118,532	117,731	115,194	61,968	62,198	117,764	30,337	77,092	12,326	64,568	64,109
Adjusted R^2	0.108	0.076	0.083	0.061	0.056	0.083	0.096	0.083	0.119	0.091	0.091

Table B2
Continued

Panel B1: Positive Surprises, Size plus a Competing Variable

Regression Specification: Abs Thirty-Day Return$_{i,t}$ = $\alpha + \beta Size_{i,t} + \gamma CompetingVariable_{i,t} + \delta_t + \varepsilon_{i,t}$

		Illiquidity		Arbitrage Constraints		Retail Participation		Information Environment/Transparency			
Competing Variable		*AmIlliq*	*Spread*	*BorrowingFee*	*Lendable*	*RetailOwn*	*RetailTrading*	*MediaCover*	*Tangibility*	*FogIndex*	*10KLength*
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Coefficient	*Size*	-0.0185***	-0.0178***	-0.0196***	-0.0191***	-0.0179***	-0.0100***	-0.0195***	-0.0207***	-0.0171***	-0.0175***
(*t*-stat)		(-44.109)	(-40.375)	(-37.747)	(-37.996)	(-42.377)	(-12.305)	(-37.511)	(-15.059)	(-32.561)	(-32.274)
Coefficient	Compet. Var.	-2.3587***	0.0213	0.1935***	-0.0478***	0.0003	0.0495***	0.0024***	-0.6469***	0.0014***	0.0045***
(*t*-stat)		(-5.331)	(0.388)	(6.527)	(-5.248)	(0.096)	(5.605)	(4.174)	(-6.153)	(5.539)	(5.166)
Time FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs		118,047	115,023	64,487	65,196	117,999	26,592	82,247	15,202	64,138	63,688
Adjusted R^2		0.092	0.091	0.087	0.086	0.091	0.091	0.090	0.128	0.102	0.102

Panel B2: Negative Surprises, Size plus a Competing Variable

Regression Specification: Abs Thirty-Day Return$_{i,t}$ = $\alpha + \beta Size_{i,t} + \gamma CompetingVariable_{i,t} + \delta_t + \varepsilon_{i,t}$

		Illiquidity		Arbitrage Constraints		Retail Participation		Information Environment/Transparency			
Competing Variable		*AmIlliq*	*Spread*	*BorrowingFee*	*Lendable*	*RetailOwn*	*RetailTrading*	*MediaCover*	*Tangibility*	*FogIndex*	*10KLength*
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Coefficient	*Size*	-0.0116***	-0.0111***	-0.0114***	-0.0114***	-0.0110***	-0.0068***	-0.0117***	-0.0071***	-0.0106***	-0.0110***
(*t*-stat)		(-44.251)	(-39.320)	(-37.428)	(-36.672)	(-39.356)	(-11.264)	(-36.231)	(-8.438)	(-32.084)	(-32.839)
Coefficient	Compet. Var.	-1.9102***	-0.0208	0.2082***	-0.0369***	0.0014	0.0294***	0.0014***	-0.5679***	0.0013***	0.0051***
(*t*-stat)		(-7.103)	(-0.722)	(11.604)	(-6.486)	(0.654)	(5.227)	(4.821)	(-8.157)	(7.279)	(8.710)
Time FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs		117,317	114,774	61,740	62,198	117,344	30,252	76,804	12,308	64,568	64,109
Adjusted R^2		0.108	0.108	0.102	0.093	0.108	0.102	0.115	0.128	0.119	0.120

Table B2
Continued

	Panel C1: Positive Surprises, Size plus Multiple Competing Variables					*Panel C2: Negative Surprises, Size plus Multiple Competing Variables*				
	*Regression Specification: Abs Thirty-Day Return$_{i,t}$ = α + βSize$_{i,t}$ + $\boldsymbol{\gamma}$**CompetingVariables**$_{i,t}$ + δ_t + $\varepsilon_{i,t}$*									
Expl. Var.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Size	-0.0184***	-0.0180***	-0.0218***	-0.0208***	-0.0075***	-0.0114***	-0.0116***	-0.0131***	-0.0135***	-0.0049***
	(-42.251)	(-32.222)	(-35.400)	(-26.698)	(-6.709)	(-39.950)	(-30.741)	(-36.331)	(-27.867)	(-6.354)
AmIlliq	-2.3460***	-4.2465***	-4.0235**	-2.0357	-6.5937***	-1.9596***	-3.5850***	-1.7223	-1.7912	-3.3915***
	(-5.163)	(-3.793)	(-2.037)	(-1.064)	(-3.740)	(-7.057)	(-5.955)	(-1.554)	(-1.391)	(-3.578)
BorrowingFee			0.1942***	0.2699***				0.2128***	0.2687***	
			(6.369)	(3.836)				(11.245)	(5.538)	
RetailOwn	0.0025	-0.0040	-0.0049	-0.0114**	0.0048	0.0032	-0.0046	-0.0090***	-0.0183***	-0.0023
	(0.772)	(-0.907)	(-1.254)	(-2.263)	(0.575)	(1.445)	(-1.572)	(-3.482)	(-5.605)	(-0.428)
RetailTrading					0.0720***					0.0602***
					(5.875)					(7.896)
MediaCover			0.0028***	0.0030***				0.0018***	0.0021***	
			(4.348)	(2.960)				(5.319)	(3.834)	
FogIndex		0.0014***		0.0005*	0.0067***		0.0013***		0.0003	0.0056***
		(5.447)		(1.771)	(6.696)		(7.138)		(1.562)	(8.312)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs	116,957	63,602	63,570	32,518	16,154	116,154	63,984	60,743	32,188	18,386
Adjusted R^2	0.092	0.103	0.090	0.104	0.101	0.109	0.120	0.105	0.109	0.114

