Rational Information Acquisition Screens*

Paul E. Fischer[1], Mirko S. Heinle[1], Jung Min Kim[2], and Christina Zhu[1]

[1]The Wharton School, University of Pennsylvania
[2]Kellogg School of Management, Northwestern University

February 2, 2024

Abstract

We develop and empirically test a model of information acquisition in capital markets. In the model, we assume that investors are uncertain about the precision of the private information *before* they acquire it. As a result, investors use prior prices and public information as a screen to estimate the value of available private information and efficiently allocate their limited information processing capacity across firms. The model predicts that larger unexplained price movements lead to more private information acquisition, higher future price volatility, and higher future trading volumes. Using fine-grained data measuring information acquisition on Edgar and Bloomberg, we provide empirical evidence in support of the model's predictions.

Keywords: Information Acquisition; Market for Information; Higher-Moment Uncertainty; Investor Attention; Edgar Downloads; Bloomberg; Earnings Announcements; Big Data

*Fischer: `pef@wharton.upenn.edu`. Heinle: `miheinle@wharton.upenn.edu`. Kim: `jungmin.kim@kellogg.northwestern.edu`. Zhu: `chrzhu@wharton.upenn.edu`. We thank seminar participants at the University of Michigan for helpful comments. We also thank Allan Ngo for excellent research assistance and Travis Dyer for helping us understand the American Registry of Internet Numbers (ARIN) data. An earlier version of this paper was circulated with the title "A theory of rational investment screens". We thank Haresh Sapra (discussant) as well as workshop participants at UC Berkeley, Columbia University, UNC, Northwestern University, the UCLA Accounting Mini Conference, and the LBS Accounting Symposium for valuable comments and suggestions on the earlier version of the paper.

1 Introduction

Researchers have documented that investors allocate differential amounts of attention across firms, which is consistent with information processing constraints forcing investors to choose the extent to which they follow each firm.[1] However, little is known about how investors manage information processing constraints and the implications of how they do so. We posit that that some investors employ information acquisition screens to efficiently identify firms with the most profitable private information acquisition opportunities, which then guides how the investors allocate their information processing efforts. We develop a theoretical model that provides insight into how screens can be useful for allocating investor attention. From the model, we extract predictions regarding the relation between a simple screen outcome and future investor attention, trading volume, and price volatility. Using a large sample of earnings announcements of publicly traded U.S. firms from 2004 to 2022, we test the model's empirical implications.

An underlying premise of our theory is that some investors are uncertain about which firms offer the most promising private information acquisition opportunities and, hence, expected information acquisition profits, but they do have access to databases containing past prices, earnings, book values, etc. that can be employed in a screen to resolve some of that uncertainty. Those investors then use past prices and other public information to determine which firms offer the best information acquisition opportunities. To illustrate our idea formally, we employ a two-period perfectly competitive model of trade for two firms in which there are three types of traders: speculators (or active traders), rational uninformed investors (or passive traders), and noise traders. Prior to the first-period trades, each firm releases a public signal and the speculators observe private information. Speculators, rational uninformed investors, and noise traders then trade in the shares of the firm. After the first period of trade, a new set of speculators arrives. Those new speculators are resource-constrained and must choose which of the two firms to follow.[2] After they make their choices, all speculators, both established and new, obtain private information about the firm they follow and engage in a second period of trade with the uninformed investors and noise traders. The critical assumption in our model is that, when the new speculators make their firm-following choice,

[1]For example, studies have documented an increase in investors' information acquisition activities for firms with negative returns (e.g., Drake, Roulstone, and Thornock, 2015; Kempf, Manconi, and Spalt, 2017), positive earnings surprises (e.g., Brown, Hillegeist, and Lo, 2009; Koester, Lundholm, and Soliman, 2016), negative earnings shocks (e.g., Drake, Roulstone, and Thornock, 2016), and location or education commonalities with investors (e.g., Chen, Cohen, Gurun, Lou, and Malloy, 2020; Dyer, 2021).

[2]In other words, in lieu of assuming an exogenous cost of information acquisition, we assume that speculators face an endogenous opportunity cost because they have to choose one of two firms to follow.

they do not know the quality of the private information that can be attained by following each firm. This implies that new speculators will try to increase their expected trading profits by increasing the chance that they follow a firm with a higher quality of private information. When they make their firm-following choice, new speculators have access to past public information, which are earnings and past prices in our model. That information, coupled with the belief that the quality of information is correlated across time, informs their firm following decision. In effect, not only do prices allow uninformed investors to glean information about a firm's future cash flows, as is common in rational expectations equilibria, but they also allow new speculators to glean information about the value of acquiring private information. In this sense, our study differs from prior literature in which past prices are typically not informative about the potential profits to private information acquisition.

At a broad level, our model suggests that screens can be useful when the value to acquiring private information is unknown to some market participants and is correlated over time. The explicit, endogenous screen in the model is optimally contingent on the deviation of the first-period price from its expectation conditioned upon all public information. The screen arises because informed investors in the first period trade more aggressively when they have access to higher quality private information, which creates greater deviations from the expected price-to-earnings relation. As a consequence, the deviation is informative about the quality of private information and new speculators who rely on the screen naturally follow the firm that has the largest deviation from the expected pricing pattern. Hence, the model predicts that firms with greater unexplained pricing patterns are more likely to attract the attention of active investors or speculators, which implies that those firms should exhibit greater price volatility and trading volume in subsequent periods of trade.[3] In summary, we identify a role for price-based screens within a fairly standard market setting and show how they influence private information gathering efforts, which, in turn, have implications for subsequent price volatility and trading volume.

We empirically test the model's predictions using a comprehensive sample of U.S. publicly traded firms' earnings announcements from 2004 to 2022. Earnings announcements offer many advantages for testing the model's predictions. First, they represent a routine public release of fundamental information that corresponds to the public disclosure in the model, and they have a well-established relation with prices. That price-to-earnings relation can then form a basis for identifying when

[3]This intuition is similar to the result in Andrei and Hasler (2015), where an exogenous increase in investor attention increases volatility. Similarly, a firm's risk disclosure in Smith (2022) helps investors estimate when it is more lucrative to acquire information.

prices are more likely being driven by private information, which is what the model suggests will attract investor attention. Second, they occur at regular intervals, which mitigates the alternative explanation that investors are responding to underlying events at other times as opposed to the reflection of those events in earnings. Third, a broad sample of firms announces earnings, allowing us to form screen benchmarks based on a simple price-to-earnings relation. Although the price-to-earnings benchmarks admittedly reflect a very primitive screen, most screens likely rely in part on the use of some sort of price-to-earnings ratio and, as a consequence, we expect our simple empirical screen should be a reasonable proxy for more complex screens that also employ prices. Accordingly, deviations from price-to-earnings benchmarks should predict subsequent investor attention if those more complex screens are being used to allocate constrained information gathering capacity.

Using a sample of 87,493 earnings announcements, we test whether future information acquisition on Edgar increases with larger deviations of price from expected price upon the release of earnings. Consistent with the model's predictions, we find that speculators acquire more information on Edgar in the few months after the earnings announcement when the price at the earnings announcement deviates more from expected price. Similarly, using a sample of 59,895 earnings announcements, we find that speculators acquire more information on Bloomberg when unexplained price deviations are larger.[4] These findings are consistent with speculators using screens based on past prices to allocate their information acquisition efforts to firms where higher quality private information can be acquired.

We also test whether this efficient allocation of information acquisition efforts is a mechanism through which screen outcomes are related to future price volatility and future trading volume. Using a mediated (path) analysis, we test the model's implications for price volatility and trading volume. Larger unexplained price deviations are associated with higher return variance and trading volume, which suggests more informed trade when more speculators are attracted to firms with large price deviations. Consistent with the model's predictions, the mediated (path) analysis finds that information acquisition appears to be an important underlying mechanism for the relation between unexplained price deviations and informed trade.

We conduct several additional tests to assess the robustness of our results and explore their implications. First, we use alternative measures of information acquisition. An advantage of using Edgar search volume as a measure of information acquisition is that we are able to observe the IP

[4]The sample sizes for the Edgar and Bloomberg samples differ due to differences in data availability for the measures of information acquisition.

addresses downloading Edgar filings. We use the IP address information in the log files to identify Edgar downloads originating from new speculators (i.e., those IP addresses that did not download any of the firm's Edgar filings in the previous period). Using alternative measures of information acquisition based on these IP addresses, we find that information acquisition from *new* speculators is positively associated with larger price deviations. This result provides further support for the model's prediction that new speculators use screens to decide which firm to follow, and it alleviates concerns that our information acquisition results can be attributed solely to the persistent following by established speculators. In addition, we use IP addresses to identify Edgar downloads initiated from financial institutions and internet service providers (ISPs), and the results are consistent with both types of investors using screens to manage their information acquisition efforts.

Second, we use an alternative measure of a screen outcome. As noted above, we acknowledge that the exact screen investors use is unobservable. Although the main empirical tests employ a specific screen, investors need not use this specific screen to make their information acquisition allocation decisions. The intuition of the model is that investors use unexplained prices to assess the expected quality of private information that they could acquire, and therefore an association between future information acquisition and other screen outcomes measuring unexplained prices would further corroborate this intuition. Using an alternative screen, which is a changes-based version of our abnormal price screen, we continue to find that Edgar and Bloomberg information acquisition subsequent to earnings announcements increases with the screen outcomes. This result is consistent with the interpretation that investors can use screens that are similar to, and not necessarily exactly the same as, the price-to-earnings screen we study. Although the use of other price-based screens is consistent with our intuition, a limitation inherent to our setting is that we cannot observe whether, instead of using screens directly, investors only use private signals that are correlated with the screen outcomes (e.g., private meetings or data, as in Solomon and Soltes, 2015; Bushee, Gerakos, and Lee, 2018; Zhu, 2019). Although such a possibility exists, we believe the screens we use are simple and readily available, making them a reasonable option for investors to use in deciding where to allocate their attention. Furthermore, even if investors do not use screens, an alternative interpretation of our results is that investors could reasonably use screens to improve their attention allocation. Our screen outcomes predict future information acquisition, which is a measure that is suggestive of high-quality private information to be acquired in that future period. Nevertheless, our results should be interpreted with this caveat in mind.

Third, we show that our results are robust to a shorter information acquisition window, which supports the link between the screen outcome and information acquisition immediately after that screen outcome becomes public. Fourth, we assess the model's implication that attention to one firm comes at the expense of attention to another firm. Within the context of the model, this implication falls from the assumption that speculators follow only one firm. Given this structure, the attention a firm receives cannot come from another source, such as foregone leisure or increased expenditures on increased attention capacity (e.g., adding personnel or technology). In spite of this simplifying assumption in the model, it is still plausible that, at the margin, increases in attention to one firm might, in part, come at the expense of attention paid to another. To assess whether there is empirical evidence consistent with this trade off, we test whether larger unexplained deviations of a *peer* firm's price are associated with reduced information acquisition for another focal firm. To do so, we perform a matched-peer firm analysis, matching each announcement to an announcement on the same day by a peer firm in the same industry and with a comparable asset size. Consistent with the model's predictions, our results imply that information acquisition increases in the focal firm's abnormal price and decreases in the matched-peer firm's abnormal price. These results provide support for the idea that firms compete for investors' limited information acquisition capacity.

Finally, we explore the persistence of the effects we document. We find that screen outcomes have some persistence over time; firms tend to have a similar relative ranking in abnormal price from one period to the next, suggesting a positive feedback loop between abnormal price deviation and private information acquisition. Consistent with the allocation of information acquisition efforts further signaling high quality private information in a future period, we find that screen outcomes also predict information acquisition one and two quarters ahead, although the predictability decreases over time. These exploratory analyses provide suggestive evidence of dynamic predictions related to how investors allocate their attention, which could be a meaningful avenue for future theoretical and empirical research.

Our study contributes to multiple streams of literature. First, our model and empirical results help explain patterns documented in prior empirical research. The notion that institutional investors dedicate more attention to stocks that exhibit greater shocks to returns has gained traction in the empirical literature. For example, Drake et al. (2015) document an empirical pattern of greater information acquisition for firms with negative returns. Consistent with this pattern, Kempf et al. (2017) and Abramova, Core, and Sutherland (2020) assume that institutional investors focus attention on holdings that experience return shocks, and consequently become distracted with respect to other, unrelated, parts of their portfolio. Our study provides a rational explanation for this assumption, by introducing the idea that *unexplained* price movements suggest that there is something valuable to be learned. In doing so, our analysis complements prior empirical findings consistent with information intermediaries efficiently allocating their information processing outputs to satisfy user demand (e.g., Akbas, Markov, Subasi, and Weisbrod, 2018; Blankespoor, deHaan, and Zhu, 2018).

Second, our analysis contributes novel insights to the literature on private information acquisition decisions in asset markets by introducing uncertainty about private information quality. Within that literature, various determinants of private information acquisition decisions have been studied, including the direct cost of acquiring the information (e.g., Grossman and Stiglitz, 1980 or Verrecchia, 1982), the nature of available public information (e.g., Demski and Feltham, 1994; Kim and Verrecchia, 1994; McNichols and Trueman, 1998; Gao and Liang, 2013), the mechanisms for profiting from that information through trading or indirect sale (e.g., Garcia and Vanden, 2009), information processing biases (e.g., Ko and Huang, 2007), status concerns (e.g., Garcia and Strobl, 2011), and the nature of the information choice set and nature of information chosen by other investors (Froot, Scharfstein, and Stein, 1992; Fischer and Verrecchia, 1998; Van Nieuwerburgh and Veldkamp, 2009; Kacperczyk, Van Nieuwerburgh, and Veldkamp, 2016). Our analysis is most related to the latter determinants, in that we consider an information choice from a constrained set of information choices. In that work, the investors generally choose how much information to acquire (i.e., how much variance to eliminate) or the degree to which one's information overlaps with others (i.e., the covariance of the private information). The attributes of the information that investors can obtain, such as its precision, are commonly known in these models and the only uncertainty pertains to the realization of that information. In those settings, past prices are typically not informative about the potential profits to private information acquisition. As a result, screens based on public information and past prices would not be valuable in those settings. This observation holds in our model as well if private information quality is common knowledge. However, when we assume that prospective informed investors are uncertain about the realization *as well as* about the quality or precision of the information, investors use prior prices and earnings announcements as a screen to reduce the prior uncertainty about the value of private information acquisition.

