

Mispricing and the Demand for Fundamental Information

By

Samuel S. Anderson

B.S. Accounting, Brigham Young University (2017)

MAcc, Brigham Young University (2017)

S.M. Management Research, Massachusetts Institute of Technology (2021)

Submitted to the Sloan School of Management
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Management

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2022

© Massachusetts Institute of Technology 2022. All rights reserved.

Author
Department of Management
April 26, 2022

Certified by
Eric C. So
Sloan Distinguished Professor of Management
Thesis Supervisor

Certified by
Joseph P. Weber
George Maverick Bunker Professor of Management
Thesis Supervisor

Accepted by
Catherine Tucker
Sloan Distinguished Professor of Management
Professor, Marketing
Faculty Chair, MIT Sloan PhD Program

Mispricing and the Demand for Fundamental Information

by

Samuel S. Anderson

Submitted to the Department of Management
on April 26, 2022, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Management

Abstract

I provide evidence that investor demand for accounting information intensifies following nonfundamental shocks to prices. Using quasi-exogenous variation in security prices due to forced mutual fund sales, I find that mispricing triggers an increase in the consumption of accounting information, especially among institutional investors. This increase in information consumption subsequently predicts both the speed and extent to which prices return to their pre-shock levels, as well as price informativeness around future earnings events. Taken together, these findings not only demonstrate that mutual fund flow-induced mispricing shapes investors' information consumption, but also highlight the useful role of accounting information in enhancing the informational efficiency of securities markets following temporary mispricing.

Thesis Supervisor: Eric C. So

Title: Sloan Distinguished Professor of Management

Thesis Supervisor: Joseph P. Weber

Title: George Maverick Bunker Professor of Management

Contents

1	Introduction	6
2	Related Literature	13
2.1	Information Consumption	13
2.2	Mutual Fund Flows	15
3	Data and Sample Construction	16
3.1	Mutual fund flows	16
3.2	Information Consumption	18
3.3	Descriptive Statistics	19
4	Results	20
4.1	Mispricing and Information Consumption	20
4.2	Mispricing, Information Consumption, and Price Discovery	25
4.3	Mispricing, Information Consumption, and Price Discovery around Earnings Events	30
5	Conclusion	32
6	Appendix A	51
	Variable Description	51
	EDGAR log cleaning	57
	IP Matching and Investor Classification	58
7	Appendix B: Supplementary Analyses	60
	Additional Tests	60
	Tests using Wardlaw (2020) Flow-to-Volume measure	65

List of Figures

1	Figure 1	39
2	Figure 2	40
3	Figure 3	41
4	Figure 4	55
5	Figure 5	55
6	Figure 6.1	56
7	Figure 6.2	56

1 Introduction

I study whether mispricing affects investors’ consumption of accounting information. Accounting plays a central role in financial markets and provides investors with a means of estimating firms’ fundamental value. The Financial Accounting Standard Board (FASB) states in its strategic plan that it seeks to serve “a common group of stakeholders that have a compelling interest in promoting financial reporting that accurately and neutrally reflects economic activity and promotes the efficiency of capital markets” ([FASB 2015](#)), suggesting that accounting systems are designed, in part, to help facilitate efficient price discovery. Thus, understanding both what drives the consumption of accounting information and its effect on market outcomes, such as price discovery, is of particular importance. Empirically evaluating these links, however, is challenging, due to the endogenous nature of information consumption and security prices, which are jointly determined (Grossman and Stiglitz, [1980](#)), as well as the confounding effect of firms’ fundamentals.¹

In this paper, I examine two related questions: (i) whether temporary, nonfundamental shocks to prices as a result of mutual fund flows affect investors’ consumption of accounting information and (ii) whether this consumption then predicts the path of price discovery for mispriced firms. In doing so, I build on studies that document a relation between the consumption of accounting information and returns (Drake et al., [2016](#); Drake et al., [2020](#)) and attempt to provide causal evidence of prices, and more specifically, mispricing, as significant drivers of investors’ information consumption.

I predict that nonfundamental price shocks will trigger an increase in the consumption of accounting information by drawing the attention of specialized investors, particularly those who did not previously track the firm. A key tenet of efficient markets is that price inefficiencies will be competed away by sophisticated investors (arbitrageurs) searching for profitable investments. The high dimensionality of this task often leads arbitrageurs to specialize (Shleifer and Vishny, [1997](#)). These specialists are attracted to strategies in which expected returns are high, such that the marginal benefit of investing covers the costs of

¹In addition, only recently have researchers been able to obtain direct measures of investor attention and information consumption through the use of novel data sources, such as Google, Yahoo! Finance, EDGAR, and Bloomberg (Da et al., [2011](#); Lawrence et al., [2017](#); Drake et al., [2015](#); Ben-Rephael et al., [2017](#))

information acquisition and implementation (Grossman and Stiglitz, 1980).

In my setting, mutual fund forced sales temporarily depress prices, increasing expected returns, *ceteris parabus*, and potentially drawing the attention of specialized investors by providing a salient signal. I predict that, upon identifying a potentially mispriced firm, specialized investors will increase their consumption of accounting reports. This prediction hinges on two implicit assumptions supported by prior literature: (i) specialized investors focus on a subset of firms and, are therefore less likely to possess extensive prior knowledge of the fundamentals of the newly shocked firms, including a detailed understanding of their most recent accounting filings, and (ii) specialized investors glean value from consuming accounting information. Specialists face resource constraints that may prevent them from being informed at the time of a price shock.² In addition, research suggests that even pure arbitrage opportunities may carry risk and that specialized investors are exposed to idiosyncratic risk, since pursuing narrow strategies leaves them under-diversified (De Long et al., 1990; Shleifer and Vishny, 1997; Abreu and Brunnermeier, 2003). Accounting information provides value to these investors by allowing them to identify potential winners and losers and manage their exposure to different risks (Richardson et al., 2010; Piotroski and So, 2012).

To test my hypotheses, I measure mispricing using mutual fund flow-induced price pressure, an approach pioneered by Coval and Stafford (2007) and refined by Edmans et al. (2012) and Khan et al. (2012). The intuition behind this mutual fund flow measure of mispricing is that investor redemptions create liquidity shocks for mutual funds. If these shocks are large enough, the mutual fund may be forced to sell a portion of its positions, exerting downward price pressure on the stocks in its portfolio. Aggregating the effects of these liquidity shocks across all funds' holdings, it is possible to derive a firm-level measure of price pressure due solely to mutual fund flows. Because the price pressure arises due to shocks at the fund level, it is less likely to be correlated with firm-specific risk.³ Indeed,

²Many researchers have linked investor specialization and limited attention. For instance, Merton (1987) presents a model in which investors choose to specialize due to their inability to acquire information about all firms (See also Hirshleifer and Teoh, 2003; Van Nieuwerburgh and Veldkamp, 2009; Van Nieuwerburgh and Veldkamp, 2010). In addition, a large accounting literature suggests the existence of non-trivial disclosure processing and awareness costs (see Blankespoor et al. (2020) for a review).

³If investors wished to trade on the basis of firm-specific information, it would be more beneficial to trade directly in the underlying. Moreover, in creating the measure, Edmans et al. (2012) screen out international

studies show that firms subject to extreme mutual fund flow-induced price pressure (“fire sale” firms) experience steep price drops followed by predictable reversals over the next 12 to 24 months, suggesting that this pressure is temporary and is not driven by changes in risk.

I focus on mutual fund flow-driven mispricing for two reasons. First, mutual fund fire sales can have an economically large impact on prices. In my sample, the average fire sale firm experiences a price drop of 5.2% in the quarter that the fire sale occurs. Second, it is important to identify movements in price that are unrelated to fundamentals in order to avoid the confounding effects of fundamentals on information consumption and prices as documented in prior studies such as Drake et al. (2015). Given these points, as well as the support in the literature for the quasi-exogenous nature of mutual fund flows as a shock to prices, I believe that using mutual fund flow-driven mispricing provides a well-identified and economically significant setting in which to study the effects of mispricing on investors’ consumption of accounting information.