Table B3

Absolute Thirty-Day Return at Earnings Announcements, Single Sort by Stock Characteristics

This table shows how the stock price correction at earnings announcements varies across the quintiles of different stock characteristics (described in Table 2). Each quarter we assign a stock to one of the five groups (quintiles) according to the value of the sorting variable prior to the earnings announcement. Q1 (Q5) indicates the quintile with the lowest (highest) values. The table reports the mean of *Abs Thirty-Day Return* (defined in Table B2) in each quintile as well as the difference in means of *Abs Thirty-Day Return*, with the corresponding *t*-statistic, between the highest and the lowest quintile. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively. Panel A (Panel B) shows the analysis for positive (negative) surprises. The sample period for each sorting variable is reported in Table 1.

Panel A: Positive Surprises

Abs Thirty-Day Return by Quintiles of Stock Characteristics

		Illiquidity		Arbitrage Constraints		Retail Participation		Information Environment/Transparency			
Sorting Variable	*Size* (1)	*AmIlliq* (2)	*Spread* (3)	*BorrowingFee* (4)	*Lendable* (5)	*RetailOwn* (6)	*RetailTrading* (7)	*MediaCover* (8)	*Tangibility* (9)	*FogIndex* (10)	*10KLength* (11)
Q1	0.0751	0.0468	0.0509	0.0569	0.1503	0.0629	0.0496	0.1345	0.0601	0.1221	0.1308
Q2	0.0695	0.0589	0.0577	0.0586	0.1276	0.0607	0.0508	0.1429	0.0582	0.1342	0.1364
Q3	0.0642	0.0654	0.0634	0.0614	0.1055	0.0610	0.0547	0.1281	0.0561	0.1348	0.1394
Q4	0.0570	0.0698	0.0691	0.0645	0.1039	0.0617	0.0591	0.1287	0.0549	0.1378	0.1322
Q5	0.0452	0.0703	0.0735	0.0783	0.1094	0.0655	0.0719	0.1123	0.0548	0.1377	0.1278
Q5-Q1	-0.0834***	0.0689***	0.0715***	0.0644***	-0.0409***	0.0352***	0.0533***	-0.0222***	-0.0053*	0.0155***	-0.0031
(*t*-stat)	(-53.010)	(45.282)	(42.732)	(26.761)	(-18.661)	(21.680)	(16.227)	(-11.979)	(-1.604)	(7.799)	(-1.598)

Panel B: Negative Surprises

Abs Thirty-Day Return by Quintiles of Stock Characteristics

		Illiquidity		Arbitrage Constraints		Retail Participation		Information Environment/Transparency			
Sorting Variable	*Size* (1)	*AmIlliq* (2)	*Spread* (3)	*BorrowingFee* (5)	*Lendable* (4)	*RetailOwn* (6)	*RetailTrading* (7)	*MediaCover* (8)	*Tangibility* (9)	*FogIndex* (10)	*10KLength* (11)
Q1	0.1330	0.0861	0.0913	0.0832	0.1208	0.1056	0.1067	0.1116	0.0929	0.1059	0.1094
Q2	0.1236	0.1048	0.1021	0.0843	0.1042	0.1023	0.1116	0.1203	0.0939	0.1124	0.1167
Q3	0.1144	0.1158	0.1113	0.0938	0.0901	0.1061	0.1187	0.1075	0.0909	0.1157	0.1183
Q4	0.1010	0.1228	0.1211	0.1032	0.0882	0.1146	0.1227	0.1060	0.0868	0.1193	0.1143
Q5	0.0825	0.1256	0.1326	0.1354	0.0927	0.1273	0.1432	0.0980	0.0794	0.1195	0.1141
Q5-Q1	-0.0506***	0.0395***	0.0413***	0.0522***	-0.0281***	0.0218***	0.0365***	-0.0136***	-0.0135***	0.0136***	0.0047***
(*t*-stat)	(-55.169)	(43.291)	(42.074)	(38.980)	(-22.557)	(22.125)	(18.271)	(-11.619)	(-5.743)	(10.400)	(3.606)