As such, our third contribution is to a stream of literature where investors are uncertain about the precision of other investors' beliefs. For example, in Blume, Easley, and O'Hara (1994) one set of investors does not know the precision of other investors' private information and in Schneider

(2009) investors do not know the precision of the aggregate information in price.[5] Investors in both models know their own demand and information, and can then use the aggregate trading volume to infer information about the other traders' precision. Finally, in Banerjee and Green (2015) rational uninformed investors do not know whether other investors trade on information or on noise. As a result, larger price surprises are associated with higher fundamental uncertainty and, thus, with higher risk premia. This leads to a seemingly stronger reaction to negative news. In our model, all investors that actively trade know all parameters of the model; that is, conditional on following a firm, there is no uncertainty about the information endowment of other investors following the firm, but investors do not know the precision of private information when they decide which firm to follow.[6]

Finally, our focus on screens connects our study to antecedent studies that employ publicly available data to predict future market returns, which naturally creates a form of investment screen. Those studies use such screens to allocate funds to securities to construct superior investment portfolios, as opposed to allocating attention to acquire superior private information. One view of portfolio formation investment screens is that they facilitate the generation of excess risk-adjusted returns, which is consistent with their use in quantitative investment strategies. Nevertheless, the evidence of such excess returns is not overly compelling.[7] Another view of these portfolio investment screens is that they facilitate the formation of portfolios with desired risk characteristics. Anecdotally, this use of investment screens is consistent with the presence of some relatively passive investment funds, such as high dividend yield funds, that implicitly rely on simple screens. In contrast to portfolio formation screens, our theory of information acquisition screens focuses on their value for allocating resources (e.g., limited attention) to private information acquisition in the market for information. Accordingly, we note that the term "screen" refers to an information acquisition screen rather than a portfolio formation screen.

The remainder of paper begins with an introduction of the model and a discussion of its critical

[5] Somewhat related are Hughes and Pae (2004), Penno (1996), and Subramanyam (1996), where a firm discloses information with an unknown precision. In contrast to our study, in this stream of literature all investors have homogeneous information.

[6] Dye and Sridhar (2007) and Michaeli (2017) allow for an unobservable precision of information in a capital market. Dye and Sridhar (2007) assume that all investors receive the information and do not learn the precision. Michaeli (2017) assumes that a manager acquires information and provides it to users for free. That is, investors do not choose to acquire information (whether or not they observe its precision), and they have as much information as the manager chooses to provide.

[7] For example, Green, Hand, and Zhang (2017) simultaneously studies a wide variety of previously documented screening variables, 94 in total, and finds a lesser number of reliably independent variables from that set. Furthermore, they document that the hedge-portfolio returns derived from those independent screening variables are largely nonexistent in more recent years.

assumptions in Section 2. In Section 3, we characterize the equilibrium for the model, describe the intuition for why and how unexplained prices inform information acquisition decisions, and identify some testable empirical implications. Section 4 describes the empirical data and measures used in the study and presents descriptive statistics. In Section 5, we discuss the empirical strategy and present results. Section 6 provides concluding remarks.

2 Model

Consider a setting where the claims to two separate firms, a and b, are traded at two separate periods. Each firm's terminal value is the sum of cash flows at three points in time: (1) before the first-period trade, (2) before the second-period trade, and (3) after the second-period trade. That is, the uncertain terminal value of firm $i \in \{a, b\}$ is $\tilde{v}_i = \tilde{e}_{i1} + \tilde{e}_{i2} + \tilde{e}_{i3}$, where $\tilde{e}_{it} = \tilde{e}_{it-1} + \tilde{\varepsilon}_{it}$, $\tilde{\varepsilon}_{it} \sim N\left(0, s_i^2\right)$, $\tilde{\varepsilon}_{it}$ is independent of $\tilde{\varepsilon}_{j\tau}$ for all $j \in \{a, b\}$, $\tau \in \{1, 2, 3\}$ and $\{j, \tau\} \neq \{i, t\}$, and e_{i0} is normalized to 0.

Figure 1 summarizes the timeline. Period 1 has three stages. At stage 1.1, prior to the first-period trade at stage 1.2, e_{a1} and e_{b1} are disclosed, which we interpret as a period-1 earnings release.[8] At stage 1.3, new speculators arrive and make their firm-following choice. Period 2 has two stages. At stage 2.1, the period-2 earnings release, e_{a2} and e_{b2}, occurs for each firm. Then, the second period of trade occurs at stage 2.2. The terminal value of the firm is the sum of cash flows at period 3.

$t = 1$			$t = 2$		$t = 3$
Stage 1.1	Stage 1.2	Stage 1.3	Stage 2.1	Stage 2.2	
disclosure of e_{i1}, speculators observe x_{i1}	trade at P_{i1}	new speculators choose firm to follow π_{i2}	disclosure of e_{i2}, speculators observe x_{i2}	trade at P_{i2}	terminal cash flows v_i

Figure 1 Timeline. This figure summarizes the timeline of the model.

The markets for the two sets of claims involve speculators, passive investors, and noise traders.

[8] One can interpret this as the public information disclosed in an earnings release. Therefore, in practice this release represents a multitude of variables, but we denote each firm's release with one variable (e.g., earnings) to keep the model analytically simple. If we were to introduce additional dimensions of firm performance, the deviation of observed price from the pricing norm would still convey the same information about the quality of private information. Hence, there is little incremental insight from formally considering such additions to the screen in our model. Furthermore, throughout the paper we denote random variables with a tilde (˜) and their respective realizations without.

At each trading period for firm i's shares: the set of noise traders has measure 1; there are π_{i1} established speculators per noise trader; and there are $m \rightarrow \infty$ passive investors per noise trader. We let $m \rightarrow \infty$ to derive simple, risk-neutral pricing representations, although this assumption is not necessary for the model results. Finally, we normalize the supply of shares per investor to 0; that is, the cumulative holdings of speculators, noise traders, and passive investors have to equal zero in equilibrium.

Speculators receive private information before trading in the assets. That is, concurrent with the earnings release at stage 1.1, each speculator in the market for firm i privately observes the realization $\tilde{x}_{i1} = x_{i1}$, where $\tilde{x}_{i1}$ is normally distributed with mean 0, variance $q_i^2 < s_i^2$, covariance with $\tilde{\varepsilon}_{i2}$ of q_i^2, and is independent of all other random variables in the model. In effect, q_i^2 represents the quality of the speculators' private information. Furthermore, each speculator for firm i privately observes the realization $\tilde{x}_{i2} = x_{i2}$ prior to the second-period trade at stage 2.2, where $\tilde{x}_{i2}$ is normally distributed with mean 0, variance $q_i^2 < s_i^2$, covariance with $\tilde{\varepsilon}_{i3}$ of q_i^2, and is independent of all other random variables in the model.[9] Therefore, we assume that the quality of the speculators' private information is perfectly correlated over the two trading periods. Furthermore, we assume that passive investors in the market for firm i know the quality of the private information obtained by speculators trading in firm i, but not the realization of that information.

Prior to the second-period trade, at stage 1.3, a new set of speculators with a measure of 1 arrive. The new speculators can follow only one of the two firms, and a speculator who follows firm i learns x_{i2} in addition to the public disclosures.[10] We denote the proportion of new speculators who choose to follow firm i with π_{i2}. It follows that, if π_{i2} of the new speculators follow firm i, there are $\pi_{i1} + \pi_{i2}$ speculators per noise trader in the market for firm i.

An established speculator or passive investor in the market for firm i's shares chooses holdings d_{i1} and d_{i2} to maximize the expectation of:

$$d_{i1}\left(P_{i2} - P_{i1}\right) - \frac{c}{2} d_{i1}^2 + d_{i2}\left(v_i - P_{i2}\right) - \frac{c}{2} d_{i2}^2. \tag{1}$$

The term $\frac{c}{2} d_{i1}^2$ ($\frac{c}{2} d_{i2}^2$) reflects some cost of holding a position after period-1 (2) trade to period-2 trade (terminal date), which crudely reflects the cost of being exposed to the risks of holding i over that time frame. From a modeling perspective, the introduction of this cost is a parsimonious

[9]In lieu of this setup, we could instead assume a structure where the signal is the terminal value plus noise, and we would obtain the same results. Our setup is more notationally parsimonious.

[10]That speculators can only follow one of the two firms but can glean information from prices for free is similar to the literature on rationally inattentive investors (e.g., Mondria, 2010 or Kacperczyk et al., 2016).

way to bound demands. Similarly, a new speculator who follows firm i for the second-period trade chooses holdings d_{i2} to maximize the expectation of:

$$d_{i2}\left(v_i - P_{i2}\right) - \frac{c}{2}d_{i2}^2. \tag{2}$$

We assume that noise traders' aggregate demand in the market for firm i in period t is given by $\frac{n_{it}}{c}$, where n_{it} is the realization of the random variable $\tilde{n}_{it}$, which is normally distributed with mean 0 and variance σ_{it}^2, and is independent of all other random variables. We scale n_{it} by c (which effectively scales the variance of the noisy demand by c^2) to simplify the price expression.

The critical, and novel, assumption in our model concerns the knowledge the new speculators possess when they decide which firm to follow. We assume these investors know all of the model primitives except for the values of q_a^2 and q_b^2. Their priors regarding q_a^2 and q_b^2 are that they are independent and identically distributed random variables with two equally likely outcomes, q_h^2 and q_l^2, where $q_h^2 > q_l^2$. In addition, these investors observe the disclosures of e_a and e_b, as well as the market clearing prices from the first-period trade, P_{a1} and P_{b1}. This information is consistent with the kind of data that actual investors could easily access prior to deciding where to focus their information gathering efforts. With this information, they try to assess which firm offers the greatest opportunity for profitable information acquisition.

In our model, we have four main assumptions that warrant some discussion. First, we assume that each investor can only acquire information about one firm. Alternatively, we could assume a cost of information acquisition for each firm. Because a speculator's (gross) expected benefit from following a firm would be unchanged, our results would remain the same under this assumption.

Second, we assume that each investor trades in the shares of only one firm, which is similar to Merton (1987). The nature of the investor objective functions, coupled with an infinitely large number of passive investors, implies that this assumption is of no particular incremental import. In particular, the same equilibrium outcomes arise if all speculators and passive investors are potentially participating in both markets, with speculators and passive investors maximizing the expectation of $\sum_{i\in\{a,b\}} d_{i1}\left(P_{i2} - P_{i1}\right) - \frac{c}{2}d_{i1}^2 + d_{i2}\left(v_i - P_{i2}\right) - \frac{c}{2}d_{i2}^2$ and new speculators maximizing the expectation of $\sum_{i\in\{a,b\}} d_{i2}\left(v_i - P_{i2}\right) - \frac{c}{2}d_{i2}^2$.

Third, we assume that investors face a cost of holding a position for one period, $\frac{c}{2}d_{it}^2$, which guarantees that traders have finite demands. Alternatively, to finitely bound investor demands, we could: (i) assume that the speculators believe that their demand affects price as in Kyle (1985); (ii) assume that investors are risk averse with a traditional negative exponential utility function; or (iii) exogenously bound an investor's demand, for example by assuming that $d_{i1} \in [-1, 1]$. Assumption (i) implies that the speculators in period 1 will strategically change their trading behavior to lower the probability that new speculators enter the market, which drastically complicates the analysis without altering the primary economic force highlighted in our model. Assumption (ii) also significantly complicates the analysis because the prices at the second period of trade are not normally distributed. Specifically, second-period prices are not normally distributed because new speculators follow a firm based on first-period prices and earnings from both firms, which influences the returns distribution for second-period prices. Finally, assumption (iii) effectively assumes a convex cost that is zero within the bounds but infinite for any demand outside of the bounds, which is somewhat analogous to our setup, which effectively assumes a convex but smooth cost function.

Finally, we assume that the quality of private information, q_i^2, is unknown to new speculators but that the variance of noise traders' demand, σ_{it}^2, is known. These two parameters are the two main determinants of speculators' expected profits. The important assumption in our model is that the benefit of information acquisition is unknown and correlated over time. Therefore, we would obtain similar results if we assumed that q_i^2 is known and that σ_{it}^2 is unknown and correlated across periods.

3 Equilibrium Characterization and Empirical Predictions

We first characterize the unique equilibrium for our model and discuss why and how the screen works within that equilibrium. After characterizing the equilibrium, we then discuss the model's testable empirical implications.

3.1 Equilibrium Derivation

An equilibrium is characterized by establishing equilibrium behavior for each of three stages: the first-period trade (stage 1.2), the second-period trade (stage 2.2), and the stage where the new second-period speculators make a decision about which firm to follow (stage 1.3). For the two stages involving trade, we restrict attention to noisy rational expectations equilibria as in Grossman and Stiglitz (1980). That is, we assume that the passive investors learn from price. The equilibrium condition for the firm-following decisions simply requires that no new speculator would alter their firm-following decision given their rational conjecture of the proportion of speculators following each firm.

3.1.1 Second-period Equilibrium Pricing

Assume that the proportion of new speculators who choose to follow firm i is given by the function $\pi_{i2}(\Omega)$,where Ω is the outcomes from the first-period trade $\{e_{a1}, P_{a1}, e_{b1}, P_{b1}\}$. Taking $\pi_i(\Omega)$ as given, we determine the demands of each investor and then establish the second-period price using a market clearing condition. The first-order condition for a firm i established or new speculator's objective function yields an optimal demand for firm i claims of

$$d_{i2S} = \frac{E[\tilde{v}_i | e_{i1}, e_{i2}, x_{i2}] - P_{i2}}{c} = \frac{e_{i1} + 2e_{i2} + x_{i2} - P_{i2}}{c}. \tag{3}$$

Similarly, firm i passive investor demand is

$$d_{i2P} = \frac{E[\tilde{v}_i | e_{i1}, e_{i2}, P_{i1}, P_{i2}] - P_{i2}}{c} = \frac{e_{i1} + 2e_{i2} + E[\tilde{x}_{i2} | e_{i1}, e_{i2}, P_{i1}, P_{i2}] - P_{i2}}{c}. \tag{4}$$

The market clearing condition is

$$(\pi_{i1} + \pi_{i2}(\Omega)) d_{i2S} + m d_{i2P} + \frac{n_{i2}}{c} = 0. \tag{5}$$

Substituting in for the demands and rearranging terms implies that the market clearing price must satisfy:

$$P_{i2} = e_{i1} + 2e_{i2} + \frac{(\pi_{i1} + \pi_{i2}(\Omega)) x_{i2}}{\pi_{i1} + \pi_{i2}(\Omega) + m} + \frac{mE[\tilde{x}_{i2} | e_{i1}, e_{i2}, P_{i1}, P_{i2}]}{1 + \pi_i(\Omega) + m} + \frac{n_{i2}}{1 + \pi_i(\Omega) + m}. \tag{6}$$

In order to complete the characterization of the equilibrium, we must determine $E[\tilde{x}_{i2} | e_{i1}, e_{i2}, P_{i1}, P_{i2}]$. We determine this expectation in the same manner as Grossman and Stiglitz (1980). Specifically, given the relationship between price and demands, knowledge of second-period price allows the passive investors to infer the statistic $y_{i2} = (\pi_{i1} + \pi_{i2}(\Omega)) x_{i2} + n_{i2}$, which is a sufficient statistic for $\{e_{i1}, e_{i2}, P_{i1}, P_{i2}, y_{i2}\}$ with respect to $\tilde{x}_{i2}$, $E[\tilde{x}_{i2} | e_{i1}, e_{i2}, P_{i1}, P_{i2}, y_{i2}] = E[\tilde{x}_{i2} | y_{i2}]$. The statistic y_{i2} is the realization of a random variable with mean 0, variance $(\pi_{i1} + \pi_{i2}(\Omega))^2 q_i^2 + \sigma_{i2}^2$, and covariance with $\tilde{x}_{i2}$ of $(\pi_{i1} + \pi_{i2}(\Omega)) q_i^2$. It follows that

$$E[\tilde{x}_{i2}|y_{i2}] = \frac{(\pi_{i1} + \pi_{i2}(\Omega)) q_i^2}{(\pi_{i1} + \pi_{i2}(\Omega))^2 q_i^2 + \sigma_{i2}^2} ((\pi_{i1} + \pi_{i2}(\Omega)) x_{i2} + n_{i2}). \tag{7}$$

Letting $m \to \infty$ yields the second-period price, which is characterized in Observation 1.