Following Drake et al. (2015), I measure the consumption of accounting information using investors’ access of accounting reports hosted on the Electronic Data Gathering, Analysis, and Retrieval system (EDGAR), an online system maintained by the Securities and Exchange Commission (SEC) that contains a comprehensive collection of all financial reports filed with the SEC. In particular, the SEC maintains a daily server log, which records every server request, or click, made on EDGAR. Each request is registered to a unique IP address and provides a highly accurate timestamp as well as a record of which form was accessed.⁴ In an effort to foster research efforts on the use of EDGAR, the SEC publicly released these server logs for the period spanning 2003 to 2017. Using these logs, I construct a firm-level measure of investors’ consumption of accounting information by counting the number of unique requests for a firm each day.⁵

I find that mutual fund fire sales increase investors’ consumption of accounting information in the subsequent quarter by an additional 1.4% (approximately 24 additional down-

and sector mutual funds with holdings concentrated in a single market or industry, further reducing the likelihood that fund flows are correlated with changes in firm-specific risk.

⁴See Drake et al. (2015) for a comprehensive description of the EDGAR data. Using a FOIA request, they first obtained this data for a sample of firms from 2008 to 2011.

⁵I attempt to screen out sources of noise in the server logs created by web crawlers and other irrelevant activity following Ryans (2017). A more detailed description of this measure and the filtering process can be found in Section 3.2 and in Appendix A, respectively.

loads) compared to non-fire sale firms. To provide context for this result, I find that, on average, firms in my sample experience approximately 131 additional downloads in the 30-day window surrounding an earnings announcement compared to the 30-day non-earnings period.⁶ Thus, the effect of fire sales on information consumption is about 18% of the effect of an earnings announcement, consistent with prices, and mispricing, representing an economically significant driver of investors' consumption of accounting information.

I then perform several tests designed to capture the mechanism that mispricing increases information consumption through its ability to draw the attention of specialized investors.⁷ First, I examine how mispricing shapes information consumption across investors with varying levels of expertise. This provides both a plausibility check and a natural placebo test, since I expect institutional (retail) investors more (less) closely approximate specialized investors. Moreover, research generally supports the idea that sophisticated, institutional investors compete away price inefficiencies, whereas retail investors often trade for other, nonfundamental reasons. As such, I expect that following a shock to prices, institutional investors will be more active consumers of fundamental information. To test this, I link IP addresses in the EDGAR logs to specific users using IP registration data obtained from the American Registry for Internet Numbers (ARIN) and separate EDGAR users into institutional investors (e.g., investment banks, hedge funds, and other financial institutions) and retail investors (e.g., those using a home router).⁸ I find that mispricing increases information consumption for institutional investors by 2.2% (roughly 25% of the effect of an earnings announcement on institutional attention), but that information consumption for retail investors remains unchanged. This result suggests that mispricing increases the consumption of accounting information—primarily among expert investors—and is consistent

⁶To calculate the effect of earnings announcements on downloads, I perform a simple comparison of means. I total up the number of downloads in the 30-day window surrounding an earnings announcement [-15,15] and compare that to the average number of downloads in the 30-day windows immediately surrounding this period ([-45,-16] and [16,45]). For robustness, I repeat this analysis with more granular windows and find economically similar results.

⁷EDGAR data provides an interesting setting to understand investors' information consumption decisions in that all activity can be linked to both a unique IP address and a specific filing. This feature of the data allows me to study how mispricing shapes the information consumption of investors with varying levels of sophistication as well as how it shapes demand for specific accounting reports.

⁸All IP addresses must be registered with ARIN. I obtain bulk WhoIs data from ARIN, which includes all current IP registrations and their owners. Details of these matching and classification processes can be found in [Appendix A](#)

with mispricing attracting the attention of specialized investors.

Second, I show that mispricing attracts the attention of *specific types* of institutional investors, particularly institutions with less historical knowledge of the firm and fewer portfolio holdings. I do so by separating information consumption into downloads originating from new and existing users based on their IP addresses and search history. Specifically, I classify new users as those that search for a firm in the current quarter, but not in the prior quarter. I find that mispricing increases information consumption for new institutional investors by 2.0%, but that mispricing has no effect on information consumption for existing institutional investors. Next, I count the number of unique portfolio holdings across 13F filers and classify filers in the bottom and top terciles of holdings as small and large institutional investors, respectively. I find that mispricing increases information consumption for small institutional investors by 2.4%, but that mispricing has no statistical effect on information consumption for large institutional investors. Finally, I find that among new investors, the effect is more pronounced for smaller investors. Taken together, these results are consistent with mispricing attracting the attention of new, specialized investors.

Finally, I show that mispricing spurs increases in ownership among investors that did not previously hold the firm (extensive margin). Using 13F filings, I identify a new investor as anyone reporting an ownership stake in the current period but not in the prior period. Summing across filers, I obtain a firm-quarter measure of the number of new investors and find that mutual fund flow-induced mispricing increases the number of new investors by approximately 3.9%, consistent with mispricing attracting new institutional investors to the firm. Overall, these tests lend credibility to the proposed mechanism that mispricing increases the consumption of accounting information through its ability to attract new, specialized investors.

Having documented that mispricing increases the consumption of accounting information, I go on to examine the role of information consumption in price discovery, following a mutual fund fire sale. Specifically, I study price efficiency in terms of (i) the extent to which prices incorporate fundamental information and (ii) the speed at which prices incorporate information. It is possible that as new investors consume and trade on the fundamental information in accounting reports, they provide liquidity to firms experiencing nonfundamental shocks

and help to alleviate mispricing (Coval and Stafford, 2007, Greenwood and Thesmar, 2011). As such, I predict that, following a fire sale, the consumption of accounting information will be positively associated with both the extent to which prices return to pre-fire sale levels as well as the speed at which these reversals occur.

I test this prediction in three ways. First, I rank firms into deciles based on information consumption in the period following a fire sale and examine subsequent return patterns for firms with varying levels of post-fire sale information consumption. I find that, following a fire sale, mispriced firms in the highest decile of information consumption experience 12-month buy-and-hold market-adjusted returns of 7.5%, compared to 3.0% for the median firm, suggesting that information consumption positively predicts return reversals for mispriced firms. Second, I use a hazard model to examine the likelihood that a firm makes a full recovery over the next 90, 180, and 365 days. I find that, following a fire sale, mispriced firms in the highest decile of information consumption are 17%, 21%, and 10% more likely to make a full recovery over the next 90, 180, and 365 days, respectively. Third, I evaluate the timeliness of returns using an intra-period timeliness (IPT) measure developed in prior literature.⁹ I find that following a fire sale, mispriced firms in the highest decile of information consumption exhibit 3% higher intra-period timeliness (IPT), relative to the median firm. Collectively, these results provide consistent evidence that information consumption positively predicts both the extent and speed of return reversals for mispriced firms.

In my final set of tests, I examine whether information consumption plays a role in price discovery around earnings events for mispriced firms. Earnings events represent an opportunity for mispriced firms to attract investor attention which might help to alleviate undervaluation. Mispriced firms that have had very little attention prior to earnings may experience larger price movements upon the release of earnings as the market reacts to both new information and pre-existing undervaluation. In contrast, mispriced firms that receive more attention in the periods prior to earnings may have relatively smaller reactions since much of the undervaluation will have been corrected previously. If this is true, following a fire

⁹I measure the intra-period timeliness (IPT) of returns over the 12 months following a fire sale. IPT is an area under the curve metric that measures cumulative returns at regular intervals to determine the percentage of total cumulative returns realized at each interval. Specifically, I use the adjusted IPT measure proposed by Blankespoor et al. (2018).

sale, I predict that firms with high amounts of information consumption prior to the earnings announcement will have prices that lead earnings to a greater extent as well as lower ERCs at the time of the event. I find evidence consistent with these predictions. In particular, I find that for firms in the highest decile of information consumption prior to the event, the relation between unexpected earnings and pre-earnings returns nearly doubles and that ERCs are roughly half as large, compared to the median firm. This is in contrast to prior studies such as Drake et al. (2015) who find that, in general, firms with high information consumption prior to earnings events have larger ERCs.

I continue to find evidence that mispricing increases the consumption of accounting information using the Wardlaw (2020) flow-to-volume alternative measure of mutual fund flow-induced price pressure and after including a robust set of controls. This mitigates concerns that my results are driven purely by mechanical correlations with past returns or potential selection biases that arise in the construction of the mutual fund flow measure as discussed by Wardlaw (2020) and Berger (2020). In addition, I continue to find economically similar results when I restrict my analysis to forms 10-K and 10-Q, which have fixed supply. A detailed description of these tests and their corresponding results can be found in Appendix 7.¹⁰

Overall, my study contributes to the literature by providing novel evidence that mispricing increases the consumption of accounting information. In doing so, I advance our understanding of what drives investors to search for fundamental information contained in accounting reports. Moreover, by studying information consumption in response to plausibly exogenous variation in security prices, I remove the confounding effect that changes in firms' fundamentals are driving both prices and information consumption and attempt to provide evidence consistent with a causal interpretation. Finally, my results show that information acquisition facilitates price discovery for mispriced firms in the periods following a fire sale as well as in the days leading up to future information events such as earnings announcements. As a result, my findings speak to the literature on price formation following nonfundamental

¹⁰I do find a reduction in the magnitude of the effect of mispricing on information consumption when using the *FlowToVolume* measure, however, the effects continue to be economically significant. I also note that, although tests of the role of information consumption in the earnings-return relationship produce qualitatively similar results, they are statistically insignificant at the 10% level.

shocks to prices (Lee, 2001; Shleifer and Vishny, 2011; Duffie, 2010) as well as the literature on the earnings-return relation (Kothari and Sloan, 1992; Drake et al., 2015).