Table B4

Absolute Thirty-Day Return at Earnings Announcements, Sequential Double Sort

Panel A of this table shows the effect of stock size on the earnings announcement price reaction conditional on sorting by different stock characteristics (described in Table 2). Each quarter we assign a stock to one of the five groups (quintiles) according to the value of a given characteristic prior to the earnings announcement. Q1 (Q5) indicates the quintile with the lowest (highest) values. Then, within each of these quintiles, we further sort stocks into five quintiles of size. The table reports the difference in means of *Abs Thirty-Day Return* (defined in Table B2), with the corresponding *t*-statistic, between the highest and the lowest size quintile. Panel A1 (Panel A2) shows the results for positive (negative) surprises. Panel B reports the output of the reverse procedure: first, we sort stocks by size, and then, within each size quintile, we further sort stocks by a given characteristic. Panel B1 (Panel B2) shows the analysis for positive (negative) surprises. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively. The sample period for each sorting variable is reported in Table 1.

Panel A1: Positive Surprises, Pre-Sort by a Competing Variable, Main Sort by Size

Q5-Q1 Difference in Means of Abs Thirty-Day Return

Pre-sort by Compet. Var. Quintile	Main sort by *Size*	Illiquidity		Arbitrage Constraints		Retail Participation		Information Environment/Transparency			
		AmIlliq	*Spread*	*BorrowingFee*	*Lendable*	*RetailOwn*	*RetailTrading*	*MediaCover*	*Tangibility*	*FogIndex*	*10KLength*
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	Q5-Q1	-0.0436***	-0.0442***	-0.0726***	-0.0732***	-0.0568***	-0.0570***	-0.0753***	-0.0815***	-0.0736***	-0.0718***
	(*t*-stat)	(-20.366)	(-17.653)	(-18.637)	(-11.868)	(-20.394)	(-10.829)	(-14.703)	(-9.938)	(-17.333)	(-16.707)
2	Q5-Q1	-0.0457***	-0.0397***	-0.0696***	-0.0930***	-0.0696***	-0.0493***	-0.0732***	-0.0748***	-0.0769***	-0.0748***
	(*t*-stat)	(-17.510)	(-15.667)	(-23.338)	(-22.610)	(-25.615)	(-9.437)	(-17.195)	(-8.876)	(-15.355)	(-14.431)
3	Q5-Q1	-0.0524***	-0.0431***	-0.0734***	-0.0901***	-0.0812***	-0.0353***	-0.0961***	-0.0684***	-0.0750***	-0.0860***
	(*t*-stat)	(-16.235)	(-15.909)	(-21.052)	(-25.806)	(-26.188)	(-5.689)	(-24.631)	(-8.715)	(-15.587)	(-16.926)
4	Q5-Q1	-0.0650***	-0.0362***	-0.0682***	-0.0741***	-0.0906***	-0.0316***	-0.0975***	-0.0561***	-0.0825***	-0.0822***
	(*t*-stat)	(-17.582)	(-10.505)	(-17.291)	(-20.744)	(-26.694)	(-4.392)	(-24.241)	(-7.984)	(-17.915)	(-18.080)
5	Q5-Q1	-0.0613***	-0.0484***	-0.0781***	-0.0693***	-0.0736***	-0.0436***	-0.0913***	-0.0661***	-0.0905***	-0.0897***
	(*t*-stat)	(-13.304)	(-9.906)	(-11.684)	(-17.976)	(-16.293)	(-5.094)	(-20.599)	(-8.000)	(-16.902)	(-18.062)

Panel A2: Negative Surprises, Pre-Sort by a Competing Variable, Main Sort by Size

Q5-Q1 Difference in Means of Abs Thirty-Day Return

Pre-sort by Compet. Var. Quintile	Main sort by *Size*	Illiquidity		Arbitrage Constraints		Retail Participation		Information Environment/Transparency			
		AmIlliq	*Spread*	*BorrowingFee*	*Lendable*	*RetailOwn*	*RetailTrading*	*MediaCover*	*Tangibility*	*FogIndex*	*10KLength*
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	Q5-Q1	-0.0383***	-0.0353***	-0.0405***	-0.0337***	-0.0362***	-0.0509***	-0.0465***	-0.0332***	-0.0425***	-0.0441***
	(*t*-stat)	(-21.849)	(-18.675)	(-16.120)	(-10.924)	(-17.870)	(-13.042)	(-17.454)	(-6.466)	(-15.834)	(-16.330)
2	Q5-Q1	-0.0401***	-0.0302***	-0.0404***	-0.0580***	-0.0448***	-0.0286***	-0.0376***	-0.0252***	-0.0401***	-0.0440***
	(*t*-stat)	(-18.898)	(-15.246)	(-21.446)	(-21.684)	(-23.657)	(-7.303)	(-13.044)	(-4.417)	(-14.035)	(-15.059)
3	Q5-Q1	-0.0428***	-0.0275***	-0.0436***	-0.0560***	-0.0503***	-0.0161***	-0.0573***	-0.0196***	-0.0423***	-0.0440***
	(*t*-stat)	(-19.296)	(-13.403)	(-16.606)	(-24.192)	(-25.573)	(-3.977)	(-21.890)	(-3.598)	(-14.688)	(-14.555)
4	Q5-Q1	-0.0360***	-0.0199***	-0.0413***	-0.0485***	-0.0487***	-0.0107**	-0.0536***	-0.0161***	-0.0511***	-0.0548***
	(*t*-stat)	(-15.883)	(-8.996)	(-15.834)	(-20.046)	(-22.909)	(-2.523)	(-21.426)	(-3.278)	(-16.972)	(-19.091)
5	Q5-Q1	-0.0326***	-0.0241***	-0.0322***	-0.0385***	-0.0323***	-0.0151***	-0.0603***	-0.0080	-0.0580***	-0.0639***
	(*t*-stat)	(-13.936)	(-9.876)	(-9.519)	(-15.740)	(-13.596)	(-3.119)	(-24.771)	(-1.552)	(-19.060)	(-21.218)