Observation 1. Given any equilibrium $\pi_{i2}(\Omega)$, the second-period price for firm $i \in \{a, b\}$ is uniquely characterized by a function of the form

$$P_{i2} = e_{i1} + 2e_{i2} + \beta_{i2x}(\Omega) x_{i2} + \beta_{i2n}(\Omega) n_{i2}, \tag{8}$$

where $\beta_{i2x}(\Omega) = \frac{(\pi_{i1}+\pi_{i2}(\Omega))^2 q_i^2}{(\pi_{i1}+\pi_{i2}(\Omega))^2 q_i^2+\sigma_{i2}^2}$ *and* $\beta_{i2n}(\Omega) = \frac{(\pi_{i1}+\pi_{i2}(\Omega)) q_i^2}{(\pi_{i1}+\pi_{i2}(\Omega))^2 q_i^2+\sigma_{i2}^2}$.

Note that the second-period price in Observation 1 has the linear structure inherent to noisy rational expectations models with normally distributed random variables. The coefficient on the disclosed value e_{i2} equals two because of the AR-1 cash flow process.

3.1.2 First-period Equilibrium Pricing

We derive the demands from the investors at period 1 analogously to those at period 2. A feature of the model that greatly facilitates the derivations is that the speculators' and passive investors' expectation of $\beta_{i2x}(\Omega) x_{i2} + \beta_{i2n}(\Omega) n_{i2}$ is always 0 regardless of how first-period earnings and prices, Ω, determine $\pi_{i2}(\Omega)$ in equilibrium. A firm i speculator's first-period demand is

$$d_{i1S} = \frac{E\left[\tilde{P}_{i2}|e_{i1}, x_{i1}\right] - P_{i1}}{c} = \frac{3e_{i1} + 2x_{i1} - P_{i1}}{c}. \tag{9}$$

A passive investor's demand is

$$d_{i1P} = \frac{E\left[\tilde{P}_{i2}|e_{i1}, P_{i1}\right] - P_{i1}}{c} = \frac{3e_{i1} + 2E[\tilde{x}_{i1}|e_{i1}, P_{i1}] - P_{i1}}{c}. \tag{10}$$

The first-period market-clearing condition, requires that $\pi_{i1} d_{i1S} + m d_{i1P} + \frac{n_{i1}}{c} = 0$. Substituting demand orders and rearranging terms yields an equilibrium price of

$$P_{i1} = 3e_{i1} + \frac{\pi_{i1}}{\pi_{i1} + m} 2x_{i1} + \frac{m}{\pi_{i1} + m} 2E[\tilde{x}_{i1}|e_{i1}, P_{i1}] + \frac{1}{\pi_{i1} + m} n_{i1}. \tag{11}$$

To determine $E\left[\tilde{x}_{i1}|e_{i1}, P_{i1}\right]$, observe that the first-period price allows the passive investors to infer the statistic $y_{i1} = 2\pi_{i1}x_{i1} + n_{i1}$, and y_{i1} is a sufficient statistic for $\{e_{i1}, P_{i1}, y_{i1}\}$ with respect to $\tilde{x}_{i1}$, $E\left[\tilde{x}_{i1}|e_{i1}, P_{i1}, y_{i1}\right] = E\left[\tilde{x}_{i1}|y_{i1}\right]$. It follows that

$$E\left[\tilde{x}_{i1}|e_{i1}, P_{i1}\right] = \frac{2\pi_{i1}q_i^2}{4\pi_{i1}^2 q_i^2 + \sigma_{i1}^2}\left(2\pi_{i1}x_{i1} + n_{i1}\right). \tag{12}$$

Letting $m \to \infty$ yields the first-period price, which is characterized in Observation 2, and which has the standard structure.

Observation 2. In any equilibrium the first-period price for firm $i \in \{a, b\}$ is uniquely characterized by a function of the form

$$P_{i1} = 3e_{i1} + \beta_{i1x}x_{i1} + \beta_{i1n}n_{i1}, \tag{13}$$

where $\beta_{i1x} = \frac{8\pi_{i1}^2 q_i^2}{4\pi_{i1}^2 q_i^2 + \sigma_{i1}^2}$ and $\beta_{i1n} = \frac{4\pi_{i1} q_i^2}{4\pi_{i1}^2 q_i^2 + \sigma_{i1}^2}$.

3.1.3 Expected Profits from Following Firm i

To facilitate the characterization of equilibrium, it is useful to compute the second-period expected profits for a new speculator who follows firm i conditional upon the first-period statistics available for making the firm following decision, Ω, and an equilibrium $\pi_{i2}(\Omega)$. Given a second-period pricing function of the form $P_{i2} = e_{i1} + 2e_{i2} + \beta_{i2x}(\Omega)\, x_{i2} + \beta_{i2n}(\Omega)\, n_{i2}$, which is characterized in Observation 1, a speculator who observes x_{i2} and experiences a price determined by n_{i2} has expected payoffs

$$d_{i2}\left(E\left[\tilde{v}_i|e_{i1}, e_{i2}, x_{i2}\right] - P_{i2}\right) - \frac{c}{2}d_{i2}^2 = \frac{1}{2c}\left(\left(1 - \beta_{i2x}(\Omega)\right)x_{i2} - \beta_{i2n}(\Omega)\, n_{i2}\right)^2. \tag{14}$$

It follows that the expected payoffs prior to observing x_{i2} and n_{i2}, but with knowledge of q_i^2, σ_{i2}^2, π_{i1} and $\pi_{i2}(\Omega)$, are

$$\Pi\left(q_i^2, \pi_{i2}(\Omega)\right) = \frac{1}{2c}\frac{q_i^2\sigma_{i2}^2}{\left(\pi_{i1} + \pi_{i2}(\Omega)\right)^2 q_i^2 + \sigma_{i2}^2}. \tag{15}$$

The expected payoffs have intuitive properties in that they are increasing in the quality of the speculators' private information, $\frac{\partial \Pi}{\partial q_i^2} > 0$, and in the extent of noise trade, $\frac{\partial \Pi}{\partial \sigma_{i2}^2} > 0$, which serves to obfuscate the informed trading activity in price. Note that the quality of private information and the extent of noise trade structurally enter the same way in a speculator's expected payoffs from information acquisition. This suggests that uncertainty about the extent of noise trade, coupled with a positively correlated extent of noise trade over time, would also make screens based on price and earnings valuable to speculators and have similar comparative statics. Finally, note that expected payoffs are decreasing in the proportion of new speculators who follow firm i, $\frac{\partial \Pi}{\partial \pi_{i2}(\Omega)} < 0$, which is useful for establishing a unique equilibrium.

When new speculators arrive, they do not know q_a^2 and q_b^2. Hence, they must assess their expected payoffs from following each firm given the probability that the quality of private information is high or low. These expected payoffs are

$$E\left[\Pi\left(q_i^2, \pi_{i2}(\Omega)\right)|\Omega\right] = \Pi\left(q_l^2, \pi_{i2}(\Omega)\right) + \Pr\left(q_i^2 = q_h^2|\Omega\right)\left(\Pi\left(q_h^2, \pi_{i2}(\Omega)\right) - \Pi\left(q_l^2, \pi_{i2}(\Omega)\right)\right) \quad (16)$$

where, for $r \in \{h, l\}$, $\Pr\left(q_i^2 = q_r^2|\Omega\right)$ is the probability $q_i^2 = q_r^2$ conditional on Ω. Not surprisingly, the expected profits are increasing in the probability that the quality of private information is high.

Characterizing $\Pr\left(q_i^2 = q_h^2|\Omega\right)$ is necessary to assess how the expected payoffs from following firm i, $E\left[\Pi\left(q_i^2, \pi_{i2}(\Omega)\right)|\Omega\right]$, are affected by the equilibrium $\pi_{i2}(\Omega)$. As derived in Appendix A, the probability that $q_i^2 = q_h^2$ is

$$\Pr\left(q_i^2 = q_h^2|\Omega\right) = \frac{f\left(P_{i1}|e_{i1}, q_h^2\right)}{f\left(P_{i1}|e_{i1}, q_h^2\right) + f\left(P_{i1}|e_{i1}, q_l^2\right)} = \frac{1}{1 + \sqrt{\frac{V_h^2}{V_l^2}} \exp\left[-\left(P_{i1} - 3e_{i1}\right)^2 \frac{V_h^2 - V_l^2}{2V_h^2 V_l^2}\right]}, \quad (17)$$

where $f\left(P_{i1}|e_{i1}, q_i^2\right)$ is the probability density function for P_{i1} conditional upon e_{i1} and $q_i^2 \in \left\{q_h^2, q_l^2\right\}$, a normally distributed random variable with mean $3e_{i1}$ and variance $V_i^2 = \frac{16\pi_{i1}^2 q_i^2}{4\pi_{i1}^2 q_i^2 + \sigma_{i1}^2} q_i^2$.

The mathematical characterization of the probability that $q_i^2 = q_h^2$, $\Pr\left(q_i^2 = q_h^2|\Omega\right)$, implies that a greater deviation of the realized first-period price, P_{i1}, from the price that would be expected to prevail given the public data employed for the screen, $3e_{i1}$, signals that higher quality private information is more likely to be obtained from following firm i. The deviation provides the signal of private information quality because higher quality private information induces more informed trade, which causes the equilibrium price to deviate from expectation. However, the deviation is not a perfect signal, because the deviation can also be attributable to noise trade. Within the context of our model, then, the screen conveys information about the returns to speculative information gathering by conveying information about the uncertain quality of private information within a firm's operating environment.

3.1.4 Equilibrium Characterization

Observation 1 characterizes the unique equilibrium second-period pricing functions given an equilibrium $\pi_{a2}(\Omega)$ and $\pi_{b2}(\Omega)$, and Observation 2 characterizes the unique first-period pricing functions, which are not determined by $\pi_{a2}(\Omega)$ and $\pi_{b2}(\Omega)$. We complete the characterization of equilibrium by showing that there is a unique $\pi_{a2}(\Omega)$ and $\pi_{b2}(\Omega) = 1-\pi_{a2}(\Omega)$, $\pi^*_{a2}(\Omega)$ and $\pi^*_{b2}(\Omega) = 1-\pi^*_{a2}(\Omega)$, such that no new speculators can strictly increase the expected profits by changing their firm following decision.

Formally, $\pi^*_{a2}(\Omega)$ and $\pi^*_{b2}(\Omega) = 1 - \pi^*_{a2}(\Omega)$ must satisfy

$$E\left[\Pi\left(q_a^2, \pi^*_{a2}(\Omega)\right)|\Omega\right] = E\left[\Pi\left(q_b^2, \pi^*_{b2}(\Omega) = 1 - \pi^*_{a2}(\Omega)\right)|\Omega\right], \tag{18}$$

if $\pi^*_{a2}(\Omega) \in (0,1)$ and $\pi^*_{b2}(\Omega) \in (0,1)$. If $\pi^*_{a2}(\Omega) = 1$ and $\pi^*_{b2}(\Omega) = 0$, the equal sign in equation (18) is replaced by a greater than or equal sign, and if $\pi^*_{a2}(\Omega) = 0$ and $\pi^*_{b2}(\Omega) = 1$, the equal sign is replaced by a less than or equal sign. Because $E\left[\Pi\left(q_a^2, \pi^*_{a2}(\Omega)\right)|\Omega\right]$ is decreasing in $\pi^*_{a2}(\Omega)$ and $E\left[\Pi\left(q_b^2, \pi^*_{b2}(\Omega) = 1 - \pi^*_{a2}(\Omega)\right)|\Omega\right]$ is increasing in $\pi^*_{a2}(\Omega)$, there is a unique equilibrium $\pi^*_{a2}(\Omega)$ and $\pi^*_{b2}(\Omega) = 1 - \pi^*_{a2}(\Omega)$. Proposition 1 naturally follows.

Proposition 1. *There exists a unique equilibrium characterized by the pricing functions in Observations 1 and 2, and an allocation of new speculators for the second-period trade that is determined by a screen tied to realized first-period prices and earnings.*

3.2 Empirical Implications

Our model provides three testable empirical implications. First, the equilibrium characterization itself implies that investor attention, specifically new speculator following, will be increasing in the screen realization, which is the difference between the observed price and the expected price conditional upon the screening variables.

Corollary 1. *In equilibrium, the number of informed speculators following firm i weakly increases in the deviation of firm i's first-period price from the expected price conditioned on earnings, $|P_{i1} - 3e_{i1}|$, and weakly decreases in the deviation of firm j's first-period price from the expected price conditioned on earnings, $|P_{j1} - 3e_{j1}|$, where $i,j \in \{a, b\}$ and $j \neq i$.*

Corollary 1 suggests that firms whose prices deviate from valuation norms to a relatively larger degree should naturally attract more attention from speculative investors who rely on screens to determine which firms to follow. Within the context of our model, the valuation norm is a simple multiple of earnings, three times earnings. Deviations from that norm suggest to potential informed investors that there is more private information being impounded into the price. However, the deviation from the pricing norm could also be due to noise trade, so the presence of significant private information is not guaranteed.