2 Related Literature

2.1 Information Consumption

Despite a vast literature linking accounting information to capital market outcomes (see Kothari, 2001), only recently have researchers obtained large-scale data on investors’ actual consumption of accounting reports. Most notable is a study by Drake et al. (2015), who use EDGAR data obtained from the SEC to study investor downloads of accounting reports and document that investors’ consumption of accounting information increases around corporate events and following negative return patterns.¹¹ Since then, a number of studies have used EDGAR data to study investors’ consumption of accounting information. For instance, Drake et al. (2016) investigate the usefulness of historical financial reports and find that historical reports give context to current information conveyed in earnings as well as following large price jumps (crashes). Lee et al. (2015) find that investors’ EDGAR search patterns reveal similarities between firms, beyond standard industry classifications, and Drake et al. (2017) note a relationship between EDGAR search patterns and return comovement.

Apart from providing researchers with access to large-scale data on the actual use of accounting information, another notable feature of the EDGAR data is that all activity can be traced to unique IP addresses. Although the SEC partially masks these IP addresses, researchers have begun to de-anonymize these IP addresses, allowing them to sort information consumption into different groups of investors and study their behaviors more directly. For instance, Drake et al. (2020) separate EDGAR users into retail and more sophisticated institutional investors and find that institutional search activity predicts returns and is a leading indicator of portfolio holdings. Bernard et al. (2020) find that peer firms derive

¹¹Earlier work by Da et al. (2011) provides a direct measure of investor attention using Google searches. Similarly, Ben-Rephael et al. (2017) provides a direct measure of attention for institutional investors using Bloomberg search intensity. Using a Yahoo! Finance field experiment, Lawrence et al. (2017) show that prominent ad placement increases investor attention. Although these studies represent major improvements in the ability to directly measure investor attention, Drake et al. (2015) were the first to directly measure investors’ use of *accounting* information.

value from the information contained in competitors’ accounting reports. Chen et al. (2020) link EDGAR activity to specific 13F institutions and document a relationship between performance and search patterns for insider filings. Dyer (2021) localizes IP addresses and finds that investors exhibit a local bias, searching more for firms that are geographically close. Crane et al. (2018) identifies IP addresses belonging to hedge funds and finds that funds that acquire more EDGAR data tend to outperform their peers. Similarly, Bowles (2020) explores the relationship between hedge funds’ information consumption and short sale constraints.

Perhaps most similar to my paper is a working paper by Brunner and Ungeheuer (2019), who study whether salient returns attract investor attention. To disentangle salience from returns, they examine the timing of EDGAR activity. Specifically, they examine firms that announce earnings after the market is closed and find that EDGAR activity is greater for firms with large earnings surprises, relative to those with small earnings surprises, and that this difference does not materialize at the time of the announcement but when markets open and the returns are realized. Although related, my work differs along three important dimensions. First, I specifically focus on how *nonfundamental* movements in prices affect information consumption. This is a critical distinction as not all salient returns can be divorced from changes in firms’ fundamentals. Moreover, studies have shown that firms’ timing of earnings announcements is endogenous (DeHaan et al., 2015). By focusing on mutual fund flow-induced price pressure, I attempt to hold fundamentals constant and circumvent potential endogeneity concerns that changes in fundamentals or firm activities are driving changes in information consumption. Second, I break down information consumption by investor type and search history and provide an alternative channel—mispricing attracts the attention of new specialized investors. In doing so, I provide additional details regarding who consumes accounting information, following nonfundamental shocks to prices. Third, I examine the role of the consumption of accounting information in price discovery for firms that are affected (unaffected) by nonfundamental price pressure and show that information consumption is positively associated with various dimensions of price discovery.

2.2 Mutual Fund Flows

Early studies, such as the work of Scholes (1972) and Shleifer and Vishny (1992), suggest that liquidity shocks can lead to significant price pressure for firms, even absent changes in the underlying economics. Coval and Stafford (2007) note that the transparency of open-ended mutual funds and their reliance on external capital make them a good setting for studying nonfundamental price pressure in equity markets. Specifically, they note that investor redemptions create liquidity shocks for mutual funds that, if large enough, may force funds to sell a portion of their existing positions, often at discounted prices (fire sales). Moreover, because this price pressure arises due to liquidity shocks at the mutual fund level, it is unlikely to be correlated with firm attributes. Aggregating these liquidity-driven sales across mutual funds, they derive a firm-level measure of price pressure due solely to mutual fund fire sales and find that fire sale firms experience a sharp declines in price, followed by return reversals in subsequent periods, consistent with fire sales causing temporary, nonfundamental price pressure (mispricing).

Building on Coval and Stafford (2007), Edmans et al. (2012) note that, during a fire sale, managers may exercise discretion in which securities to offload or hold. Edmans et al. (2012) refine the Coval and Stafford (2007) measure, which uses shocked mutual funds' actual sales of securities and instead estimate the measure using shocked mutual funds' *hypothetical* sales of securities. In doing so, they remove the potentially confounding effects of managerial information present when using actual sales. The development of these quasi-exogenous measures of mutual fund flow price pressure have helped financial economists to overcome issues of endogeneity inherent in prices and has contributed to the rapid expansion in research on market feedback effects. As a result, a large and growing literature, which documents that security prices and mispricing affect corporate actions, is built upon these measures.¹²

¹²For instance, Edmans et al. (2012) find that undervaluation, as a result of mutual fund fire sales, leads to an economically significant increase in the threat of takeover. Foucault and Fresard (2014) show that firms reduce investment in response to nonfundamental shocks to peer firms' stock prices (see also Dessaint et al., 2019). Khan et al. (2012) and Hau and Lai (2013) document that mispricing can significantly impact firms' decision and ability to raise capital. Ali et al. (2011) find that insiders influence the timing of their option grants in response to mispricing.

In addition to analyzing firms’ investment and financing decisions, researchers have also explored whether mispricing impacts firms’ information environment. Sletten (2012) finds that firms increase voluntary disclosure following mispricing due to restatement spillover effects. Abramova et al. (2020) find that firms respond to shocks in institutional investors’ portfolio returns by altering disclosures. Jayaraman and Wu (2020) show that managers use prices to learn about optimal CAPEX investment and that this relationship strengthens following nonfundamental shocks to prices. They go on to document that mispricing increases the likelihood that managers issue a CAPEX forecast. Similarly, Jiang et al. (2021) show that mispricing significantly affects firms’ information production and decisions to issue guidance.¹³ Finally, Lee and So (2017) find that analysts tend to cover undervalued firms, resulting in predictable improvements in future performance. Likewise, Sulaeman and Wei (2019) find that a subset of analysts issue price correcting recommendations following mutual fund fire sales and help stabilize markets. Although differing in their methodologies and execution, these studies provide consistent evidence that mispricing affects the production of information.

3 Data and Sample Construction

My study examines whether mispricing drives investors’ consumption of accounting information. To test this, I measure mispricing using mutual fund flow-driven price pressure and information consumption using EDGAR downloads. I discuss each of these measures below.

3.1 Mutual fund flows

I measure mispricing using mutual fund flow-driven price pressure, following the methodology of Edmans et al. (2012). Specifically, I compute mutual fund flow-induced price pressure on a quarterly basis using data on mutual fund holdings from Thomson Reuters and mutual fund flows from CRSP. The measure is constructed as follows.

¹³Two other concurrent working papers, Heater et al. (2017) and Kadach (2016), perform similar analyses and produce similar results.