Table B4
Continued

Panel B1: Positive Surprises, Pre-Sort by Size, Main Sort by a Competing Variable
Q5-Q1 Difference in Means of Abs Thirty-Day Return

Pre-sort by *Size* Quintile	Main sort by Compet. Var.	Illiquidity *AmIlliq* (1)	Illiquidity *Spread* (2)	Arbitrage Constraints *BorrowingFee* (3)	Arbitrage Constraints *Lendable* (4)	Retail Participation *RetailOwn* (5)	Retail Participation *RetailTrading* (6)	Information Environment/Transparency *MediaCover* (7)	Information Environment/Transparency *Tangibility* (8)	Information Environment/Transparency *FogIndex* (9)	Information Environment/Transparency *10KLength* (10)
1	Q5-Q1	-0.0259***	0.0258***	0.0385***	0.0159***	-0.0021	0.0704***	0.0483***	-0.0361***	0.0183***	0.0229***
	(*t*-stat)	(-6.051)	(5.564)	(6.557)	(2.764)	(-0.492)	(8.059)	(8.043)	(-3.520)	(2.911)	(3.819)
2	Q5-Q1	-0.0226***	0.0185***	0.0285***	-0.0108**	0.0026	0.0105	0.0286***	-0.0249***	0.0192***	0.0164***
	(*t*-stat)	(-6.399)	(4.970)	(5.475)	(-2.499)	(0.721)	(1.337)	(6.537)	(-3.608)	(3.896)	(3.527)
3	Q5-Q1	-0.0186***	0.0101***	0.0414***	-0.0245***	0.0044	0.0053	0.0211***	-0.0320***	0.0154***	0.0126***
	(*t*-stat)	(-6.402)	(3.099)	(8.566)	(-5.743)	(1.395)	(0.794)	(5.566)	(-4.687)	(3.767)	(3.311)
4	Q5-Q1	-0.0080***	0.0072***	0.0266***	-0.0011	-0.0107***	-0.0028	0.0184***	-0.0250***	0.0126***	0.0075**
	(*t*-stat)	(-3.224)	(2.754)	(8.280)	(-0.373)	(-4.143)	(-0.430)	(6.152)	(-4.056)	(3.587)	(2.234)
5	Q5-Q1	0.0164***	0.0134***	0.0085***	0.0030	-0.0142***	0.0087*	-0.0033	-0.0043	0.0128***	0.0036
	(*t*-stat)	(8.268)	(5.873)	(3.705)	(1.369)	(-6.765)	(1.910)	(-1.252)	(-0.975)	(4.561)	(1.268)

Panel B2: Negative Surprises, Pre-Sort by Size, Main Sort by a Competing Variable
Q5-Q1 Difference in Means of Abs Thirty-Day Return

Pre-sort by *Size* Quintile	Main sort by Compet. Var.	Illiquidity *AmIlliq* (1)	Illiquidity *Spread* (2)	Arbitrage Constraints *BorrowingFee* (3)	Arbitrage Constraints *Lendable* (4)	Retail Participation *RetailOwn* (5)	Retail Participation *RetailTrading* (6)	Information Environment/Transparency *MediaCover* (7)	Information Environment/Transparency *Tangibility* (8)	Information Environment/Transparency *FogIndex* (9)	Information Environment/Transparency *10KLength* (10)
1	Q5-Q1	-0.0137***	0.0151***	0.0427***	-0.0061*	0.0089***	0.0434***	0.0147***	-0.0332***	0.0181***	0.0235***
	(*t*-stat)	(-5.718)	(6.193)	(12.899)	(-1.899)	(3.726)	(9.704)	(4.701)	(-5.640)	(5.405)	(7.155)
2	Q5-Q1	-0.0266***	0.0018	0.0313***	-0.0057*	0.0044*	0.0180***	0.0174***	-0.0248***	0.0150***	0.0140***
	(*t*-stat)	(-11.587)	(0.760)	(9.059)	(-1.935)	(1.907)	(3.788)	(6.009)	(-4.694)	(4.865)	(4.525)
3	Q5-Q1	-0.0181***	0.0061***	0.0413***	-0.0189***	0.0037*	-0.0027	0.0093***	-0.0297***	0.0175***	0.0174***
	(*t*-stat)	(-8.264)	(2.687)	(13.594)	(-7.028)	(1.722)	(-0.591)	(3.508)	(-5.305)	(5.909)	(5.973)
4	Q5-Q1	-0.0141***	0.0008	0.0275***	-0.0048**	-0.0088***	-0.0056	0.0229***	-0.0068	0.0132***	0.0120***
	(*t*-stat)	(-6.920)	(0.375)	(10.837)	(-2.063)	(-4.304)	(-1.232)	(9.569)	(-1.224)	(4.861)	(4.317)
5	Q5-Q1	0.0110***	0.0071***	0.0090***	0.0058***	-0.0111***	0.0148***	-0.0003	-0.0057	0.0075***	0.0057**
	(*t*-stat)	(6.531)	(3.894)	(4.167)	(3.039)	(-6.094)	(4.221)	(-0.132)	(-1.213)	(3.207)	(2.473)