As an aside, we emphasize that the deviation does not suggest whether the private information is good or bad news relative to what uninformed market participants (i.e., the passive investors) believe because the equilibrium price reflects a correct expectation given all of the public information. That is, the deviation itself does not suggest a trading opportunity. Instead, it signals an information acquisition opportunity. Hence, the deviation is informative in the market for information even though it is not informative in the market for cash flows. This further implies that when there is no uncertainty about the quality of private information (i.e., when $q_h = q_l$), then new speculators will not use past prices at all.

Corollary 1 establishes a link between unexplained pricing and informed speculative activity, which in turn, should influence other observable market characteristics. Hence, the model predicts a relation between unexplained price and those characteristics. The first characteristic we consider is price variance.

Corollary 2. *In the unique equilibrium, an increase in the deviation of firm i's first-period price from the expected price conditioned on earnings,* $|P_{i1} - 3e_{i1}|$*, is associated with an increase in the second-period price variance for firm i's claims and a decrease in second-period price variance for firm j's claims.*

The intuition underlying Corollary 2 is quite straightforward: a larger realization for $|P_{i1} - 3e_{i1}|$ attracts more informed speculators to the market for firm i claims and away from the market for firm j claims. As a consequence there is more (less) informed trade for firm i (j) claims, which leads to more (less) movement in prices.

The second characteristic we consider is trading volume.

Corollary 3. *In the unique equilibrium, an increase in the deviation of firm i's first-period price from the expected price conditioned on earnings,* $|P_{i1} - 3e_{i1}|$*, is associated with an increase in the second-period new speculator trading volume for firm i's claims and a decrease in second-period new speculator trading volume in firm j's claims.*

The intuition underlying Corollary 3 is identical to that for Corollary 2. That is, a larger unexplained price attracts more informed speculators to the market, which increases trading volume from those new informed speculators.

4 Data

4.1 Sample

We use a comprehensive sample of U.S. publicly traded firms' earnings announcements from 2004 to 2022 to test Corollaries 1 to 3.

Our sample of earnings announcements comes from the intersection of Compustat and IBES. We use the next trading day if the announcement is after trading hours and use the earlier of the Compustat and IBES announcement dates if they differ, following DellaVigna and Pollet (2009). We also use Compustat to obtain control variables such as financial leverage, total assets, and short interest. We also obtain management guidance information from IBES.

We download Edgar search volume data from the SEC's Edgar website and remove robot downloads, following Drake et al. (2015).[11] We download information on the intensity of Bloomberg user attention to firms from Bloomberg terminals. We obtain stock price and trading volume data from CRSP. Institutional ownership information comes from Thomson Reuters. Information on voluntary disclosures comes from company filing data obtained from WRDS SEC Analytics Suite. We obtain data on media stories from RavenPack. Our primary sample consists of 87,493 firm-quarters when examining Edgar search volume, due to the Edgar search volume coverage from 2004 to 2016.[12] Our primary sample when examining Bloomberg attention consists of 59,895 firm-quarters, due to the Bloomberg data coverage from 2010 to 2022.

4.2 Measures and Descriptive Statistics

We empirically test whether deviations of price from expected price (e.g., based on earnings realizations) are related to investors' subsequent information acquisition activities. If investors are using these types of screens to allocate their information acquisition efforts, we would observe that investors acquire more information when a firm has large positive or negative price deviations, because such deviations suggest the availability of high quality private information. Although we expect investors in practice to base expected price on a multitude of variables, given that P/E ratios are simple and commonly available, our empirical tests base these expectations on P/E ratios. Our measure of abnormal price, *AbnPrice*, captures deviations of price from the expected price, measured using a benchmark P/E ratio. Specifically, abnormal price is the absolute value of the

[11] Raw data of Edgar filing downloads is available at https://www.sec.gov/about/data/edgar-log-file-data-sets. Ryans (2017) provides a processed version of the log file that removes robot downloads.

[12] We begin our sample in 2004, because Edgar downloads are sparsely populated in 2003.

difference between the firm's P/E ratio at the earnings announcement and the median P/E ratio, measured four quarters prior, of firms in the same industry and size quintile, scaled by this median P/E ratio.[13]

Our tests focus on two measures of information acquisition, one based on the SEC Edgar downloads and the other based on Bloomberg terminal attention. These measures relate to investors' acquisition of information about firms' activities and fundamental performance. In addition, both measures have the advantage of capturing information acquisition from investors with some degree of sophistication. Users downloading company filings from Edgar are sophisticated enough to be aware of and read financial statements, and Bloomberg terminals are primarily used by institutional investors. These measures are available for different sample periods and they encompass information acquisition from different sources, enhancing the robustness of our findings through triangulation. We measure Edgar search volume and Bloomberg abnormal institutional investor attention over the quarter, beginning after the earnings announcement date and ending before the next earnings announcement date. This window maps directly to the second period in the model. In additional analyses, we assess the robustness of our results to using a shorter window of five days, days [+1,+5] after the earnings announcement.

Table 1 reports summary statistics for the variables used in our analysis. The unit of analysis is a firm-quarter. Panel A reports statistics for the 2004-2016 sample used in the Edgar analysis. Panel B reports statistics for the 2010-2022 sample used in the Bloomberg analysis. All variables are defined in Appendix B.

The summary statistics in Panel A reveal that the average firm-quarter has 2,025.11 Edgar downloads, with an average of 183.63 of those downloads in days [+1,+5] after the earnings announcement. The summary statistics in Panel B reveal that the average firm-quarter has a sum of 21.50 in Bloomberg readership, which is the sum of the Bloomberg heat index ranging from 0 to 4 on each firm-day. The average Bloomberg readership sum over days [+1,+5] after the earnings announcement is 1.54.

[13] A disadvantage of this measure is that it is difficult to interpret a P/E ratio when earnings is negative. Therefore, we calculate $AbnPrice$ only for those observations where earnings is positive. In robustness tests, we use an alternative measure that does not require us to drop firm-quarters with negative earnings.

5 Empirical Results

5.1 Main Results: Abnormal Price and Information Acquisition

We begin by testing Corollary 1. Specifically, we test whether larger deviations of price from expected price attract greater information acquisition from speculative investors. We estimate the following regression using the sample of firm-quarters with available data to compute each information acquisition proxy:

$$InfoAcq_{i,t+1} = \beta_1 AbnPrice_{i,t} + \gamma Controls_{i,t} + \Sigma\beta_i Firm_i + \Sigma\beta_t Year\text{-}Quarter_t + \epsilon_{i,t+1}, \tag{19}$$

where $InfoAcq_{i,t+1}$ is one of several measures of information acquisition for firm i in a particular period $t+1$ following the earnings announcement. $AbnPrice_{i,t}$ is the within-quarter percentile rank of the abnormal price, calculated as the absolute difference between firm i's P/E ratio at the earnings announcement and the median P/E ratio, measured four quarters prior, of firms in the same industry and size quintile, scaled by the median P/E ratio. Ranking this variable mitigates the impact of outliers and, compared to using the raw values, better reflects the model's setup.[14] A positive β_1 suggests that larger deviations from expected price attract greater information acquisition from speculative investors using information acquisition screens to allocate their efforts.

Our specification includes a vector of control variables, $Controls_{i,t}$, including financial leverage, firm size, the presence of a bundled management forecast, institutional ownership, short interest, voluntary 8-K disclosures, and media stories. The inclusion of these control variables accounts for variation in information acquisition absent the use of price-based screens. For example, $log(Assets)$, $BundledForecast$, $Voldisc8k$, and $MediaStories$ control for the general information environment or the availability of information about the firm. We include $Leverage$, $Instown$, and $Shortsell$ to control for the type and sophistication of investors that acquire information about the firm (i.e., creditors, institutional investors, and short sellers, following Drake et al., 2015).[15] We also

[14] Specifically, in the model, a speculative investor's capacity to acquire information is limited to only one of two firms. Therefore, the relative ranking of the firm's price deviation matters to investors (e.g., if investors only have enough capacity to acquire information about the top X firms, whether or not a firm is in the set of top X firms matters more than the difference in raw price deviation between the firm with rank X and the firm with rank $X+1$).

[15] As each of these control variables affects the supply of and demand for information, investors potentially use the control variables as inputs to their information acquisition screens. Our intuition from the model is that a simple screen based on stock price (i.e., greater deviations from expected price reflect more aggressive trading from speculators) can help investors assess the quality of private information that can be attained. Therefore, by controlling for non-price-based variables, our empirical tests measure the relation between future information acquisition and price deviations that is incremental to other information environment indicators. Nevertheless, for completeness, we present results with and without the inclusion of these control variables.

include Firm and Year-Quarter fixed effects, which subsume any variation constant for each firm and each quarter. In some specifications, we control for lagged information acquisition and replace Firm fixed effects with Industry fixed effects.[16] Firm fixed effects account for firm-specific variation in information acquisition (e.g., in established speculators) and the control for lagged information acquisition directly accounts for established speculators' information acquisition from the previous period; thus, our empirical specifications map directly to the model's prediction that *new* speculators use screens to rationally allocate their information acquisition efforts.[17] We also include Year-Quarter fixed effects to account for potential time trends in the nature of information in the market and noise trade properties (e.g., due to technological or regulatory changes, or other macroeconomic shocks). We cluster standard errors by Firm, which allows for correlation across time for a given firm.

Table 2 presents our primary regression results estimating Equation (19). Across all specifications, we find support for the prediction in Corollary 1. Information acquisition is greater when $AbnPrice$ reflects larger deviations from expected price. The results in Table 2 imply that an increase in $AbnPrice$ is associated with more Edgar downloads (columns 1 to 3) and more Bloomberg news attention (columns 4 to 6) in the subsequent quarter. Collectively, our empirical findings suggest that new speculators rely on abnormal price deviations as screens to allocate their information acquisition efforts, consistent with the model's intuition that such deviations imply the existence of high-quality private information.

5.2 Main Results: Abnormal Price and Market Outcomes

An important outcome of the model is that, as larger price deviations attract more information acquisition, such information acquisition generates more trade. Thus, Corollary 2 (Corollary 3) posits that subsequent price variance (trading volume) increases with larger price deviations.

We empirically test Corollaries 2 and 3, including the mechanism underlying their predictions, by performing a mediated (path) analysis. Specifically, we test whether information acquisition is an important mechanism underlying the association between larger price deviations and increased

[16]Industries are defined using 2-digit SIC codes.

[17]Price deviations and future information acquisition could be positively correlated, absent the use of screens, because of persistence in information acquisition from established speculators. Therefore, removing variation in information acquisition from established speculators addresses this alternative explanation. Firm fixed effects and the control for the lagged dependent variable both accomplish this goal, but we do not implement both in the same specification because they require different identifying econometric assumptions. Angrist and Pischke (2009) suggest assessing the robustness of the findings to multiple identifying assumptions and using them to bound the effect size.

trading volume and return variance (MacKinnon, 2012).[18] We first estimate the association between $AbnPrice$ and future market outcomes as follows:

$$MarketOutcome_{i,t+1} = \beta_1 AbnPrice_{i,t} + \\ \gamma Controls_{i,t} + \Sigma\beta_i Firm_i + \Sigma\beta_t Year\text{-}Quarter_t + \epsilon_{i,t+1}, \quad (20)$$

where $MarketOutcome_{i,t+1}$ is either trading volume ($TradingVol$) or the standard deviation of returns ($RetVariance$), measured over the subsequent quarter. Then, to estimate the importance of information acquisition as the mechanism for this association, we include information acquisition in the regression:

$$MarketOutcome_{i,t+1} = \beta_1 AbnPrice_{i,t} + \beta_2 InfoAcq_{i,t+1} + \\ \gamma Controls_{i,t} + \Sigma\beta_i Firm_i + \Sigma\beta_t Year\text{-}Quarter_t + \epsilon_{i,t+1}, \quad (21)$$

where $InfoAcq_{i,t+1}$ is either Edgar downloads or Bloomberg investor attention, measured over the subsequent quarter.

Table 3 reports the results of the path analysis. Panels A and B use Edgar downloads as the information acquisition proxy and mediating variable. Panel A reports the indirect and total effects, and Panel B reports the underlying regressions for the effects. The results in Table 3 Panel A, columns 1 and 2, suggest that the direct path (path I) between abnormal price and subsequent trading volume is positive but insignificant after accounting for the mediator variable $EdgarSearch$ (see also Panel B column 2). The mediated path has two components, the path between abnormal price and information acquisition (path II) and the path between information acquisition and trading volume (path III). We report the underlying regression for Path II in Table 2 column 2 and the underlying regression for path III in Table 3 Panel B column 2. The product of these two paths is the indirect effect (0.008), which is more than half of the total effect of 0.014. Note that the total effect is the sum of the direct and indirect effects. Overall, the estimation results support our inference that information acquisition on Edgar is an important underlying mechanism for the positive association between abnormal price deviations and subsequent trading volume.

The results in Table 3 Panel A, columns 3 and 4, suggest that the direct path (path I) between

[18] Prior literature in accounting has used path analysis to formally test whether a relationship between X and Y arises through path Z. For example, Bonsall IV, Green, and Muller III (2018) uses path analysis to study whether business press coverage is an important mechanism through which credit rating agencies increase ratings stringency.

abnormal price and subsequent return variance is positive and significant after accounting for the mediator variable *EdgarSearch* (see also Panel B column 4). The mediated path has two components, the path between abnormal price and information acquisition (path II) and the path between information acquisition and return variance (path III). We report the underlying regression for path II in Table 2 column 2 and the underlying regression for path III in Table 3 Panel B column 4. The product of these two paths is the indirect effect (0.0001), which is 3.3% of the total effect of 0.003. Overall, the findings support our inference that information acquisition on Edgar is also an important underlying mechanism for the positive association between abnormal price deviations and subsequent return variance.

Panels C and D use Bloomberg attention as the information acquisition proxy and mediating variable to find similar results. Panel C reports the indirect and total effects, and Panel D reports the underlying regressions for the effects. The results in Table 3 Panel C, columns 1 and 2, suggest that the direct path (path I) between abnormal price and subsequent trading volume is positive and significant after accounting for the mediator variable *Bloomberg* (see also Panel D column 2). The mediated path has two components, the path between abnormal price and information acquisition (path II) and the path between information acquisition and trading volume (path III). [19] The product of these two paths is the indirect effect (0.004), which is 13.8% of the total effect of 0.029. This finding suggests that information acquisition on Bloomberg is an important underlying mechanism for the positive association between abnormal price deviations and subsequent trading volume.