First, I calculate fund flows as a percentage of prior net assets by measuring the quarterly change in net assets adjusted for returns:

$$Outflow_{j,t} = -\frac{F_{j,t}}{TNA_{j,t-1}}, \quad (1)$$

where $j = (1...m)$ represents mutual funds, t indexes time in quarters, and

$$F_{j,t} = TNA_{j,t} - TNA_{j,t-1}(1 + RET_{j,t}) \quad (2)$$

is the dollar value of fund j 's flow. Next, I match fund flows to prior period holdings and apply flows evenly across holdings to obtain hypothetical flow-induced sales. I sum this across all funds for which outflows are greater than 5% to obtain hypothetical flow-induced sales for each firm-quarter:¹⁴

$$MFFlow_{i,t} = \left(\sum_{j=1}^m \frac{F_{j,t} s_{i,j,t-1}}{VOL_{i,t}} \middle| Outflow_{j,t} \geq 5\% \right), \quad (3)$$

where $i = (1...n)$ represents individual firms and $VOL_{i,t}$ is the dollar trading volume of stock i in quarter t and

$$s_{i,j,t-1} = \frac{SHARES_{i,j,t-1} \times PRC_{i,t-1}}{TA_{j,t-1}} \quad (4)$$

is the fraction that firm i represents of fund j 's portfolio. In essence, one can think of this measure as follows.

$$MFFlow_{i,t} = \left(\sum_{j=1}^m \frac{DollarHypotheticalVolume_{j,i,t}}{DollarVolume_{i,t}} \middle| Outflow_{j,t} \geq 5\% \right) \quad (5)$$

such that firms with high amounts of $MFFlow$ are experiencing abnormally high amounts of fund flow-driven trading as a percentage of all trading volume. Following the literature, I classify all firm-quarters in the highest decile of $MFFlow$ as fire sales. To validate that I have constructed the measure appropriately, I replicate a central result of Coval and Stafford

¹⁴Edmans et al. (2012) note that using hypothetical flow-induced sales is superior to using actual sales since using actual sales may contaminate the measure with security-level fundamentals due to the informational content of managers' selection of which firms to sell.

(2007) and Edmans et al. (2012) in Figure 1, which plots cumulative market-adjusted returns around the fire sale for an equal-weighted portfolio of fire sale firms. In the figure, we see that fire sale firms exhibit a rapid decline in cumulative abnormal returns at the time of the fire sale followed by a slow reversal, consistent with the findings of Coval and Stafford (2007) and Edmans et al. (2012).

[Insert Figure 1 Here]

3.2 Information Consumption

I measure information consumption using the EDGAR server logs to capture all EDGAR search activity from 2003 to 2017. These logs record each time a user clicks a link within the company filings page, registering a unique IP address, a time stamp, and the form or information requested. For simplicity, I use the words “requests” and “downloads” interchangeably. To reduce the amount of noise in the data, I screen out automated requests following the literature (Ryans, 2017). These and all additional cleaning steps are described in Appendix A. After applying these filters, I am left with a sample of roughly 800 million downloads.

I link this download data to CRSP and Compustat using linking tables provided by the SEC and WRDS. After merging, I am left with roughly 470 million downloads.¹⁵ Next, I aggregate across users and forms to obtain a measure of total downloads for a given firm-day. I then aggregate across time to obtain a measure of total downloads for a given firm-quarter, yielding a final attention sample containing 215,219 firm-quarters and representing 8,674 unique firms, with the average (median) firm-quarter receiving approximately 1,793 (901) downloads.

¹⁵The SEC uses CIK as a unique firm identifier and provides a linking table for mapping to Compustat. Having linked firms to Compustat, I use the linking table provided by WRDS to map to CRSP. The linking results in a substantial loss in the sample. This could be due to at least two reasons. First, the SEC hosts filings for nonpublic entities that might not appear in Compustat. Second, the mapping is not perfectly static and may at times lose coverage. Overall, my level of attrition comports with that of prior studies.

[Figure 2](#) plots information consumption over time and reveals that investor downloads have generally increased over time.¹⁶

[Insert [Figure 2](#) Here]

One notable feature of the EDGAR data is that all activity can be traced to unique IP addresses. Although IP addresses are partially masked to protect users’ identity, it is possible to determine the organization to which they belong. Chen et al. ([2020](#)) suggest a method for de-masking these IP addresses by cross-examining the EDGAR server logs with those of another widely used website.¹⁷ Following their methodology, I de-mask IP addresses and link them to organizations using IP address data obtained from the American Registry for Internet Numbers (ARIN). As a result, I can separate EDGAR users into institutional investors, such as investment banks, hedge funds, and other financial institutions, as well as retail investors, such as those using a home router, and study information consumption decisions across each of these investor types. A more complete description of this process can be found in [Appendix A](#).

3.3 Descriptive Statistics

[Table 1](#) presents descriptive statistics for my final sample, which contains 215,219 firm-quarter observations, representing 8,647 unique firms over the period starting in January 2003 and extending to July 2017. All variables are described in [Appendix A](#). Variables exhibiting excessive skewness are log-transformed to more closely approximate a normal distribution and continuous variables are winsorized at the first and 99th percentiles to reduce the impact of small denominators and extreme values. The mean (median) firm in my sample receives approximately 1,793 (901) download requests in a given quarter.

¹⁶[Figure 2](#) also reveals what appears to be a gap in the coverage for 2006. The SEC reports that the server log file containing 2006 download information became corrupted and much of the data from this year is missing. In an official statement, the agency comments: “Due to certain limitations, including the existence of lost or damaged files, the information assembled by DERA may not capture all SEC.gov website traffic.”

¹⁷Drake et al. ([2020](#)) and other studies also note that organizations tend to buy large IP blocks and, as a result, hiding the final octet does little to hide the owner of the IP address.

[Insert **Table 1** Here]

4 Results

4.1 Mispricing and Information Consumption

I begin my study by focusing on how nonfundamental price shocks affect the consumption of accounting information. I predict that these shocks will lead to an increase in information consumption. This prediction is rooted in the idea that investors are constantly searching for profitable investment opportunities, but have limited resources such that they are not likely to possess a detailed understanding of all firms' fundamentals. Mutual fund forced sales push prices down temporarily, increasing expected returns while both investment and information acquisition costs remain fixed, resulting in a potentially profitable investment opportunity. These mispriced firms will exhibit signals (e.g. salient price movements, shifting price-fundamental ratios, direct tracking of mutual fund flows, etc.) that attract the attention of new, specialized investors. Upon identify a potentially mispriced firm, I expect that these investors will use the fundamental information contained in accounting reports to determine the extent of any mispricing and maximize investment efficiency.¹⁸

There are three sources of tension for my main hypotheses. First, investors tracking the firm may have already performed their due diligence and obtained an estimate of fundamental value. As a result, shocks to prices may influence investors' trading behavior without affecting their decision to acquire additional accounting information. Second, mispricing may be difficult to detect such that it does not immediately attract the attention of specialized investors. Third, shocks to prices that occur in the absence of new accounting information may prompt research into current events or changes in the broader economy, rather than

¹⁸Although it is difficult to disentangle these signals, I attempt to do so in supplemental analyses. First, I examine the effect of information consumption for fire sale firms with large price declines relative to those that experience small price declines. I find that information consumption is nearly 3 times larger for those with large price shocks. In addition, I examine whether information consumption increases for firms that shift to extreme quantiles of price-fundamental ratios such as BTM, CFP, and EP. Again, I find that information consumption increases incrementally for these firms.

potentially stale accounting information. Thus, information consumption following mispricing may be unrelated to accounting reports or flow through other outlets, such as Twitter, Google, or Bloomberg.

In my first test, I examine whether mispricing affects investors' consumption of accounting information. Because EDGAR downloads tend to be serially correlated, I employ a first differences model, as shown in equation (6) below,

$$\Delta \text{LnAtt}_{iq+1} = \beta_1 \Delta \text{Mispricing}_{iq} + \sum_k \beta_k \Delta \text{Controls}_{iq+1} + \gamma_t + \varepsilon_{iq} \quad (6)$$

where Mispricing_{iq} is measured using mutual fund flows ($MFFlow$), and fire sales ($FireSale$). This specification allows me to both control for serial correlation and allows me to scale downloads since small log differences approximate percentage changes. I expect the result to be strongest in fire sale firms, as these are the firms experiencing the most extreme mutual fund flow-driven price pressure. I include quarter-year fixed effects and cluster standard errors by firm and year. [Table 2](#) reports the results of estimating this regression.