Table B5

Size vs. Other Stock Characteristics as Drivers of Institutional Attention, Multivariate Regressions

This table reports the results of quarterly panel regressions of institutional attention on stock size and other stock characteristics. The main dependent variable in the left pane of the table is *InstAtten*, defined as in Table 5. The right pane of the table shows the results for an alternative measure of institutional attention—*EDGAR*, calculated as the natural logarithm of the average daily number of searches for the firm's filings in the EDGAR database in a quarter. The competing explanatory variables include *AmIlliq*, *Spread*, *BorrowingFee*, *Lendable*, *RetailOwn*, *MediaCover*, *FogIndex*, *EarnVolatility*, *BM*, *Leverage*, and *R&D* (defined as in Table 5). The sample period in column (1) is from February 2010 to December 2016. The sample period in columns (2)-(4) is from February 2010 to December 2012. The sample period in columns (5)-(8) is from January 2003 to December 2016. The inclusion of time (year-quarter) fixed effects is indicated at the bottom of the table. *T*-statistics are reported in parentheses and are derived from standard errors clustered by stock. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Panel C: Size plus Multiple Competing Variables

Regression Specification: $Attention_{i,t} = \alpha + \beta Size_{i,t} + \boldsymbol{\gamma CompetingVariables_{i,t}} + \delta_t + \varepsilon_{i,t}$

	Dependent Variable							
	InstAtten				*EDGAR*			
Expl. Var.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Size	0.2158***	0.2449***	0.2462***	0.2353***	0.3026***	0.3057***	0.3196***	0.3499***
	(37.607)	(33.207)	(25.257)	(19.193)	(28.547)	(33.844)	(24.432)	(23.691)
AmIlliq	0.0130***		-0.0146	-0.0226**	0.0009		-0.0118*	-0.0098**
	(6.510)		(-1.590)	(-2.542)	(0.526)		(-1.862)	(-2.166)
Spread		13.1509***	30.2034***	35.3693***		1.2274**	12.5860***	11.6186***
		(6.695)	(2.687)	(3.486)		(1.972)	(4.748)	(3.275)
BorrowingFee	2.0464***		2.7691***	2.4784***	1.6547***		3.6292***	3.4752***
	(9.030)		(4.062)	(2.916)	(5.796)		(8.331)	(6.446)
Lendable		-0.2585	-0.0202	-0.1403		0.2019**	0.1096	0.0512
		(-1.361)	(-0.091)	(-0.730)		(2.031)	(0.898)	(0.366)
RetailOwn	-0.0418	-0.2099***	-0.1964**	-0.1503	-0.1046**	0.0576	0.0827	0.2882***
	(-1.268)	(-2.919)	(-2.066)	(-1.364)	(-2.520)	(1.242)	(1.340)	(4.103)
MediaCover	0.0224***		0.0082***	0.0085***	0.0698***		0.0582***	0.0452***
	(4.776)		(4.744)	(6.078)	(4.758)		(3.339)	(3.263)
FogIndex		0.0010	0.0021	0.0015		0.0055***	0.0063***	0.0069***
		(0.968)	(1.516)	(0.826)		(3.487)	(3.076)	(2.807)
EarnVolatility	0.1072***		0.1035***	0.2089**	0.1388***		0.0858***	0.3444***
	(4.471)		(3.465)	(2.391)	(2.724)		(2.943)	(5.326)
BM		0.0132	0.0029	-0.0115		0.0458***	0.0620***	0.0799***
		(1.041)	(0.188)	(-0.589)		(4.197)	(5.228)	(5.437)
Leverage			0.4206***	0.3687***			0.4709***	0.3794***
			(6.375)	(3.657)			(8.558)	(4.802)
R&D				1.0794				1.4303***
				(1.435)				(3.429)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs	37,550	11,246	7,816	4,705	103,365	80,332	46,715	29,179
Adjusted R^2	0.577	0.533	0.551	0.548	0.637	0.611	0.666	0.683