The results in Table 3 Panel C, columns 3 and 4, suggest that the direct path (path I) between abnormal price and subsequent return variance is positive and significant after accounting for the mediator variable *Bloomberg* (see also Panel D column 4). The mediated path has two components, the path between abnormal price and information acquisition (path II) and the path between information acquisition and return variance (path III). [20] The product of these two paths is the indirect effect (0.0001), which is 10% of the total effect of 0.001. This finding suggests that information acquisition on Bloomberg is also an important underlying mechanism for the positive association between abnormal price deviations and subsequent return variance.

[19] We report the underlying regression for path II in Table 2 column 5 and the underlying regression for path III in Table 3 Panel D column 2.

[20] We report the underlying regression for path II in Table 2 column 5 and the underlying regression for path III in Table 3 Panel D column 4.

5.3 Robustness Tests and Exploratory Analysis

Our main results show that information acquisition is greater when firms' prices deviate more from expected prices and that such information acquisition is an important mechanism underlying the association between price deviations and subsequent market outcomes. In this section, we present several additional analyses to assess the robustness of our results and provide exploratory insights for future research.

First, we assess the robustness of our results to alternative measures of information acquisition. Our specifications in Section 5.1 include Firm fixed effects or a control for lagged information acquisition to control for established speculators' information acquisition. To further isolate changes in information acquisition at the extensive margin, we estimate a version of Equation (19) that investigates whether information acquisition by new speculators increases with abnormal price deviations. Our measures of $InfoAcq$ in these tests are $EdgarNew$ and $EdgarNewIPs$, the number of Edgar downloads from new IP addresses and the number of unique new IP addresses, respectively, summed over the subsequent quarter. New IP addresses are those that did not download the firm's filings in the previous period. Columns 1 and 2 of Table 4 report results showing that greater price deviations attract information acquisition from new speculators, consistent with the model's intuition that new speculators use screens to allocate their information acquisition efforts.

To distinguish between downloads from financial institutions and those from non-professional investors acquiring information on Edgar, we also refine our measures of Edgar downloads by identifying the organizations to which the IP addresses downloading filings are registered. We use WhoWas reports from the American Registry of Internet Numbers (ARIN) to identify the names of the organizations associated with the top 1,500 IP address blocks downloading Edgar filings during our sample period. We identify organizations owning a block of IP addresses, because the last three digits of the IP address are masked. We manually identify financial institutions based on the list of organization names matched to these IP address blocks.[21] Following Drake et al. (2020), we also identify internet service providers (ISPs), such as Comcast and Verizon, that could reflect non-professional investors or professional investors working outside of the office. We then restrict our Edgar downloads variables to the subset of downloads originating from IP addresses identified as financial institutions and the subset of downloads originating from IP addresses owned by ISPs. We expect both subsets to increase with larger price deviations, as both professional

[21] Following Drake, Johnson, Roulstone, and Thornock (2020), these financial institutions include investment banks, hedge funds, commercial banks, insurance companies, and other financial institutions.

and non-professional investors can use screens to allocate their attention. In addition, even the non-professional investors are relatively "sophisticated," because they actively acquire information on Edgar. Columns 3 and 4 of Table 4 present results. We find that abnormal price deviations are associated with more information acquisition on Edgar from financial institutions (column 3) and with more information acquisition on Edgar from non-professional investors (column 4). These results provide further support for the model's prediction that new speculators (e.g., sophisticated investors allocating their attention) use screens to choose which firms to follow. Furthermore, these results are consistent with prior literature's findings that both types of investors' information acquisition activities are associated with other firm outcomes (e.g., Kim, 2023).

Second, we present a robustness test addressing concerns that investors may use an alternative screen. It is important to note that the use of alternative screens is not an alternative explanation for our results. Although our main empirical tests employ a specific screen based on P/E ratios, investors need not use this specific screen to rationally allocate their attention to exhibit behavior consistent with the model. The intuition of the model only suggests that they use a screen, based on past prices, that should reveal the expected quality of private information that could be gleaned.

To this end, we assess whether our results are robust to an alternative price-based screen. An advantage of our main proxy for abnormal price, *AbnPrice*, is that it is a simple proxy for the extent of deviation of a firm's price from what an investor expects the price to be, based on the firm's realized earnings and a benchmark P/E ratio.[22] Our alternative proxy for large price deviations is the changes version of *AbnPrice*, *AbnPriceChange*, which uses the change in price (e.g., 3-day return) over the short window surrounding the earnings announcement, scaled by the unexpected earnings. This variable effectively calculates a short-window earnings response coefficient on each earnings announcement day using the pre-announcement expectations as the benchmark. Measuring returns in this short window around the release of earnings information has the advantage of controlling for potential correlated omitted variables that stay constant in this short window. In addition, we compare the 3-day return per unit of unexpected earnings to the firm's historical average of this value, which means that this screen does not require an assumption about the set of comparable firms investors use to form price expectations.[23]

[22] The benchmark P/E ratio is computed using comparable firms, those in the same industry and size profile.

[23] An additional advantage of this measure is that, unlike *AbnPrice*, it is available for firm-quarters with negative earnings. Therefore, in our regressions using *AbnPriceChange*, we include an additional control for loss quarters. A disadvantage of this alternative measure is that it is difficult to interpret the short-window earnings response coefficient when earnings surprise (UE) and abnormal returns (CAR) move in opposite directions, and that such cases cannot be directly compared to the majority of the cases where the two move in the same direction. Therefore, we calculate this variable only for those observations where UE and CAR have the same sign.

We report the results of estimating Equation (19) after replacing $AbnPrice$ with $AbnPriceChange$ in Table 5. Across all specifications, we find support for the prediction in Corollary 1 using this alternative measure. Information acquisition is greater when $AbnPriceChange$ reflects larger unexplained price movements. The results in Table 5 imply an increase in $AbnPriceChange$ is associated with more Edgar downloads (columns 1 to 3) and more Bloomberg news attention (columns 4 to 6) in the subsequent quarter.

Third, we assess the robustness of our results to measuring information acquisition using a shorter window of five days, days [+1,+5] after the earnings announcement. An advantage of using this shorter window is that information acquisition immediately after the earnings announcement is more likely to be motivated by abnormal price at the earnings announcement, thus mitigating confounds. In Table 6, we report the results of estimating Equation (19) after replacing the dependent variables with Edgar search and Bloomberg news attention measured in the five days after the earnings announcement. Across all specifications, we find support for the prediction in Corollary 1 using this alternative measurement window. Information acquisition in a short window after the announcement is greater when $AbnPrice$ reflects larger unexplained price movements. The results in Table 6 imply an increase in $AbnPrice$ is associated with more Edgar downloads (columns 1 to 3) and more Bloomberg news attention (columns 4 to 6) in days [+1,+5] after the earnings announcement.

Fourth, we corroborate the intuition of the model by showing that abnormal prices elsewhere in the market draw attention away from the focal firm. When interpreted broadly in the context of a multi-firm setting, Corollary 1 not only predicts that information acquisition about firm i is greater when its $AbnPrice$ reflects larger deviations from its expected price, but also that it is smaller when a peer firm j's $AbnPrice$ reflects larger deviations from firm j's expected price. In effect, these two firms compete for investors' limited information acquisition capacity.[24]

To test this prediction, we conduct a matched-peer analysis, in which we match each firm with a peer firm from the same industry that announces earnings on the same day as the focal firm and is closest in asset size. To ensure that the two firms are comparable in terms of asset size, we retain matches where the larger firm's asset size is less than 1.5 times the smaller firm's, and the size

[24] Our model only includes two firms, and it does not account for the possibility that deviations in a peer firm's price can draw additional attention to the focal firm, if investors acquire information about comparable firms. Although outside of the model, we acknowledge that this possibility makes it unclear ex-ante whether the empirical tests will find a negative association between a peer firm's abnormal price and the focal firm's information acquisition. We use Industry fixed effects to address industry-specific information acquisition and Year-Quarter fixed effects to address the potential impact of abnormal price on market-wide information acquisition.

difference between the two firms is less than \$100 billion. When there are multiple such matches, we retain the matched peer with the smallest difference in asset size. This process results in a significantly smaller sample than in Table 2, consisting of 45,619 firm-quarters in the Edgar sample and 29,597 firm-quarters in the Bloomberg sample. Using this sample, we estimate a version of Equation (19) that includes the abnormal price of the peer firm ($AbnPricePeer$) as an additional explanatory variable.[25] Table 7 reports the estimation results. Across all specifications, we find results consistent with Corollary 1's prediction that a larger $AbnPricePeer$ is associated with reduced information acquisition about the focal firm. Table 7 finds that $AbnPrice$ and $AbnPricePeer$ have opposite associations with Edgar downloads (columns 1 to 3) and Bloomberg news attention (columns 4 to 6) in the subsequent quarter. In sharp contrast to the results for the focal firm's abnormal price, $AbnPricePeer$ is associated with *fewer* Edgar downloads about the focal firm. Similarly, $AbnPricePeer$ is associated with *reduced* Bloomberg attention about the focal firm. Combined with the mediated (path) analysis results from Table 3, the results in this table also imply that subsequent price variance and trading volume are increasing in the focal firm's abnormal price and decreasing in the matched-peer firm's abnormal price through their effects on information acquisition.

Finally, we also conduct exploratory analyses to assess the persistence of the effects we document. Our intuition is that, as speculators rely on screens to identify firms with high quality private information, their information acquisition efforts for these firms further increase price deviations in the subsequent period. Thus, it is possible that the extent of price deviation from expected price exhibits some persistence. We investigate this possibility by presenting a rank transition matrix in Table 8 Panel A. The matrix shows that firms with $AbnPrice$ decile rank X in one period (row) are more likely to have the same or similar $AbnPrice$ decile rank in the next period (column) than they are to move several decile rankings up or down.

This persistence in the relative ranking of abnormal price deviations over time also suggests that current period price deviations could predict information acquisition for multiple subsequent periods. We investigate this possibility by estimating versions of Equation (19) that replace $InfoAcq_{i,t+1}$ with $InfoAcq_{i,t+2}$ and $InfoAcq_{i,t+3}$. Table 8 Panel B, columns 1 to 3 present results using Edgar downloads. Column 1 reports the main specification for the subsequent quarter (also see Table 2 column 2), column 2 reports results using two-quarters-ahead Edgar downloads, and

[25] Consistent with the rest of the analyses, we percentile-rank all firms' abnormal prices within each Year-Quarter, such that $AbnPricePeer$ is defined as the percentile rank of the matched peer's abnormal price.

column 3 reports results using three-quarters-ahead Edgar downloads. Compared to the coefficient of interest in column 1, column 2's coefficient is significantly positive, but smaller in magnitude and column 3's coefficient is significantly positive and even smaller in magnitude. Table 8 Panel B, columns 4 to 6 present results using Edgar downloads. Column 4 reports the main specification for the subsequent quarter (also see Table 2 column 5), column 5 reports results using two-quarters-ahead Bloomberg attention, and column 6 reports results using three-quarters-ahead Bloomberg attention. Compared to the coefficient of interest in column 4, column 5's coefficient is significantly positive, but smaller in magnitude, and column 6's coefficient is significantly positive, and smaller in magnitude (but larger than column 5's). Our findings suggest that the use of information acquisition screens has some degree of persistence, which is consistent with speculators relying on the screens to allocate their information acquisition efforts, which in turn further increases the quality of private information in the subsequent period. Consequently, large price deviations predict greater future information acquisition, although this predictability generally decreases over time. These exploratory analyses suggest that a meaningful avenue for future research is to formalize dynamic predictions related to information acquisition screens.

6 Conclusion

An information event that offers profitable trading opportunities to investors who know about it and can assess its implications can occur for any tradeable asset at almost any time. As a result, it is challenging for a given investor to know where to focus their information acquisition efforts. Price-based screens that quickly sort large numbers of assets can guide these efforts and increase investors' expected trading profits. To provide some insight into why, how, and when screens are effective, and to identify some implications of their use for asset prices, we develop and analyze a model in which speculators can only follow one of two firms and are uncertain about the quality (i.e., profitability) of the private information that can be obtained by following each firm. When those speculators have access to past prices and earnings, they optimally employ a screen using those statistics to inform their firm following decisions. Within the context of this model of screens, speculators are more inclined to follow firms that have a larger deviation between price and the expectation of price conditioned on earnings. They do so because larger unexplained price deviations suggest that there is more private information in the marketplace. While intuitive, this observation stands in contrast to the prior literature in which past prices are irrelevant for information acquisition decisions. In addition, because firms whose prices deviate from database-driven pricing norms are more likely to attract the attention of speculators, our model also predicts that those firms will be more inclined to experience increased price volatility and trading activity. Our empirical analysis using abnormal price deviations at earnings announcements and investors' information acquisition on Edgar and Bloomberg find results consistent with these predictions.

Future research could explore a more general version of our framework by allowing for screens that include more information than just earnings to derive unexplained price movements and identify profitable information acquisition opportunities. Additionally, while we have focused on investors' response to past prices, empirical research has investigated managers' reactions in the presence (or absence) of institutional investor attention. Our model allows for an endogenous attention such that managers may be able to influence the attention that their firm receives from investors. For example, it would be interesting to analyze a setting where managers can manipulate (first period) reported earnings with the intent of attracting or discouraging institutional investors' information acquisition.