[Insert [Table 2](#) Here]

In [Table 2](#), I find that mispricing leads to an increase in EDGAR activity using both mutual fund flows and fire sales. In column (1), I proxy for mispricing using mutual fund flows and find that a one standard deviation change in $\Delta Mfflow$ increases the consumption of accounting information by 0.55%. In column (2), I repeat this analysis using mutual fund fire sales and find that fire sales increase information consumption by 1.36% (or approximately 24 additional quarterly downloads). Although this result may appear small initially, I find that firms in my sample experience approximately 131 additional downloads in the 30-day window surrounding an earnings announcement compared to the 30-day non-earnings period such that 24 additional downloads represents roughly 18% of the effect of an earnings announcement. This comparison highlights the important role of mispricing as an economically significant driver of investors' consumption of accounting information.¹⁹

¹⁹In additional tests, I examine what types of documents investors tend to download and find that the

In my next set of tests, I aim to provide support for the mechanism through which mispricing increases information consumption, namely, that mispricing attracts the attention of new, specialized investors with limited prior knowledge of the firm.

I begin by investigating the effect of mispricing on information consumption for investors with varying levels of sophistication under the assumption that institutional investors are more likely to represent specialists. I predict that, following a mutual fund fire sale, specialized investors will be more active consumers of accounting information. This prediction is rooted in the idea that specialized investors are constantly competing for investment opportunities, and upon identifying a potentially profitable investment, such as a firm that has experienced a mutual fund fire sale, will engage in fundamental analysis. Additionally, studying information consumption across investor types provides a nice placebo test, since retail investors' information consumption and trading patterns are often noisy and nonfundamentally-driven (Drake et al., 2020; Barber and Odean, 2008). As such, I do not expect to observe an effect of mispricing on retail investors. To test these predictions, I classify EDGAR downloads into institutional investors (e.g., hedge funds, investment banks, and other financial professionals) and retail investors (e.g., those using a home router) based on the unique IP address associated with each download. Summing across firm-quarters, I obtain measures of retail and institutional consumption, which I call $LnRetailAtt_{iq}$ and $LnInstAtt_{iq}$, respectively, and repeat the previous analysis.

[Insert [Table 3](#) Here]

In [Table 3](#), I find that mispricing increases information consumption for institutional investors but not for retail investors. In terms of economic magnitudes, I find that mutual fund fire sales increase institutional investors' consumption of accounting information by 2.2%, (approximately 25% of the effect of an earnings announcement on institutional information consumption). In contrast, mutual fund fire sales appear to have no effect on retail investors' consumption of accounting information. Together, these findings provide evidence

increase in information consumption is concentrated in forms 10-Q and 8-K. Moreover, investors tend to download both recent (filed within the last 365 days) and historical filings, although the increase is markedly larger for more recent filings. These analyses can be found in [Appendix B: Supplementary Analyses](#).

that the effects of mispricing are concentrated among professional investors, consistent with mispricing attracting the attention of specialized investors.

As a second test of the mechanism, I examine the effects of mispricing on information consumption for *specific types* of institutional investors. In particular, I separate institutional investors along two important dimensions: historical knowledge of the firm and size. I begin by examining investors' historical knowledge. If mispricing increases information consumption through its ability to draw the attention of new specialized investors, I expect the effects of mispricing to be concentrated among institutional investors with less prior knowledge of the firm (extensive margin). To test this, I separate information consumption into downloads originating from new and existing users based on their IP addresses and search history. Specifically, I classify new users as those that search for a firm in the current quarter, but not in the prior quarter. Following this classification, I obtain a measure of information consumption on the extensive and intensive margin, which allows me to identify the information consumption of new and existing institutional investors. Using this measure, I re-estimate the regression in equation (6). The results can be found in [Table 4](#).

In [Table 4](#) Panel A, I find that mispricing increases the consumption of accounting information on the extensive margin but has comparatively less of an effect on the intensive margin. Importantly, I find that mutual fund fire sales increase information consumption among new, institutional investors by approximately 1.95%, but appear to have no effect on the information consumption decisions of existing institutional investors. In addition, there appears to be no effect on retail investors along either the extensive or the intensive margin. Together, these findings are consistent with the idea that mispricing increases information consumption by drawing the attention of new, specialized investors.

I now turn to consider the size of institutional investors. Research suggests that specialized investors actively seek to discover mispricing in order to capture excess returns. These specialists tend to follow narrow strategies and, as a result, tend to have fewer holdings than other institutional investors. Moreover, small institutional investors are less likely to possess prior information about a large set of firms, track broad indexes, or be engaged in other passive management strategies. This is in contrast to large institutional investors. As a result, I expect the increase in information consumption following a fire sale will be concentrated

among smaller institutions. To test this, I count the number of unique portfolio holdings across 13F filers and classify filers in the bottom and top terciles of holdings as small and large institutional investors, respectively.²⁰

I find that mispricing increases information consumption for small institutional investors by 2.4%, but that it appears to have little effect on information consumption for large institutional investors, consistent with my predictions. [Table 4](#) Panel B displays the results of these tests. Finally, I find that among new investors, the effect is more pronounced for smaller institutions. Overall, these findings suggest that nonfundamental price shocks increase the consumption of accounting information, especially among new, specialized investors.²¹

[Insert [Table 4](#) Here]

Finally, to further triangulate the mechanism, I study whether mutual fund fire sales attract new investors to the firm by analyzing investors' portfolio holdings. Drake et al. (2020) suggest that information consumption is a leading indicator of investors' portfolio holdings. Thus, if mispricing increases information consumption by attracting new investors to the firm, I expect to observe an increase in ownership among new investors (extensive margin). Using 13F filings, I identify a new investor as any filer that reports an ownership stake in the current period but not in the prior period. Summing across all investors in a given quarter, I obtain a firm-level measure of the number of new investors, $NewInvestors_{iq}$, which I log transform. I then run the regression in equation (7),

$$\Delta \ln NewInvestors_{iq+1} = \beta_1 \Delta FireSale_{iq} + \sum_k \beta_k \Delta Controls_{iq+1} + \gamma_t + \varepsilon_{iq+1} \quad (7)$$

[Insert [Table 5](#) Here]

²⁰See [Mispricing and Information Consumption](#) for a more detailed explanation of this classification process. The cutoff for the number of holdings for institutional investors in the bottom (top) tercile is 75 (329). In untabulated analyses, I repeat these tests using a subjective breakpoint of 100 and continue to find similar results.

²¹One alternative mechanism is that increased information consumption following a fire sale may be due to an increase in information production by the firm. In robustness tests, I attempt to control for this using downloads of forms 10-K and 10-Q, which have fixed supply, and continue to find economically similar results.

In [Table 5](#), I find that mispricing increases the number of new investors that hold the firm. Specifically, in column (1), I find that mutual fund fire sales increase the number of new investors that report an ownership stake in the firm by approximately 3.9%, consistent with mutual fund fire sales attracting the attention of investors on the extensive margin.²² Overall, the tests in [Tables 3-5](#) lend credibility to the mechanism that mispricing increases the consumption of accounting information through its ability to attract the attention of new, specialized investors.²³

4.2 Mispricing, Information Consumption, and Price Discovery

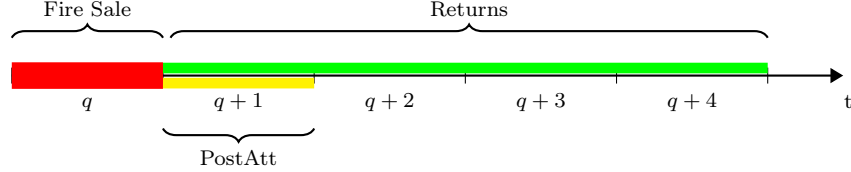
Having documented that mispricing drives investors' consumption of information, I proceed to examine the link between information consumption and price discovery for mispriced firms. I investigate two aspects of price discovery: (i) the extent to which prices revert to pre-fire sale levels and (ii) the speed with which this occurs. An advantage of studying the role of information consumption for price discovery in the context of mispricing is that there is a clear role for fundamental analysis and a clear path for returns (predictable reversals). In addition, the effects of mutual fund flow-induced mispricing can be economically significant, leading to negative externalities, such as increased threat of takeover and cost of capital (Edmans et al., [2012](#); Hau and Lai, [2013](#)). A quick recovery may help to mitigate these threats. Coval and Stafford ([2007](#)) and Edmans et al. ([2012](#)) show that firms that undergo a fire sale tend to experience return reversals over the next 24 months. Because fire sales do not alter firms' existing fundamentals, the consumption of accounting information may help accelerate reversals and alleviate mispricing. As such, I expect that information consumption will be positively associated with both the extent to which prices return to pre-fire sale levels as well as the speed of reversals.

To test this hypothesis, I measure investors' consumption of accounting information in the quarter following a fire sale and examine whether increased information consumption

²²These results are robust to measuring the dependent variable using number of shares as well as dollar volume of investment.