Table B6

The Effect of Institutional Attention vs. Other Stock Characteristics on the Price Correction at Earnings Announcements, Thirty-Day Price Correction

This table reports the results of the panel regressions of the stock price correction at earnings announcements on institutional attention and other explanatory variables. The dependent variable is *Abs Thirty-Day Return*, defined as in Table B2. Panel A (Panel B) shows the analysis for positive (negative) surprises. The main independent variable is *InstAtten*, defined as in Table 6. The competing explanatory variables include *AmIlliq*, *Spread*, *BorrowingFee*, *Lendable*, *RetailOwn*, *MediaCover*, *FogIndex* (defined as in Table 2), and *GoogleSearch*, defined as in Table 6. The sample period in columns (1)-(7) is from February 2010 to December 2016 and the sample period in column (8) is from February 2010 to December 2012. The inclusion of time (year-quarter) fixed effects is indicated at the bottom of the table. *T*-statistics are reported in parentheses and are derived from standard errors clustered by stock. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Panel A: Positive Surprises, Attention vs. Other Stock Characteristics

Regression Specification: $Abs\ Thirty\text{-}Day\ Return_{i,t} = \alpha + \beta InstAtten_{i,t} + \gamma CompetingVariable_{i,t} + \delta_t + \varepsilon_{i,t}$

		Illiquidity		Arbitrage Constraints		Retail Participation		Information Environment/Transparency	
Competing Variable		*AmIlliq*	*Spread*	*BorrowingFee*	*Lendable*	*RetailOwn*	*GoogleSearch*	*MediaCover*	*FogIndex*
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Coefficient	*InstAtten*	-0.0187***	-0.0148***	-0.0192***	-0.0151***	-0.0160***	-0.0154***	-0.0193***	-0.0154***
(*t*-stat)		(-11.637)	(-9.229)	(-12.222)	(-9.719)	(-10.016)	(-7.744)	(-11.407)	(-8.351)
Coefficient	Compet. Var.	33.6564***	3.0778***	0.2904***	-0.1665***	0.0582***	-0.0042***	-0.0004	-0.0002
(*t*-stat)		(5.863)	(8.506)	(7.983)	(-13.586)	(10.790)	(-6.096)	(-1.009)	(-0.753)
Time FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs		32,540	32,839	32,796	32,663	32,402	27,020	32,839	7,321
Adjusted R^2		0.021	0.029	0.026	0.033	0.027	0.019	0.020	0.029

Panel B: Negative Surprises, Attention vs. Other Stock Characteristics

Regression Specification: $Abs\ Thirty\text{-}Day\ Return_{i,t} = \alpha + \beta InstAtten_{i,t} + \gamma CompetingVariable_{i,t} + \delta_t + \varepsilon_{i,t}$

		Illiquidity		Arbitrage Constraints		Retail Participation		Information Environment/Transparency	
Competing Variable		*AmIlliq*	*Spread*	*BorrowingFee*	*Lendable*	*RetailOwn*	*GoogleSearch*	*MediaCover*	*FogIndex*
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Coefficient	*InstAtten*	-0.0119***	-0.0101***	-0.0115***	-0.0094***	-0.0100***	-0.0104***	-0.0120***	-0.0101***
(*t*-stat)		(-10.192)	(-8.524)	(-10.261)	(-8.036)	(-8.339)	(-7.106)	(-9.518)	(-6.465)
Coefficient	Compet. Var.	18.6474***	1.4975***	0.2362***	-0.1149***	0.0368***	-0.0016***	-0.0004	-0.0001
(*t*-stat)		(4.659)	(8.873)	(12.308)	(-15.882)	(10.544)	(-3.149)	(-1.353)	(-0.338)
Time FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs		30,630	30,942	30,888	30,726	30,486	25,458	30,942	7,969
Adjusted R^2		0.022	0.028	0.044	0.038	0.030	0.019	0.021	0.017

Table B7
Trading Discreteness and Stock Characteristics

This table reports the results of quarterly panel regressions of trading discreteness on stock characteristics. The trading discreteness measures are computed as in Panel C of Table 7; however, we divide *TimeBetweenTrades* by 100 and multiply *DistanceBetweenLots1* and *DistanceBetweenLots2* by 100 before running the regressions. The independent variables include *AmIlliq*, *Spread*, *BorrowingFee*, *MediaCover*, and *FogIndex*, defined as in Table 5. The sample periods are as follows: from January 1999 to December 2010 in columns (1)-(2) and (5) of Panels A and B, from June 2002 to December 2010 in column (3) of Panels A and B, from January 2000 to December 2010 in column (4) of Panels A and B, from January 1999 to June 2013 in columns (1)-(2) of Panels C and D, from June 2002 to June 2013 in column (3) of Panels C and D, from January 2000 to June 2013 in column (4) of Panels C and D, and from January 1999 to December 2012 in column (5) of Panels C and D. The inclusion of time (year-quarter) fixed effects is indicated at the bottom of the table. *T*-statistics are reported in parentheses and are derived from standard errors clustered by stock. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Panel A: Dependent Variable - Average Time between Trades within Fund