Appendix A Proofs

Derivation of probability that $q_i^2 = q_h^2$

Given all prices and earnings, $\Pr\left(q_i^2 = q_h^2 | \Omega\right) = \frac{\frac{1}{2}k\left(P_{i1}, e_{i1}, P_{j1}, e_{j1} | q_h^2\right)}{\frac{1}{2}k\left(P_{i1}, e_{i1}, P_{j1}, e_{j1} | q_h^2\right) + \frac{1}{2}k\left(P_{i1}, e_{i1}, P_{j1}, e_{j1} | q_l^2\right)}$, where $k\left(P_{i1}, e_{i1}, P_{j1}, e_{j1} | q_i^2\right)$ denotes the joint density of $\{P_{i1}, e_{i1}, P_{j1}, e_{j1}\}$ conditional upon q_i^2. Note that $\{P_{j1}, e_{j1}\}$ and e_{i1} are independent of q_i^2, and that $\{P_{i1}, e_{i1}\}$ and $\{P_{j1}, e_{j1}\}$ are independent of each other. It follows that $k\left(P_{i1}, e_{i1}, P_{j1}, e_{j1} | q_i^2\right)$ can be written as

$$f\left(P_{i1} | e_{i1}, q_i^2\right) g\left(e_{i1}\right) z\left(P_{j1}, e_{j1}\right), \tag{22}$$

where $f\left(P_{i1} | e_{i1}, q_i^2\right)$ is the probability density for P_{i1} conditioned upon e_{i1} and q_i^2, $g\left(e_{i1}\right)$ is the probability density for e_{i1}, and $z\left(P_{j1}, e_{j1}\right)$ is the joint probability density for $\{P_{j1}, e_{j1}\}$. Furthermore, given that $P_{i1} = 3e_{i1} + \beta_{i1x}x_{i1} + \beta_{i1n}n_{i1}$, where $\beta_{i1x} = \frac{8\pi_{i1}^2 q_i^2}{4\pi_{i1}^2 q_i^2 + \sigma_{i1}^2}$ and $\beta_{i1n} = \frac{4\pi_{i1} q_i^2}{4\pi_{i1}^2 q_i^2 + \sigma_{i1}^2}$, and x_{i1} and n_{i1} are independent mean 0 normally distributed random variables, it follows that P_{i1} conditional upon e_{i1} and q_{i1} is a normally distributed random variable with mean $3e_{i1}$ and variance $V_i^2 = \frac{16\pi_{i1}^2 q_i^2}{4\pi_{i1}^2 q_i^2 + \sigma_{i1}^2} q_i^2$. Hence,

$$\begin{aligned}
\Pr\left(q_i^2 = q_h^2 | \Omega\right) &= \frac{\frac{1}{2}k\left(P_{i1}, e_{i1}, P_{j1}, e_{j1} | q_h^2\right)}{\frac{1}{2}k\left(P_{i1}, e_{i1}, P_{j1}, e_{j1} | q_h^2\right) + \frac{1}{2}k\left(P_{i1}, e_{i1}, P_{j1}, e_{j1} | q_l^2\right)} \\
&= \frac{f\left(P_{i1} | e_{i1}, q_h^2\right)}{f\left(P_{i1} | e_{i1}, q_h^2\right) + f\left(P_{i1} | e_{i1}, q_l^2\right)} \\
&= \frac{1}{1 + \frac{f\left(P_{i1} | e_{i1}, q_l^2\right)}{f\left(P_{i1} | e_{i1}, q_h^2\right)}} \\
&= \frac{1}{1 + \sqrt{\frac{V_h^2}{V_l^2}} \exp\left[-\left(P_{i1} - 3e_{i1}\right)^2 \frac{V_h^2 - V_l^2}{2V_h^2 V_l^2}\right]}.
\end{aligned} \tag{23}$$

Proof of Corollary 1

Note that only $\Pr\left(q_i^2 = q_h^2 | \Omega\right)$ in eqn. (16) is a function of $\left(P_{i1} - 3e_{i1}\right)^2$ and that

$$\partial \Pr\left(q_i^2 = q_h^2 | \Omega\right) / \partial \left(P_{i1} - 3e_{i1}\right)^2 > 0. \tag{24}$$

This implies that $\partial E\left[\Pi\left(q_i^2, \pi_{i2}\left(\Omega\right)\right) | \Omega\right] / \partial \left(P_{i1} - 3e_{i1}\right)^2 > 0$ such that, in equilibrium, $\partial \pi_{i2}^*\left(\Omega\right) / \partial \left(P_{i1} - 3e_{i1}\right)^2 > 0$ and, therefore $\partial \pi_{j2}^*\left(\Omega\right) / \partial \left(P_{i1} - 3e_{i1}\right)^2 < 0$.

Proof of Corollary 2

Holding constant the quality of private information, the variance of second-period price is given by

$$\begin{aligned} Var\left[P_{i2}\right] &= Var\left[e_{i1}+2e_{i2}+\frac{\left(\pi_{i1}+\pi_{i2}\left(\Omega\right)\right)^{2}q_{i}^{2}}{\left(\pi_{i1}+\pi_{i2}\left(\Omega\right)\right)^{2}q_{i}^{2}+\sigma_{i2}^{2}}x_{i2}+\frac{\left(\pi_{i1}+\pi_{i2}\left(\Omega\right)\right)q_{i}^{2}}{\left(\pi_{i1}+\pi_{i2}\left(\Omega\right)\right)^{2}q_{i}^{2}+\sigma_{i2}^{2}}n_{i2}\right] \\ &= 4s_{i}^{2}+\left(\frac{\left(\pi_{i1}+\pi_{i2}\left(\Omega\right)\right)^{2}q_{i}^{2}}{\left(\pi_{i1}+\pi_{i2}\left(\Omega\right)\right)^{2}q_{i}^{2}+\sigma_{i2}^{2}}\right)^{2}q_{i}^{2}+\left(\frac{\left(\pi_{i1}+\pi_{i2}\left(\Omega\right)\right)q_{i}^{2}}{\left(\pi_{i1}+\pi_{i2}\left(\Omega\right)\right)^{2}q_{i}^{2}+\sigma_{i2}^{2}}\right)^{2}\sigma_{i2}^{2}. \qquad (25) \end{aligned}$$

Finally,

$$\frac{\partial Var\left[P_{i2}\right]}{\partial\pi_{i2}\left(\Omega\right)}=\frac{2\left(\pi_{i1}+\pi_{i2}\left(\Omega\right)\right)q_{i}^{4}\sigma_{i2}^{2}}{\left(\left(\pi_{i1}+\pi_{i2}\left(\Omega\right)\right)^{2}q_{i}^{2}+\sigma_{i2}^{2}\right)^{2}}>0, \qquad (26)$$

which completes the proof because Corollary 1 shows that $\frac{\partial\pi_{i2}^{*}(\Omega)}{\partial|P_{i1}-3e_{i1}|}>0$.

Proof of Corollary 3

The second-period trading volume from new speculators for firm i's claims is given by

$$V=\pi_{i2}\left(\Omega\right)\left|d_{i2S}\right|. \qquad (27)$$

Thus,

$$cV=\left|A_{1}x_{i2}-A_{2}n_{i2}\right|, \qquad (28)$$

where $A_1=\frac{\pi_{i2}^{*}(\Omega)\sigma_{i2}^{2}}{\left(\pi_{i1}+\pi_{i2}^{*}(\Omega)\right)^{2}q_{i}^{2}+\sigma_{i2}^{2}}$ and $A_2=\frac{\pi_{i2}^{*}(\Omega)(\pi_{i1}+\pi_{i2}^{*}(\Omega))q_{i}^{2}}{\left(\pi_{i1}+\pi_{i2}^{*}(\Omega)\right)^{2}q_{i}^{2}+\sigma_{i2}^{2}}$. Note that $A_1x_{i2}-A_2n_{i2}\sim N\left(0,A_1^2q_i^2+A_2^2\sigma_{i2}^2\right)$. Therefore,

$$\begin{aligned} cE\left[V\right] &= \sqrt{\frac{2}{\pi}\left(A_{1}^{2}q_{i}^{2}+A_{2}^{2}\sigma_{i2}^{2}\right)} \qquad (29) \\ &= \sqrt{\frac{2}{\pi}\left(\frac{\pi_{i2}^{*}\left(\Omega\right)^{2}q_{i}^{2}\sigma_{i2}^{2}}{\left(\pi_{i1}+\pi_{i2}^{*}\left(\Omega\right)\right)^{2}q_{i}^{2}+\sigma_{i2}^{2}}\right)}. \qquad (30) \end{aligned}$$

Finally,

$$c\frac{\partial E\left[V\right]}{\partial\pi_{i}^{*}\left(\Omega\right)}=\sqrt{\frac{2}{\pi}}\frac{\left(\sigma_{i2}^{2}+q_{i}^{2}\pi_{i1}\left(\pi_{i1}+\pi_{i2}^{*}\left(\Omega\right)\right)\right)\left(\frac{q_{i}^{2}\sigma_{i2}^{2}(\pi_{i2}^{*}(\Omega))^{2}}{\sigma_{i2}^{2}+q_{i}^{2}(\pi_{i1}+\pi_{i2}^{*}(\Omega))^{2}}\right)^{\frac{3}{2}}}{q_{i}^{2}\sigma_{i2}^{2}\left(\pi_{i2}^{*}\left(\Omega\right)\right)^{3}}>0, \qquad (31)$$

which completes the proof because Corollary 1 shows that $\frac{\partial\pi_{i2}^{*}(\Omega)}{\partial|P_{i1}-3e_{i1}|}>0$.

Appendix B Variable Definitions

This table presents definitions of the primary variables used throughout the paper. All continuous variables are winsorized at 1% and 99% to limit the influence of outliers.

Variable	Definition
	Main analysis variables
AbnPrice	The absolute difference between a firm's P/E ratio on the first trading day after the earnings announcement and the benchmark P/E ratio, scaled by the benchmark P/E ratio. The benchmark P/E ratio is the median P/E ratio of the firms in the same 2-digit SIC industry and asset size quintile, measured four quarters prior. The P/E ratio is computed by dividing the price per share on the first trading day after the earnings announcement by the announced earnings per share, and is defined only for the observations with positive earnings per share. We percentile-rank this variable, within each quarter.
AbnPricePeer	*AbnPrice* of the matched peer firm. This matched peer is defined as the firm in the same industry (defined based on 2-digit SIC codes) as and that announces earnings on the same day as the focal firm, that is closest in asset size. In addition, we only retain matches where the larger firm's asset size is less than 1.5 times the smaller firm's, and the size difference between the two firms is less than $100 billion.
Bloomberg_5day	Sum of Bloomberg's measure of abnormal institutional investor attention between 1 and 5 trading days subsequent to this quarter's earnings announcement date. We log-transform this variable by taking log(1 + Bloomberg raw sum) and multiply the log-transformed variable by 100.
Bloomberg_lag	*Bloomberg_5day* or *Bloomberg*, measured in the previous quarter.
Bloomberg	Sum of Bloomberg's measure of abnormal institutional investor attention between this quarter's and the subsequent quarter's earnings announcement dates. We log-transform this variable by taking log(1 + Bloomberg raw sum) and multiply the log-transformed variable by 100.
BundledForecast	Indicator variable set to one if the firm issues earnings guidance on trading days [0,+2] of this quarter's earnings announcement and zero otherwise.
EdgarSearch_5day	The total number of Edgar downloads between 1 and 5 trading days subsequent to this quarter's earnings announcement date. The measure includes downloads of all filing forms and excludes bot downloads, following Drake et al. (2015). We log-transform this variable by taking log(1 + EdgarSearch raw sum) and multiply the log-transformed variable by 100.
EdgarSearch_lag	*EdgarSearch_5day* or *EdgarSearch*, measured in the previous quarter.
EdgarSearch	The total number of Edgar downloads between this quarter's and the subsequent quarter's earnings announcement dates. The measure includes downloads of all filing forms and excludes bot downloads, following Drake et al. (2015). We log-transform this variable by taking log(1 + EdgarSearch raw sum) and multiply the log-transformed variable by 100.
Instown	Proportion of institutional ownership from Thomson 13F, multiplied by 100 and measured at the end of the quarter.
Leverage	Debt divided by equity, measured at the end of the quarter.
log(*Assets*)	The natural log of one plus total assets, measured in thousands of dollars at the end of the quarter.
MediaStories	Number of unique media stories about the firm in a quarter, from RavenPack.
RetVariance	Standard deviation of daily returns between this quarter's and the subsequent quarter's earnings announcement dates, multiplied by 100.
Shortsell	Outstanding short-sell interest relative to shares outstanding, from Compustat, multiplied by 100 and measured at the end of the quarter.
TradingVol	Sum of the daily percentage turnover, calculated as the volume of shares traded scaled by shares outstanding, between this quarter's and the subsequent quarter's earnings announcement dates and multiplied by 100.
Voldisc8K	Number of voluntary Form 8-K disclosures (Items 2.02, 7.01, 8.01) issued by the firm in a quarter.

Variable	Definition
	Additional analysis variables
AbnPriceChange	The absolute difference between the 3-day earnings response coefficient and its historical average across the previous four earnings announcement dates. The 3-day earnings response coefficient is calculated as the cumulative abnormal returns (CAR) over the three trading days centered on the earnings announcement date, divided by unexpected earnings (UE). UE is the quarterly earnings surprise relative to analyst consensus, scaled by the quarter-end stock price. This measure is calculated only for those observations where CAR and UE have the same sign and is percentile-ranked within-quarter.
$Bloomberg_{t+2}$	Sum of Bloomberg's measure of abnormal institutional investor attention between next quarter's $(t+1)$ and the subsequent quarter's $(t+2)$ earnings announcement dates. We log-transform this variable by taking $\log(1 + \text{Bloomberg raw sum}_{t+2})$ and multiply the log-transformed variable by 100.
$Bloomberg_{t+3}$	Sum of Bloomberg's measure of abnormal institutional investor attention between two quarters ahead $(t+2)$ and the subsequent quarter's $(t+3)$ earnings announcement dates. We log-transform this variable by taking $\log(1 + \text{Bloomberg raw sum}_{t+3})$ and multiply the log-transformed variable by 100.
EdgarInst	The total number of Edgar downloads between this quarter's and the subsequent quarter's earnings announcement dates that are initiated from IP addresses that are identified to be a financial institution as per ARIN WhoWas reports. The measure includes downloads of all filing forms and excludes bot downloads, following Drake et al. (2015). We log-transform this variable by taking $\log(1 + \text{EdgarInst raw sum})$ and multiply the log-transformed variable by 100.
EdgarISP	The total number of Edgar downloads between this quarter's and the subsequent quarter's earnings announcement dates that are initiated from IP addresses that are identified to be an internet service provider (ISP) as per ARIN WhoWas reports. The measure includes downloads of all filing forms and excludes bot downloads, following Drake et al. (2015). We log-transform this variable by taking $\log(1 + \text{EdgarISP raw sum})$ and multiply the log-transformed variable by 100.
EdgarNew	The total number of Edgar downloads between this quarter's and the subsequent quarter's earnings announcement dates that are initiated from new IP addresses. New IP addresses refer to the ones that did not download the firm's filings in the previous quarter over the same window. The measure includes downloads of all filing forms and excludes bot downloads, following Drake et al. (2015). We log-transform this variable by taking $\log(1 + \text{EdgarNew raw sum})$ and multiply the log-transformed variable by 100.
EdgarNewIP	Number of unique IP addresses downloading Edgar filings of a firm between this quarter's and subsequent quarter's earnings announcement dates, considering only the new IP addresses that did not download the firm's filings in the previous quarter over the same window. This measure includes IP addresses accessing all filing forms and excludes IP addresses identified as bots, following Drake et al. (2015). We log-transform this variable by taking $\log(1 + \text{EdgarNewIPs raw sum})$ and multiply the log-transformed variable by 100.
$EdgarSearch_{t+2}$	The total number of Edgar downloads between next quarter's $(t+1)$ and the subsequent quarter's $(t+2)$ earnings announcement dates. The measure includes downloads of all filing forms and excludes bot downloads, following Drake et al. (2015). We log-transform this variable by taking $\log(1 + \text{EdgarSearch raw sum}_{t+2})$ and multiply the log-transformed variable by 100.
$EdgarSearch_{t+3}$	The total number of Edgar downloads between two quarters ahead $(t+2)$ and the subsequent quarter's $(t+3)$ earnings announcement dates. The measure includes downloads of all filing forms and excludes bot downloads, following Drake et al. (2015). We log-transform this variable by taking $\log(1 + \text{EdgarSearch raw sum}_{t+3})$ and multiply the log-transformed variable by 100.
Loss	Indicator variable set to one if net income is negative for the quarter and zero otherwise.