²³One caveat of this test is that it does not directly link those specific 13F filers who download following a fire sale to those who purchase securities, but rather provides evidence that a specific group (13F Filers) engage in both activities.

is associated with the likelihood and rate of recovery. To facilitate interpretation, I rank leading information consumption into deciles and transform it such that it ranges between $[-1, 1]$ similar to Drake et al. (2020). I refer to this variable as $PostAtt_{iq+1}$. The following picture illustrates this timing.



I note two limitations of these tests. First, in an ideal experiment, I could observe investors' intentions and separate downloads driven by fire sales from those driven by other factors to isolate the effect of *mispricing-induced* information consumption on price discovery. Second, due to the low frequency with which mutual fund holdings are updated (quarterly), it is difficult to pin down the timing of nonfundamental price pressure, impeding identification of the appropriate window in which to measure information consumption following mispricing. I acknowledge these limitations and take steps to address them. Most notably, in my main tests, I implement a difference-in-differences design with a robust suite of controls, which allows me to study how the relationship between information consumption and returns varies for firms that are affected (unaffected) by fire sales.²⁴ [Table 6](#) presents the results of estimating the following regression:

$$\begin{aligned}
 Ret_{i[t,t+\tau]} = & \beta_1 FireSale_{iq} + \beta_2 PostAtt_{iq+1} + \beta_3 PostAtt_{iq+1} \times FireSale_{iq} \\
 & + \sum_k \beta_k Controls_{iq+1} + \gamma_t + \delta_i + \varepsilon_{iq+1}
 \end{aligned} \tag{8}$$

[Insert [Table 6](#) Here]

In [Table 6](#) Panel A, I find that information consumption following a fire sale predicts return reversals for mispriced firms. I measure returns as buy-and-hold market-adjusted

²⁴In addition, I construct a new monthly mfflow measure using quarterly holdings and monthly TNA and measure information consumption in the month following mispricing. Although untabulated, I find qualitatively similar results, mitigating concerns about the timing of information consumption.

returns, where Ret_τ represents the period $[t, t + \tau]$ such that column (1) estimates three-month returns, column (2) estimates six-month returns, and so on. I am most interested in the coefficient on $PostAtt \times FireSale$, which indicates the relation between information consumption and return reversals for fire sale firms with high information consumption, relative to the median fire sale firm. I find that information consumption following a fire sale is positively associated with six-month and 12-month returns. In terms of economic magnitudes, I find that firms in the top decile of information consumption following a fire sale tend to experience 12-month buy-and-hold market-adjusted returns of 7.5%, compared to 3.0% for the median firm. I also find a positive coefficient on $FireSale$, suggesting that firms experience return reversals following a fire sale, consistent with findings of Edmans et al. (2012), and a positive coefficient on $PostAtt$ in the short-term, consistent with findings using Google search volume in the IPO setting (Da et al. (2011) but differing from studies using EDGAR data (Drake et al., 2015; Drake et al., 2016).

Illustrating this result, Figure 3 plots the return patterns for firms with high, medium, and low download intensity following a fire sale. We can see that all firms exhibit similar return patterns during the pre-event and event windows, mitigating potential concerns about selection biases. In the post-event window, firms with the most information consumption tend to recover to pre-fire sale levels in a more timely manner.

[Insert Figure 3 Here]

In Table 6 Panel B, I separate information consumption by investor type and find that information consumption by institutional investors predicts return reversals for mispriced firms but that information consumption by retail investors does not. Columns (1) through (4) display the results for retail investors, and columns (5) through (8) display the results for institutional investors. Examining the coefficient on $PostAtt \times FireSale$ in columns (5) through (8), I find that information consumption following a fire sale is positively associated with returns across all time windows. In contrast, estimates for retail investors are statistically insignificant. Specifically, I find that firms in the top decile of institutional information consumption following a fire sale experience 12-month buy-and-hold market-adjusted returns

of 6.1%, compared to 2.8% for the median firm. These results are consistent with specialized investors being the primary consumers of accounting information following mispricing and playing an important role in price discovery.

To further test the extent to which prices return to pre-fire sale levels, I employ a hazard model approach. To do this, I measure fire sale returns and count the number of days until the cumulative return (including the fire sale drop in price) exceeds zero for five consecutive days.²⁵ Using the hazard model, I estimate the likelihood of making a 90-day, 180-day, or 365-day recovery. The results are presented in [Table 7](#).

[Insert [Table 7](#) Here]

In [Table 7](#), I find that, following a fire sale, mispriced firms in the highest decile of information consumption are 17%, 21%, and 10% more likely to fully recover over the next 90, 180, and 365 days, respectively. All coefficients are displayed as likelihood ratios, such that coefficients greater than one indicate an increased likelihood of realizing an outcome — in this case, recovery. As in previous analyses, I test my model using institutional and retail information consumption. I find that, following a fire sale, mispriced firms in the highest decile of institutional information consumption are 9%, 8%, and 6% more likely to make a full recovery over the next 90, 180, and 365 days, respectively. In contrast, mispriced firms in the highest decile of retail information consumption are no more likely to recover across any of the windows.²⁶ Similar to [Table 6](#), these results are consistent with a positive association between information consumption, following a fire sale, and the likelihood that prices return to pre-fire sale levels. Moreover, this positive association across both short-term and long-term windows speaks to the speed at which prices are reverting, suggesting more timely recoveries.

Finally, in an effort to further reinforce the previous results, I perform an additional test aimed at studying whether information consumption following a fire sale is positively

²⁵The choice of five days is subjective and is intended to ensure that there is a true recovery, rather than capturing volatility which causes the firm to experience extreme returns. My results are robust to changes in this requirement.

²⁶In fact, the result in column (2) suggests that mispriced firms in the highest decile of retail information consumption are 6% less likely to make a short-term recovery.

associated with the speed of price discovery. I proxy for the speed of price discovery using an adjusted measure of intra-period timeliness (IPT) proposed by Blankespoor et al. (2018)²⁷. I measure adjusted IPT in the 12 months following a fire sale and estimate the regression in equation (9) below.

$$\begin{aligned} AdjIPT_{i[t,t+12]} = & \beta_1 FireSale_{iq} + \beta_2 PostAtt_{iq+1} + \beta_3 PostAtt_{iq+1} \times FireSale_{iq} \\ & + \sum_k \beta_k Controls_{iq+1} + \gamma_t + \delta_i + \varepsilon_{iq+1} \end{aligned} \quad (9)$$

[Insert **Table 8** Here]

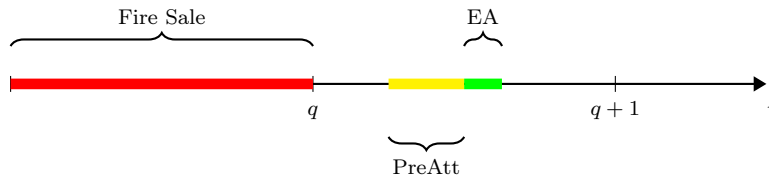
In **Table 8**, I find that information consumption following a fire sale is positively associated with the speed of return reversals. Examining the coefficient on $PostAtt_{iq+1} \times FireSale_{iq}$ in column (1), I find that, following a fire sale, mispriced firms in the highest decile of information consumption exhibit 3% higher IPT, relative to the median firm. In columns (2) and (3), I re-estimate the model using retail and institutional information consumption, respectively. As in other tests, I document a link between institutional information consumption and the speed of return reversals but find no such relation for retail information consumption.

Taken together, the findings in Tables 6 - 8 suggest that the consumption of accounting information plays an important role in the price discovery process for mispriced firms and is positively associated with both the extent to which prices return to pre-fire sale levels as well as the speed of return reversals. Note that, because I measure both information consumption and returns in the periods following the shock to prices, these tests face many potential endogeneity concerns, and, as such, it is not my intent to suppose a causal interpretation of these results.

²⁷IPT measures the percentage of a firms' cumulative return that has been realized at regular intervals and computes an area under the curve. Specifically, $IPT_{iq} = \sum_{m=1}^{11} (BH_{1,m}/BH_{1,12}) + 0.5$, where BH is the firm's market-adjusted buy-and-hold return (see Butler et al. (2007) for details on the construction of the IPT measure.) Blankespoor et al. (2018) propose an improvement to the IPT measure that penalizes firms for overreactions and drop observations with cumulative returns <2% to reduce noise. I follow their construction of the AdjIPT measure.