Regression Specification: $TimeBetweenTrades_{i,t} = \alpha + \beta ExplanatoryVariable_{i,t} + \delta_t + \varepsilon_{i,t}$

	Illiquidity		Arbitrage Constraints	Information Environment/Transparency	
Expl. Var.	*AmIlliq*	*Spread*	*BorrowingFee*	*MediaCover*	*FogIndex*
	(1)	(2)	(3)	(4)	(5)
Coefficient	0.0008***	0.4509***	0.2032***	-0.0032***	-0.0004***
(*t*-stat)	(9.981)	(20.533)	(12.278)	(-6.030)	(-3.967)
Time FE	Yes	Yes	Yes	Yes	Yes
NObs	168,604	168,505	95,490	152,576	127,342
Adjusted R^2	0.027	0.036	0.025	0.023	0.027

Panel B: Dependent Variable - Average Distance between Unique Trading Lots within Fund

Regression Specification: $DistanceBetweenLots1_{i,t} = \alpha + \beta ExplanatoryVariable_{i,t} + \delta_t + \varepsilon_{i,t}$

	Illiquidity		Arbitrage Constraints	Information Environment/Transparency	
Expl. Var.	*AmIlliq*	*Spread*	*BorrowingFee*	*MediaCover*	*FogIndex*
	(1)	(2)	(3)	(4)	(5)
Coefficient	0.0007***	0.4116***	0.0810***	-0.0016***	-0.0002*
(*t*-stat)	(8.762)	(15.734)	(8.078)	(-5.663)	(-1.884)
Time FE	Yes	Yes	Yes	Yes	Yes
NObs	164,144	164,037	93,928	148,541	124,072
Adjusted R^2	0.018	0.026	0.023	0.018	0.018

Panel C: Dependent Variable - Average Distance between Unique Trading Lots across Funds

Regression Specification: $DistanceBetweenLots2_{i,t} = \alpha + \beta ExplanatoryVariable_{i,t} + \delta_t + \varepsilon_{i,t}$

	Illiquidity		Arbitrage Constraints	Information Environment/Transparency	
Expl. Var.	*AmIlliq*	*Spread*	*BorrowingFee*	*MediaCover*	*FogIndex*
	(1)	(2)	(3)	(4)	(5)
Coefficient	0.0009***	0.4363***	0.0376***	-0.0006***	-0.0001**
(*t*-stat)	(12.722)	(20.024)	(8.700)	(-4.397)	(-2.449)
Time FE	Yes	Yes	Yes	Yes	Yes
NObs	194,090	194,089	118,468	177,560	140,075
Adjusted R^2	0.014	0.029	0.005	0.008	0.011

Panel D: Dependent Variable - Herfindahl Concentration Index Based on Frequency of Unique Trading Lots across Funds

Regression Specification: $LotConcentration_{i,t} = \alpha + \beta ExplanatoryVariable_{i,t} + \delta_t + \varepsilon_{i,t}$

	Illiquidity		Arbitrage Constraints	Information Environment/Transparency	
Expl. Var.	*AmIlliq*	*Spread*	*BorrowingFee*	*MediaCover*	*FogIndex*
	(1)	(2)	(3)	(4)	(5)
Coefficient	0.0132***	5.1604***	0.5816***	-0.0063***	-0.0015***
(*t*-stat)	(18.983)	(34.010)	(12.029)	(-4.432)	(-3.933)
Time FE	Yes	Yes	Yes	Yes	Yes
NObs	205,119	205,291	122,727	187,032	147,671
Adjusted R^2	0.082	0.168	0.030	0.031	0.034

Table B8
Sensitivity of Trading Discreteness to Profit Potential by Fund Size

This table reports the results of quarterly panel regressions of trading discreteness on stock size, competing explanatory variables, and their interactions with fund size proxies. We use two fund-stock-specific trading discreteness measures—*TimeBetweenTrades* and *DistanceBetweenLots1*, as described in Panel B of Table 7. The main independent variables are *StockSize*, *FundVolume*, and *BigFund*, defined as in Table 10. The competing explanatory variables include *AmIlliq*, *Spread*, *BorrowingFee*, *MediaCover*, and *FogIndex* (defined as in Table 2). We divide *TimeBetweenTrades* by 100 and multiply *DistanceBetweenLots1* by 100 before running the regressions; the coefficients on *StockSize* and its interactions are further scaled by 100; the coefficients on *FogIndex* and its interactions are further scaled by 10^5. The sample periods are as follows: from January 1999 to December 2010 in columns (1)-(2), (5)-(7), and (10); from June 2002 to December 2010 in columns (3) and (8); from January 2000 to December 2010 in columns (4) and (9). The inclusion of time (year-quarter) fixed effects is indicated at the bottom of the table. *T*-statistics are reported in parentheses and are derived from standard errors double-clustered by fund and stock. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Panel A: TimeBetweenTrades Regressed on Fund Size Interacted with Stock Size and Other Stock Characteristics