References

Abramova, I., J. E. Core, and A. Sutherland (2020). Institutional investor attention and firm disclosure. *The Accounting Review 95*(6), 1–21.

Akbas, F., S. Markov, M. Subasi, and E. Weisbrod (2018). Determinants and consequences of information processing delay: Evidence from the thomson reuters institutional brokers' estimate system. *Journal of Financial Economics 127*(2), 366–388.

Andrei, D. and M. Hasler (2015). Investor attention and stock market volatility. *Review of Financial Studies 28*(1), 33–72.

Angrist, J. D. and J.-S. Pischke (2009). *Mostly harmless econometrics: An empiricist's companion.* Princeton university press.

Banerjee, S. and B. Green (2015). Signal or noise? uncertainty and learning about whether other traders are informed. *Journal of Financial Economics 117*(1), 398–423.

Blankespoor, E., E. deHaan, and C. Zhu (2018). Capital market effects of media synthesis and dissemination: Evidence from robo-journalism. *Review of Accounting Studies 23*, 1–36.

Blume, L., D. Easley, and M. O'Hara (1994). Signal or noise? uncertainty and learning about whether other traders are informed. *Journal of Finance 49*(1), 153–181.

Bonsall IV, S. B., J. R. Green, and K. A. Muller III (2018). Are credit ratings more rigorous for widely covered firms? *The Accounting Review 93*(6), 61–94.

Brown, S., S. A. Hillegeist, and K. Lo (2009). The effect of earnings surprises on information asymmetry. *Journal of Accounting and Economics 47*(3), 208–225.

Bushee, B. J., J. Gerakos, and L. F. Lee (2018). Corporate jets and private meetings with investors. *Journal of Accounting and Economics 65*(2-3), 358–379.

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

DellaVigna, S. and J. M. Pollet (2009). Investor inattention and friday earnings announcements. *Journal of Finance 64*(2), 709–749.

Demski, J. S. and G. A. Feltham (1994). Market response to financial reports. *Journal of Accounting and Economics 17*(1-2), 3–40.

Drake, M. S., B. A. Johnson, D. T. Roulstone, and J. R. Thornock (2020). Is there information content in information acquisition? *The Accounting Review 95*(2), 113–139.

Drake, M. S., D. T. Roulstone, and J. R. Thornock (2015). The determinants and consequences of information acquisition via edgar. *Contemporary Accounting Research 32*(3), 1128–1161.

Drake, M. S., D. T. Roulstone, and J. R. Thornock (2016). The usefulness of historical accounting reports. *Journal of Accounting and Economics 61*(2-3), 448–464.

Dye, R. A. and S. S. Sridhar (2007). The allocational effects of the precision of accounting estimates. *Journal of Accounting Research 45*(4), 731–769.

Dyer, T. A. (2021). The demand for public information by local and nonlocal investors: Evidence from investor-level data. *Journal of Accounting and Economics 72*(1), 101417.

Fischer, P. E. and R. E. Verrecchia (1998). Correlated forecast errors. *Journal of Accounting Research 36*(1), 91–110.

Froot, K. A., D. S. Scharfstein, and J. C. Stein (1992). Herd on the street: Informational inefficiencies in a market with short-term speculation. *Journal of Finance 47*, 1451–1484.

Gao, P. and P. J. Liang (2013). Informational feedback, adverse selection, and optimal disclosure policy. *Journal of Accounting Research 51*(5), 1133–1158.

Garcia, D. and G. Strobl (2011). Relative wealth concerns and complementarities in information acquisition. *Review of Financial Studies 24*(1), 169–207.

Garcia, D. and J. M. Vanden (2009). Information acquisition and mutual funds. *Journal of Economic Theory 144*(5), 1965–1995.

Green, J., J. R. Hand, and X. F. Zhang (2017). The characteristics that provide independent information about average us monthly stock returns. *The Review of Financial Studies 30*(12), 4389–4436.

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

Hughes, J. S. and S. Pae (2004). Voluntary disclosure of precision information. *Journal of Accounting and Economics 37*(2), 261–289.

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

Kempf, E., A. Manconi, and O. Spalt (2017). Distracted shareholders and corporate actions. *Review of Financial Studies 30*(5), 1660–1695.

Kim, J. M. (2023). Economics of information search and financial misreporting. *Working Paper*.

Kim, O. and E. Verrecchia, Robert (1994). Market liquidity and volume around earnings announcements. *Journal of Accounting and Economics 17*(1-2), 41–67.

Ko, K. J. and Z. J. Huang (2007). Arrogance can be a virtue: Overconfidence, information acquisition, and market efficiency. *Journal of Financial Economics 84*(2), 529–560.

Koester, A., R. Lundholm, and M. Soliman (2016). Attracting attention in a limited attention world: Exploring the causes and consequences of extreme positive earnings surprises. *Management Science 62*(10), 2871–2896.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica 53*(6), 1315–1335.

MacKinnon, D. (2012). *Introduction to statistical mediation analysis*. Routledge.

McNichols, M. and B. Trueman (1998). Public disclosure, private information collection, and short-term trading. *Journal of Accounting and Economics 17*(1-2), 69–94.

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

Michaeli, B. (2017). Divide and inform: Rationing information to facilitate persuasion. *The Accounting Review 92*(5), 167–199.

Mondria, J. (2010). Portfolio choice, attention allocation, and price comovement. *Journal of Economic Theory 145*(5), 1837–1864.

Penno, M. (1996). Unobservable precision choices in financial reporting. *Journal of Accounting Research 34*(1), 141–149.

Ryans, J. (2017). Using the edgar log file data set. *Working Paper*.

Schneider, J. (2009). A rational expectations equilibrium with informative trading volume. *Journal of Accounting Research 64*(6), 2783–2805.

Smith, K. (2022). Risk information, investor learning, and informational feedback. *Review of Accounting Studies*, 1–39.

Solomon, D. and E. Soltes (2015). What are we meeting for? the consequences of private meetings with investors. *The Journal of Law and Economics 58*(2), 325–355.

Subramanyam, K. R. (1996). Uncertain precision and price reactions to information. *The Accounting Review 71*(2), 207–219.

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

Verrecchia, R. E. (1982). Information acquisition in a noisy rational expectations economy. *Econometrica 50*(6), 1415–1430.

Zhu, C. (2019). Big data as a governance mechanism. *The Review of Financial Studies 32*(5), 2021–2061.

Table 1

Descriptive Statistics

Panel A. Edgar Downloads Sample: 2004-2016

	Obs.	Mean	Std.Dev.	Q1	Median	Q3
AbnPrice	87,493	0.73	1.41	0.15	0.34	0.63
EdgarSearch	87,493	2,025.11	2,435.86	575.00	1,189.00	2,444.00
EdgarSearch_5day	87,493	183.63	223.19	51.00	108.00	222.00
TradingVol	87,493	54.30	47.42	24.43	41.61	68.23
RetVariance	87,493	2.26	1.40	1.36	1.89	2.71
Leverage	87,493	0.93	2.10	0.10	0.48	1.11
Assets(in millions)	87,493	11,061.84	31,861.58	563.40	1,803.47	6,269.00
BundledForecast	87,493	0.36	0.48	0.00	0.00	1.00
Instown	87,493	53.32	36.99	8.34	63.32	86.42
Shortsell	87,493	4.15	4.58	1.13	2.69	5.58
Voldisc8K	87,493	2.10	1.66	1.00	2.00	3.00
MediaStories	87,493	38.12	56.11	2.00	18.00	49.00

Panel B. Bloomberg Attention Sample: 2010-2022

	Obs.	Mean	Std.Dev.	Q1	Median	Q3
AbnPrice	59,895	0.86	1.73	0.16	0.35	0.67
Bloomberg	59,895	21.50	30.26	0.00	9.00	30.00
Bloomberg_5day	59,895	1.54	2.87	0.00	0.00	2.00
TradingVol	59,895	49.03	41.96	25.90	39.01	58.81
RetVariance	59,895	2.05	1.23	1.30	1.72	2.40
Leverage	59,895	1.02	2.46	0.23	0.63	1.21
Assets(in millions)	59,895	16,135.70	38,210.88	1,096.90	3,193.59	10,618.84
BundledForecast	59,895	0.35	0.48	0.00	0.00	1.00
Instown	59,895	49.56	39.88	0.00	62.28	87.10
Shortsell	59,895	3.78	4.02	1.30	2.47	4.73
Voldisc8K	59,895	2.20	1.58	1.00	2.00	3.00
MediaStories	59,895	68.91	97.62	0.00	42.00	90.00

This table presents descriptive statistics for the variables used in our analysis. Panel A presents summary statistics of variables, measured at the firm-quarter level, for the 2004 to 2016 sample examining Edgar downloads. Panel B presents summary statistics of variables, measured at the firm-quarter level, for the 2010 to 2022 sample examining Bloomberg attention. All variables are defined in Appendix B. For ease of interpretation, we present summary statistics for raw variables (e.g., unranked and unlogged). All continuous variables are winsorized at the 1% and 99% levels to limit the influence of outliers.

Table 2
Abnormal Price and Information Acquisition

	EdgarSearch			Bloomberg		
	(1)	(2)	(3)	(4)	(5)	(6)
AbnPrice	0.040***	0.034***	0.012**	0.090***	0.089***	0.154***
	(6.25)	(5.43)	(2.49)	(3.89)	(3.94)	(9.91)
Leverage		-0.090	-0.060		-0.813**	-0.219
		(-0.68)	(-0.78)		(-2.40)	(-1.29)
log(Assets)		16.324***	5.708***		40.688***	20.431***
		(16.70)	(27.39)		(13.53)	(34.91)
BundledForecast		-2.744***	1.787***		-1.368	2.920**
		(-3.17)	(4.71)		(-0.48)	(2.19)
Instown		-0.086***	0.006		0.027	-0.013
		(-5.09)	(1.16)		(0.63)	(-0.87)
Shortsell		0.867***	0.370***		2.791***	1.616***
		(11.17)	(11.20)		(9.64)	(13.26)
Voldisc8K		2.414***	0.089		2.303***	-1.031***
		(16.31)	(0.81)		(4.74)	(-3.31)
MediaStories		0.047***	0.078***		-0.004	0.039***
		(5.21)	(19.60)		(-0.30)	(6.23)
EdgarSearch_lag			0.764***			
			(140.83)			
Bloomberg_lag						0.700***
						(122.25)
Observations	87,493	87,493	87,493	59,895	59,895	59,895
Adj R-Squared	0.877	0.881	0.880	0.715	0.723	0.762
Firm FE	Yes	Yes	No	Yes	Yes	No
Industry FE	No	No	Yes	No	No	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes

This table presents an analysis of the relation between abnormal price movement at the earnings announcement and subsequent information acquisition. Columns 1 to 3 examine information acquisition on Edgar. Columns 4 to 6 examine information acquisition on Bloomberg. All variables are defined in Appendix B. All continuous variables are winsorized at the 1% and 99% levels to limit the influence of outliers. *t*-statistics based on standard errors clustered by Firm are in parentheses. Levels of significance are presented as follows: *p<0.1; **p<0.05; ***p<0.01.

Table 3

Abnormal Price, Trading Volume, and Return Volatility: Mediated (Path) Analysis

Panel A. Through Information Acquisition on Edgar: Effects

Outcome:	TradingVol		RetVariance	
	(1)	(2)	(3)	(4)
	Coef	Bootstrap z	Coef	Bootstrap z
Direct Path:				
I. AbnPrice → Outcome	0.006	1.62	0.003***	22.87
Mediated Path:				
II. AbnPrice → EdgarSearch	0.034***	6.58	0.034***	6.58
III. EdgarSearch → Outcome	0.247***	67.12	0.004***	36.92
Indirect Effect (II × III)	0.008***	6.61	0.0001***	6.54
Total Effect (I + II × III)	0.014***	3.38	0.003***	24.06
Controls	Yes		Yes	
Firm FE	Yes		Yes	
Year-Quarter FE	Yes		Yes	

Panel B. Through Information Acquisition on Edgar: Mediated Regressions

	TradingVol		RetVariance	
	(1)	(2)	(3)	(4)
AbnPrice	0.014***	0.006	0.003***	0.003***
	(2.65)	(1.14)	(16.94)	(16.42)
EdgarSearch		0.247***		0.004***
		(38.80)		(25.63)
Leverage	0.083	0.105	0.018***	0.019***
	(0.69)	(0.91)	(4.28)	(4.39)
log(Assets)	1.343	-2.687**	-0.040**	-0.098***
	(1.16)	(-2.41)	(-2.18)	(-5.45)
BundledForecast	-0.503	0.175	-0.085***	-0.075***
	(-0.71)	(0.26)	(-5.33)	(-4.73)
Instown	0.047***	0.068***	-0.003***	-0.003***
	(2.97)	(4.53)	(-8.00)	(-7.26)
Shortsell	2.895***	2.681***	0.025***	0.022***
	(29.62)	(29.10)	(11.45)	(10.20)
Voldisc8K	0.544***	-0.052	0.003	-0.006**
	(4.70)	(-0.48)	(0.90)	(-2.02)
MediaStories	0.019**	0.008	0.000***	0.000
	(2.34)	(0.97)	(2.58)	(1.30)
Observations	87,493	87,493	87,493	87,493
Adj R-Squared	0.631	0.665	0.632	0.641
Firm FE	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes

Table 3

Abnormal Price, Trading Volume, and Return Volatility: Mediated (Path) Analysis (cont'd)