4.3 Mispricing, Information Consumption, and Price Discovery around Earnings Events

In my final set of tests, I examine the link between information consumption and price informativeness around earnings announcements through the lens of mispriced firms. Research suggests that information consumption increases around earnings events and that EDGAR searches prior to an earnings event are associated with larger ERCs and lower PEAD (Drake et al., 2015). To my knowledge, however, very little is known about the role of information consumption in the earnings-return relation for mispriced firms. This relation may differ since mispriced firms are presumably undervalued heading into an earnings event such that there may be pricing of both new and existing information at the time of the earnings event. On the one hand, increased information consumption in the periods leading up to earnings events may be associated with larger ERCs, consistent with prior findings. On the other hand, information consumption prior to earnings may lead to significant pre-earnings price discovery related to existing information, alleviating undervaluation and resulting in lower ERCs at the time of the announcement. To examine this, I restrict my analysis to firms that experience a fire sale. I then pinpoint the date of the next earnings announcement using Compustat and IBES and measure information consumption as the number of downloads in the 10 days $[-12, -2]$ leading up to the event. The following picture illustrates this timing.



As before, I rank pre-earnings information consumption (*PreAtt*) into deciles and transform it such that it ranges between $[-1, 1]$ and examine the link between information consumption and both pre-earnings prices and earnings response coefficients (ERCs) for mispriced firms. I calculate seasonal unexpected earnings (SUE) as the earnings surprise based on a seasonal random walk. [Table 9](#) presents the results of estimating the following regres-

sions.

$$BHMAR_{i[-12,-2]} = \beta_1 SUE_{iq+1} + \beta_2 PreAtt_{iq+1} + \beta_3 PreAtt_{iq+1} \times SUE_{iq+1} + \sum_k \beta_k Controls_{iq+1} + \gamma_t + \sigma_{ind} + \varepsilon_{iq+1}, \quad (10)$$

$$BHMAR_{i[-1,1]} = \beta_1 SUE_{iq+1} + \beta_2 PreAtt_{iq+1} + \beta_3 PreAtt_{iq+1} \times SUE_{iq+1} + \sum_k \beta_k Controls_{iq+1} + \gamma_t + \sigma_{ind} + \varepsilon_{iq+1}, \quad (11)$$

where equation (10) uses pre-announcement returns to study the extent to which prices lead earnings and (11) uses announcement returns to study ERCs.

[Insert [Table 9](#) Here]

In [Table 9](#) Panel A, I find that pre-earnings information consumption is positively associated with the extent to which prices lead earnings for mispriced firms. This result suggests that, when information consumption is high in the days preceding an earnings announcement, prices tend to more accurately reflect fundamental information. In columns (2) and (3), I re-estimate the model using retail and institutional information consumption and find that institutional information consumption is positively associated with pre-earnings price informativeness.

In [Table 9](#) Panel B, I find that pre-earnings information consumption is negatively associated with the earnings response coefficient for mispriced firms. This result suggests that, when information consumption is high in the days preceding an earnings announcement, the market response tends to be muted. Notably, the negative relation between pre-earnings information consumption and ERCs is opposite that documented by Drake et al. (2015), suggesting that price discovery around earnings may differ for mispriced firms. Assuming earnings are somewhat persistent, one potential explanation for this finding is that pre-earnings information consumption alleviates undervaluation such that there is less pricing of existing information at the time of the announcement. Together, the results in Panels A and B suggest that information consumption plays a role in price formation around earnings

events, and that this role may differ following nonfundamental price shock.

5 Conclusion

A vast literature examines the extent to which accounting information is reflected in and shapes security prices. Nevertheless, there is less causal evidence on what drives the consumption of accounting information. In this paper, I build upon prior work and attempt to provide causal evidence that mutual fund flow-induced mispricing drives investors to consume information contained in accounting reports. To do this, I study information consumption in response to plausibly exogenous variation in security prices due to mutual fund flows, removing the confounding effect that changes in firms' fundamentals are driving both prices and information consumption. I find that mutual fund fire sales increase the consumption of accounting reports by 1.4%. This increase is nearly one-fifth of the effect of an earnings announcement on download activity for the quarter. The effect is concentrated in institutional investors, particularly new and small investors, consistent with the idea that mispricing increases information consumption by attracting the attention of specialized investors with less prior knowledge of the firm. In addition, I find that information consumption, following a mutual fund fire sale, predicts both the extent and speed at which prices return to pre-shock levels, as well as price informativeness around future earnings events, highlighting the important role of accounting information in facilitating price discovery, especially for mispriced firms.

References

- Abramova, Inna, John E Core, and Andrew Sutherland (2020). “Institutional investor attention and firm disclosure”. In: *The Accounting Review* 95.6, pp. 1–21.
- Abreu, Dilip and Markus K Brunnermeier (2003). “Bubbles and crashes”. In: *Econometrica* 71.1, pp. 173–204.
- Ali, Ashiq, Kelsey D Wei, and Yibin Zhou (2011). “Insider trading and option grant timing in response to fire sales (and purchases) of stocks by mutual funds”. In: *Journal of Accounting Research* 49.3, pp. 595–632.
- Barber, Brad M and Terrance Odean (2008). “All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors”. In: *The Review of Financial Studies* 21.2, pp. 785–818.
- Ben-Rephael, Azi, Zhi Da, and Ryan D Israelsen (2017). “It depends on where you search: Institutional investor attention and underreaction to news”. In: *The Review of Financial Studies* 30.9, pp. 3009–3047.
- Berger, Elizabeth (2020). “Selection Bias in Mutual Fund Fire Sales”. In: *Available at SSRN* 3011027.
- Bernard, Darren, Terrence Blackburne, and Jacob Thornock (2020). “Information flows among rivals and corporate investment”. In: *Journal of Financial Economics* 136.3, pp. 760–779.
- Blankespoor, Elizabeth, Ed deHaan, and Ivan Marinovic (2020). “Disclosure processing costs, investors’ information choice, and equity market outcomes: A review”. In: *Journal of Accounting and Economics* 70.2-3, p. 101344.
- Blankespoor, Elizabeth, Ed deHaan, and Christina Zhu (2018). “Capital market effects of media synthesis and dissemination: Evidence from robo-journalism”. In: *Review of Accounting Studies* 23.1, pp. 1–36.
- Bowles, Boone (2020). “Conditional Information Acquisition”. PhD thesis. The University of North Carolina at Chapel Hill.
- Brunner, Fabian and Michael Ungeheuer (2019). *The power of prices: Information, trade, and salient returns*. Tech. rep. Working Paper.

- Butler, Marty, Arthur Kraft, and Ira S Weiss (2007). “The effect of reporting frequency on the timeliness of earnings: The cases of voluntary and mandatory interim reports”. In: *Journal of Accounting and Economics* 43.2-3, pp. 181–217.
- Chen, Huaizhi et al. (2020). “IQ from IP: Simplifying search in portfolio choice”. In: *Journal of Financial Economics* 138.1, pp. 118–137.
- Coval, Joshua and Erik Stafford (2007). “Asset fire sales (and purchases) in equity markets”. In: *Journal of Financial Economics* 86.2, pp. 479–512.
- Crane, Alan, Kevin Crotty, and Tarik Umar (2018). “Do hedge funds profit from public information”. In: *Working Paper*.
- Da, Zhi, Joseph Engelberg, and Pengjie Gao (2011). “In search of attention”. In: *The Journal of Finance* 66.5, pp. 1461–1499.
- De Long, J Bradford et al. (1990). “Positive feedback investment strategies and destabilizing rational speculation”. In: *the Journal of Finance* 45.2, pp. 379–395.
- DeHaan, Ed, Terry Shevlin, and Jacob Thornock (2015). “Market (in) attention and the strategic scheduling and timing of earnings announcements”. In: *Journal of Accounting and Economics* 60.1, pp. 36–55.
- Dessaint, Olivier et al. (2019). “Noisy stock prices and corporate investment”. In: *The Review of Financial Studies* 32.7, pp. 2625–2672.
- Drake, Michael S, Darren T Roulstone, and Jacob R Thornock (2015). “The determinants and consequences of information acquisition via EDGAR”. In: *Contemporary Accounting Research* 32.3, pp. 1128–1161.
- (2016). “The usefulness of historical accounting reports”. In: *Journal of Accounting and Economics* 61.2-3, pp. 448–464.
- Drake, Michael S et al. (2017). “The comovement of investor attention”. In: *Management Science* 63.9, pp. 2847–2867.
- Drake, Michael S et al. (2020). “Is there information content in information acquisition?” In: *The Accounting Review* 95.2, pp. 113–139.
- Duffie, Darrell (2010). “Presidential address: Asset price dynamics with slow-moving capital”. In: *The Journal of Finance* 65.4, pp. 1237–1267.