Expl. Var.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
StockSize	7.3419***	7.3913***	8.0800***	7.3500***	7.1583***	1.2362***	1.2044***	1.2422***	1.3234***	1.1936***
	(7.474)	(7.013)	(6.525)	(6.753)	(7.397)	(12.207)	(11.799)	(10.495)	(11.618)	(11.688)
FundVolu	0.0389***	0.0394***	0.0439***	0.0383***	0.0371***					
	(5.062)	(4.695)	(4.578)	(4.509)	(4.806)					
StockSize	-0.3232***	-0.3262***	-0.3588***	-0.3191***	-0.3164***					
	(-6.736)	(-6.296)	(-5.931)	(-6.012)	(-6.708)					
BigFund						0.1320***	0.1261***	0.1418***	0.1330***	0.1304***
						(4.382)	(3.935)	(3.936)	(3.988)	(4.392)
StockSize						-1.0306***	-1.0011***	-1.1131***	-1.0356***	-1.0112***
						(-5.224)	(-4.868)	(-4.751)	(-4.687)	(-5.167)
AmIlliq	0.0100***					0.0013**				
	(2.825)					(2.159)				
AmIlliq ×	-0.0004***									
	(-2.725)									
AmIlliq ×						-0.0017**				
						(-2.368)				
Spread		0.3776					-0.3827*			
		(0.317)					(-1.688)			
Spread ×		-0.0234								
		(-0.426)								
Spread ×							0.3477			
							(1.219)			
Borrowing			-0.6060***					-0.0487		
			(-2.603)					(-1.460)		
Borrowing			0.0278**							
			(2.503)							
Borrowing								0.0817*		
								(1.791)		
MediaCov				0.0017					-0.0009*	
				(0.723)					(-1.884)	
MediaCov				-0.0001						
				(-1.341)						
MediaCov									-0.0004	
									(-0.932)	
FogIndex					-9.4134					-2.2466
					(-0.654)					(-0.742)
FogIndex					0.2891					
					(0.461)					
FogIndex										-0.8886
										(-0.283)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs	2,335,356	2,326,666	1,565,234	2,192,963	1,816,645	2,335,356	2,326,666	1,565,234	2,192,963	1,816,645
Adjusted	0.036	0.036	0.037	0.037	0.037	0.018	0.018	0.019	0.019	0.018

Table B8
Continued

Panel B: DistanceBetweenLots1 Regressed on Fund Size Interacted with Stock Size and Other Stock Characteristics

Expl. Var.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
StockSize	0.5987**	0.5193**	0.4249**	0.5665**	0.5848**	-0.1124***	-0.1092***	-0.1009***	-0.1284***	-0.1138***
	(2.534)	(2.468)	(2.162)	(2.274)	(2.238)	(-12.527)	(-12.308)	(-9.535)	(-12.278)	(-12.669)
FundVolu	0.0064***	0.0057***	0.0049***	0.0062***	0.0063***					
	(3.509)	(3.568)	(3.093)	(3.265)	(3.111)					

StockSize	-0.0393***	-0.0350***	-0.0294***	-0.0383***	-0.0385***					
	(-3.427)	(-3.434)	(-3.071)	(-3.174)	(-3.043)					
BigFund						0.0328***	0.0304***	0.0264***	0.0318***	0.0321***
						(7.113)	(7.162)	(6.033)	(6.686)	(6.626)
StockSize						-0.1968***	-0.1831***	-0.1565***	-0.1930***	-0.1924***
						(-6.900)	(-6.962)	(-5.939)	(-6.484)	(-6.332)
AmIlliq	0.0027					0.0007**				
	(0.834)					(2.406)				
AmIlliq ×	-0.0001									
	(-0.652)									
AmIlliq ×						-0.0003				
						(-0.544)				
Spread		-0.1355					0.0781***			
		(-0.347)					(3.176)			
Spread ×		0.0122								
		(0.658)								
Spread ×							0.0882*			
							(1.822)			
Borrowing			-0.1011***					0.0052**		
			(-4.339)					(2.370)		
Borrowing			0.0056***							
			(4.940)							
Borrowing								0.0239***		
								(4.352)		
MediaCov				-0.0010*					0.0002***	
				(-1.828)					(4.852)	
MediaCov				0.0001**						
				(2.313)						
MediaCov									0.0002***	
									(3.956)	
FogIndex					0.6473					0.4970**
					(0.301)					(2.354)
FogIndex					-0.0079					
					(-0.080)					
FogIndex										0.0021
										(0.009)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NObs	2,239,415	2,231,137	1,507,189	2,103,539	1,742,264	2,239,415	2,231,137	1,507,189	2,103,539	1,742,264
Adjusted	0.052	0.053	0.045	0.050	0.050	0.051	0.052	0.045	0.050	0.049