Panel C. Through Information Acquisition on Bloomberg: Effects

Outcome:	TradingVol		RetVariance	
	(1)	(2)	(3)	(4)
	Coef	Bootstrap z	Coef	Bootstrap z
Direct Path:				
I. AbnPrice → Outcome	0.024***	5.42	0.001***	11.16
Mediated Path:				
II. AbnPrice → Bloomberg	0.089***	6.32	0.089***	6.32
III. Bloomberg → Outcome	0.048***	40.83	0.008***	23.48
Indirect Effect (II × III)	0.004***	6.10	0.0001***	5.72
Total Effect (I + II × III)	0.029***	6.68	0.001***	11.58
Controls	Yes		Yes	
Firm FE	Yes		Yes	
Year-Quarter FE	Yes		Yes	

Panel D. Through Information Acquisition on Bloomberg: Mediated Regressions

	TradingVol		RetVariance	
	(1)	(2)	(3)	(4)
AbnPrice	0.029***	0.024***	0.001***	0.001***
	(4.50)	(3.93)	(9.15)	(8.73)
Bloomberg		0.048***		0.001***
		(17.58)		(16.09)
Leverage	0.123	0.163	0.009***	0.010***
	(0.94)	(1.24)	(2.79)	(3.01)
log(Assets)	0.263	-1.706*	-0.045**	-0.079***
	(0.26)	(-1.67)	(-2.45)	(-4.25)
BundledForecast	0.370	0.437	-0.039*	-0.038*
	(0.43)	(0.51)	(-1.92)	(-1.88)
Instown	0.009	0.008	-0.001***	-0.001***
	(0.79)	(0.68)	(-3.49)	(-3.58)
Shortsell	2.901***	2.766***	0.016***	0.014***
	(19.84)	(19.29)	(7.42)	(6.40)
Voldisc8K	0.465***	0.354**	-0.011***	-0.013***
	(3.28)	(2.54)	(-3.38)	(-4.02)
MediaStories	0.013***	0.013***	0.000***	0.000***
	(3.52)	(3.55)	(4.64)	(4.70)
Observations	59,895	59,895	59,895	59,895
Adj R-Squared	0.613	0.624	0.638	0.642
Firm FE	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes

This table presents a mediated (path) analysis of how abnormal price movement at the earnings announcement affects subsequent trading volume and subsequent return volatility through information acquisition. Panels A and B examine information acquisition on Edgar. Panels C and D examine information acquisition on Bloomberg. Panels A and C present the magnitude and the significance of the direct and indirect (e.g., through information acquisition) effects of abnormal price movement on trading volume and return volatility. z-statistics are based on bootstrapped standard errors clustered by Firm. Panels B and D compare the baseline regression analysis with a mediated regression analysis that includes measures of information acquisition as an additional explanatory variable. This approach follows Bonsall IV et al. (2018). All variables are defined in Appendix B. All continuous variables are winsorized at the 1% and 99% levels to limit the influence of outliers. t-statistics based on standard errors clustered by Firm are in parentheses. Levels of significance are presented as follows: *p<0.1; **p<0.05; ***p<0.01.

Table 4
Abnormal Price and Alternative Measures of Information Acquisition

	EdgarNew	EdgarNewIP	EdgarInst	EdgarISP
	(1)	(2)	(3)	(4)
AbnPrice	0.027***	0.026***	0.027***	0.034***
	(4.62)	(4.82)	(3.10)	(3.49)
Leverage	-0.031	-0.032	0.284*	0.207
	(-0.26)	(-0.29)	(1.93)	(1.12)
log(Assets)	16.880***	15.221***	26.649***	15.822***
	(18.77)	(16.91)	(21.95)	(11.12)
BundledForecast	-2.820***	-2.398***	-2.639**	-4.330***
	(-3.59)	(-3.33)	(-2.50)	(-3.54)
Instown	-0.075***	-0.058***	-0.109***	-0.136***
	(-5.00)	(-4.02)	(-5.10)	(-5.23)
Shortsell	0.870***	0.739***	1.145***	1.114***
	(12.17)	(11.26)	(11.04)	(9.56)
Voldisc8K	1.691***	1.386***	2.455***	3.161***
	(12.39)	(11.16)	(11.80)	(14.25)
MediaStories	0.021**	0.010	0.074***	0.058***
	(2.50)	(1.15)	(8.20)	(5.01)
Observations	87,493	87,493	87,493	87,493
Adj R-Squared	0.872	0.896	0.803	0.755
Firm FE	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes

This table presents an analysis of the relation between abnormal price movement at the earnings announcement and subsequent information acquisition. Columns 1 and 2 measure information acquisition as either new Edgar downloads or the number of unique new IP addresses downloading Edgar filings, considering only the new IP addresses that did not download the firm's filings in the previous quarter over the same window. The mean (median) values of raw, unlogged $EdgarNew$ and $EdgarNewIP$ are 1,287.54 (808.00) and 595.14 (366.00), respectively. Columns 3 and 4 isolate Edgar downloads by financial institutions ($EdgarInst$) versus those initiated from internet service providers ($EdgarISP$). The mean (median) values of raw, unlogged $EdgarInst$ and $EdgarISP$ are 139.50 (81.00) and 119.50 (69.00), respectively. All variables are defined in Appendix B. All continuous variables are winsorized at the 1% and 99% levels to limit the influence of outliers. t-statistics based on standard errors clustered by Firm are in parentheses. Levels of significance are presented as follows: *p<0.1; **p<0.05; ***p<0.01.

Table 5
Abnormal Price Change and Information Acquisition

	EdgarSearch			Bloomberg		
	(1)	(2)	(3)	(4)	(5)	(6)
AbnPriceChange	0.039***	0.023***	0.010**	0.234***	0.169***	0.080***
	(5.66)	(3.68)	(2.00)	(11.11)	(8.49)	(6.05)
Loss		4.806***	3.928***		0.622	10.499***
		(9.52)	(10.87)		(0.39)	(9.71)
Leverage		0.135	-0.070		-0.822***	-0.544***
		(1.45)	(-1.10)		(-3.82)	(-3.73)
log(Assets)		14.255***	5.933***		38.926***	17.488***
		(18.74)	(33.77)		(18.20)	(39.45)
BundledForecast		-2.781***	1.074***		0.235	1.993*
		(-3.61)	(2.86)		(0.09)	(1.65)
Instown		-0.093***	-0.012**		-0.000	-0.045***
		(-6.09)	(-2.48)		(-0.00)	(-3.35)
Shortsell		0.987***	0.470***		2.539***	1.242***
		(15.28)	(15.51)		(11.57)	(13.13)
Voldisc8K		2.290***	0.289***		2.349***	-0.545**
		(16.40)	(2.68)		(5.57)	(-2.00)
MediaStories		0.059***	0.091***		0.008	0.052***
		(7.49)	(23.30)		(0.68)	(9.61)
EdgarSearch_lag			0.733***			
			(155.13)			
Bloomberg_lag						0.696***
						(140.68)
Observations	87,346	87,346	87,346	62,745	62,745	62,745
Adj R-Squared	0.863	0.868	0.863	0.701	0.712	0.754
Firm FE	Yes	Yes	No	Yes	Yes	No
Industry FE	No	No	Yes	No	No	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes

This table repeats the analysis from Table 2, using a changes-based measure of abnormal price movement. Columns 1 to 3 (4 to 6) examine information acquisition on Edgar (Bloomberg). All variables are defined in Appendix B. All continuous variables are winsorized at the 1% and 99% levels to limit the influence of outliers. *t*-statistics based on standard errors clustered by Firm are in parentheses. Levels of significance are presented as follows: *p<0.1; **p<0.05; ***p<0.01.

Table 6
Alternative Measurement Window

	EdgarSearch_5day			Bloomberg_5day		
	(1)	(2)	(3)	(4)	(5)	(6)
AbnPrice	0.065***	0.058***	0.066***	0.037***	0.038***	0.102***
	(8.32)	(7.53)	(7.52)	(3.08)	(3.23)	(8.53)
Leverage		-0.154	-0.356**		-0.419**	-0.222
		(-1.05)	(-2.45)		(-2.09)	(-1.50)
log(Assets)		17.591***	12.864***		17.374***	16.352***
		(17.62)	(34.06)		(11.32)	(36.34)
BundledForecast		-5.229***	1.811**		0.029	-1.247
		(-5.66)	(2.35)		(0.02)	(-1.19)
Instown		-0.054***	-0.012		0.012	-0.079***
		(-2.92)	(-1.20)		(0.62)	(-6.58)
Shortsell		0.977***	0.969***		0.606***	0.658***
		(11.33)	(15.67)		(4.53)	(7.03)
Voldisc8K		2.957***	2.751***		1.149***	0.367
		(16.56)	(12.60)		(4.57)	(1.53)
MediaStories		0.057***	0.195***		0.001	0.071***
		(6.49)	(24.61)		(0.09)	(11.98)
EdgarSearch_lag			0.464***			
			(61.98)			
Bloomberg_lag						0.362***
						(44.40)
Observations	87,493	87,493	87,493	59,895	59,895	59,895
Adj R-Squared	0.784	0.790	0.731	0.516	0.522	0.478
Firm FE	Yes	Yes	No	Yes	Yes	No
Industry FE	No	No	Yes	No	No	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes

This table repeats the analysis from Table 2, using shorter a shorter, 5-day window to measure information acquisition. Columns 1 to 3 (4 to 6) examine information acquisition on Edgar (Bloomberg). All variables are defined in Appendix B. All continuous variables are winsorized at the 1% and 99% levels to limit the influence of outliers. *t*-statistics based on standard errors clustered by Firm are in parentheses. Levels of significance are presented as follows: *p<0.1; **p<0.05; ***p<0.01.

Table 7
Matched-Peer Firm Analysis

	EdgarSearch			Bloomberg		
	(1)	(2)	(3)	(4)	(5)	(6)
AbnPrice	0.041***	0.032***	0.020***	0.119***	0.117***	0.185***
	(5.04)	(4.12)	(3.26)	(3.54)	(3.58)	(8.81)
AbnPricePeer	-0.015**	-0.011*	-0.010*	-0.064***	-0.059**	0.003
	(-2.35)	(-1.71)	(-1.71)	(-2.69)	(-2.56)	(0.15)
Leverage		-0.550**	-0.073		-0.976	-0.239
		(-2.19)	(-0.56)		(-1.61)	(-0.78)
log(Assets)		15.254***	5.152***		48.791***	21.645***
		(13.99)	(19.75)		(11.55)	(27.85)
BundledForecast		-4.128***	0.685		-5.468	2.067
		(-4.20)	(1.44)		(-1.37)	(1.17)
Instown		-0.082***	0.008		0.028	-0.029
		(-4.16)	(1.21)		(0.49)	(-1.36)
Shortsell		0.898***	0.353***		2.854***	1.769***
		(8.84)	(7.94)		(7.08)	(10.11)
Voldisc8K		2.500***	-0.190		2.984***	-1.314***
		(13.34)	(-1.41)		(4.77)	(-3.41)
MediaStories		0.069***	0.064***		0.059***	0.079***
		(6.19)	(11.59)		(2.76)	(7.08)
EdgarSearch_lag			0.775***			
			(133.81)			
Bloomberg_lag						0.698***
						(102.85)
Observations	45,619	45,619	45,619	29,597	29,597	29,597
Adj R-Squared	0.874	0.879	0.882	0.669	0.680	0.732
Firm FE	Yes	Yes	No	Yes	Yes	No
Industry FE	No	No	Yes	No	No	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes

This table presents an analysis of the relation between the focal firm's and a matched peer firm's abnormal price movements at the earnings announcement, and subsequent information acquisition about the focal firm. Columns 1 to 3 (4 to 6) examine information acquisition on Edgar (Bloomberg). All variables are defined in Appendix B. All continuous variables are winsorized at the 1% and 99% levels to limit the influence of outliers. t-statistics based on standard errors clustered by Firm are in parentheses. Levels of significance are presented as follows: *p<0.1; **p<0.05; ***p<0.01.

Table 8
Dynamic Effects

Panel A. Rank Transition Matrix

	1	2	3	4	5	6	7	8	9	10
1	18.6%	17.9%	15.5%	12.7%	9.9%	8.3%	5.9%	4.1%	3.8%	3.3%
2	17.7%	16.2%	15.5%	13.3%	10.5%	8.9%	6.4%	4.5%	3.8%	3.2%
3	15.7%	15.1%	15.1%	13.8%	11.7%	9.3%	7.1%	5.3%	4.1%	2.9%
4	12.9%	13.2%	13.4%	13.9%	13.5%	11.1%	8.6%	5.6%	4.2%	3.6%
5	10.7%	10.8%	11.0%	13.3%	14.1%	13.1%	10.9%	7.0%	5.5%	3.6%
6	7.9%	8.8%	9.4%	10.8%	13.7%	14.9%	14.7%	10.1%	6.2%	3.6%
7	5.8%	6.4%	7.4%	8.5%	11.3%	14.3%	18.6%	15.7%	7.6%	4.5%
8	4.8%	4.8%	5.9%	6.7%	7.3%	9.8%	15.9%	22.9%	15.2%	6.6%
9	4.2%	4.8%	4.7%	4.8%	5.4%	6.3%	8.1%	16.0%	29.8%	15.8%
10	3.8%	4.2%	3.9%	3.8%	4.0%	5.6%	5.6%	7.3%	17.3%	44.8%

Panel B. Abnormal Price and Information Acquisition over Time

	EdgarSearch	EdgarSearch_{t+2}	EdgarSearch_{t+3}	Bloomberg	Bloomberg_{t+2}	Bloomberg_{t+3}
	(1)	(2)	(3)	(4)	(5)	(6)
AbnPrice	0.034^{***}	0.021^{***}	0.016^{**}	0.089^{***}	0.071^{***}	0.086^{***}
	(5.43)	(3.27)	(2.46)	(3.94)	(3.13)	(3.69)
Observations	87,493	79,139	74,938	59,895	55,776	54,150
Adj R-Squared	0.881	0.875	0.875	0.723	0.723	0.723
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes

This table examines the persistence of abnormal price movements and information acquisition over time. Panel A presents a rank transition matrix of abnormal price movements, using the Edgar Downloads Sample. Each row (column) represents a decile rank of *AbnPrice* this quarter (next quarter). The value in Cell(i,j) represents the percentage of firms with decile rank i this quarter that have decile rank j next quarter. Panel B presents an analysis of the relation between abnormal price movement at the earnings announcement and information acquisition on Edgar (columns 1 to 3) and Bloomberg (columns 4 to 6) over time. All variables are defined in Appendix B. All continuous variables are winsorized at the 1% and 99% levels to limit the influence of outliers. t-statistics based on standard errors clustered by Firm are in parentheses. Levels of significance are presented as follows: *p<0.1; **p<0.05; ***p<0.01.