- Dyer, Travis A (2021). “The demand for public information by local and nonlocal investors: Evidence from investor-level data”. In: *Journal of Accounting and Economics*, p. 101417.
- Edmans, Alex, Itay Goldstein, and Wei Jiang (2012). “The real effects of financial markets: The impact of prices on takeovers”. In: *The Journal of Finance* 67.3, pp. 933–971.
- FASB (2015). <https://www.accountingfoundation.org/strategicplan>. Accessed: 2020-09-10.
- Foucault, Thierry and Laurent Fresard (2014). “Learning from peers’ stock prices and corporate investment”. In: *Journal of Financial Economics* 111.3, pp. 554–577.
- Greenwood, Robin and David Thesmar (2011). “Stock price fragility”. In: *Journal of Financial Economics* 102.3, pp. 471–490.
- Grossman, Sanford J and Joseph E Stiglitz (1980). “On the impossibility of informationally efficient markets”. In: *The American Economic Review* 70.3, pp. 393–408.
- Hau, Harald and Sandy Lai (2013). “Real effects of stock underpricing”. In: *Journal of Financial Economics* 108.2, pp. 392–408.
- Heater, John C, Yukun Liu, and Ben Matthies (2017). “Managerial Response to Non-Fundamental Price Shocks”. In: *Available at SSRN 3029520*.
- Hirshleifer, David and Siew Hong Teoh (2003). “Limited attention, information disclosure, and financial reporting”. In: *Journal of Accounting and Economics* 36.1-3, pp. 337–386.
- Jayaraman, Sudarshan and Joanna Wu (2020). “Should I stay or should I grow? Using voluntary disclosure to elicit market feedback”. In: *The Review of Financial Studies* 33.8, pp. 3854–3888.
- Jiang, Jinglin, Vikram Nanda, and Steven Chong Xiao (2021). “Stock-market disruptions and corporate disclosure policies”. In: *Journal of Corporate Finance* 66, p. 101762.
- Kadach, Igor (2016). “Management Earnings Forecasts during Price Pressure: Evidence from Mutual Fund Trades”. In: *Available at SSRN 2798693*.
- Khan, Mozaffar, Leonid Kogan, and George Serafeim (2012). “Mutual fund trading pressure: Firm-level stock price impact and timing of SEOs”. In: *The Journal of Finance* 67.4, pp. 1371–1395.
- Kothari, SP (2001). “Capital markets research in accounting”. In: *Journal of Accounting and Economics* 31.1-3, pp. 105–231.

- Kothari, SP and Richard G Sloan (1992). “Information in prices about future earnings: Implications for earnings response coefficients”. In: *Journal of Accounting and Economics* 15.2-3, pp. 143–171.
- Lawrence, Alastair, James P Ryans, and Estelle Y Sun (2017). “Investor demand for sell-side research”. In: *The Accounting Review* 92.2, pp. 123–149.
- Lee, Charles MC (2001). “Market efficiency and accounting research: a discussion of ‘capital market research in accounting’ by SP Kothari”. In: *Journal of Accounting and Economics* 31.1-3, pp. 233–253.
- Lee, Charles MC, Paul Ma, and Charles CY Wang (2015). “Search-based peer firms: Aggregating investor perceptions through internet co-searches”. In: *Journal of Financial Economics* 116.2, pp. 410–431.
- Lee, Charles MC and Eric C So (2017). “Uncovering expected returns: Information in analyst coverage proxies”. In: *Journal of Financial Economics* 124.2, pp. 331–348.
- Loughran, Tim and Bill McDonald (2017). “The use of EDGAR filings by investors”. In: *Journal of Behavioral Finance* 18.2, pp. 231–248.
- Merton, Robert C (1987). “A simple model of capital market equilibrium with incomplete information”. In:
- Piotroski, Joseph D (2000). “Value investing: The use of historical financial statement information to separate winners from losers”. In: *Journal of Accounting Research*, pp. 1–41.
- Piotroski, Joseph D and Eric C So (2012). “Identifying expectation errors in value/glamour strategies: A fundamental analysis approach”. In: *The Review of Financial Studies* 25.9, pp. 2841–2875.
- Richardson, Scott, Irem Tuna, and Peter Wysocki (2010). “Accounting anomalies and fundamental analysis: A review of recent research advances”. In: *Journal of Accounting and Economics* 50.2-3, pp. 410–454.
- Ryans, James (2017). “Using the EDGAR log file data set”. In: *Available at SSRN 2913612*.
- Scholes, Myron S (1972). “The market for securities: Substitution versus price pressure and the effects of information on share prices”. In: *The Journal of Business* 45.2, pp. 179–211.

- Shleifer, Andrei and Robert Vishny (1992). “Liquidation values and debt capacity: A market equilibrium approach”. In: *The Journal of Finance* 47.4, pp. 1343–1366.
- (1997). “The limits of arbitrage”. In: *The Journal of Finance* 52.1, pp. 35–55.
- (2011). “Fire sales in finance and macroeconomics”. In: *Journal of Economic Perspectives* 25.1, pp. 29–48.
- Sletten, Ewa (2012). “The effect of stock price on discretionary disclosure”. In: *Review of Accounting Studies* 17.1, pp. 96–133.
- Sulaeman, Johan and Kelsey D Wei (2019). “Sell-side analysts and stock mispricing: evidence from mutual fund flow-driven trading pressure”. In: *Management Science* 65.11, pp. 5427–5448.
- Van Nieuwerburgh, Stijn and Laura Veldkamp (2009). “Information immobility and the home bias puzzle”. In: *The Journal of Finance* 64.3, pp. 1187–1215.
- (2010). “Information acquisition and under-diversification”. In: *The Review of Economic Studies* 77.2, pp. 779–805.
- Wardlaw, Malcolm (2020). “Measuring mutual fund flow pressure as shock to stock returns”. In: *The Journal of Finance* 75.6, pp. 3221–3243.

Tables and Figures



Figure 1

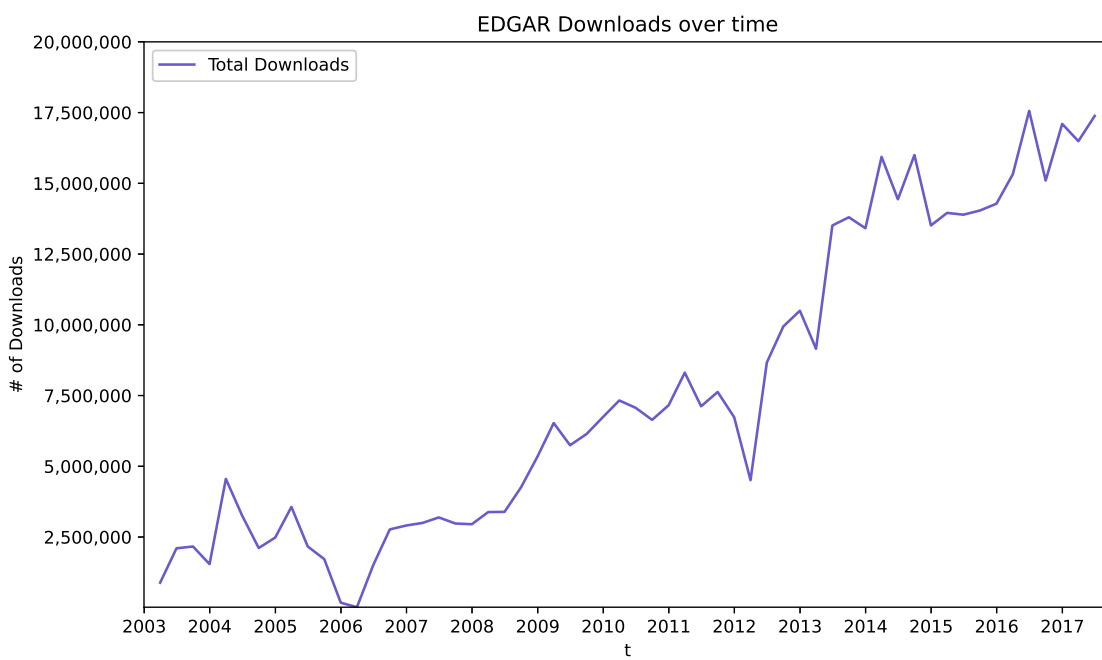


Figure 2

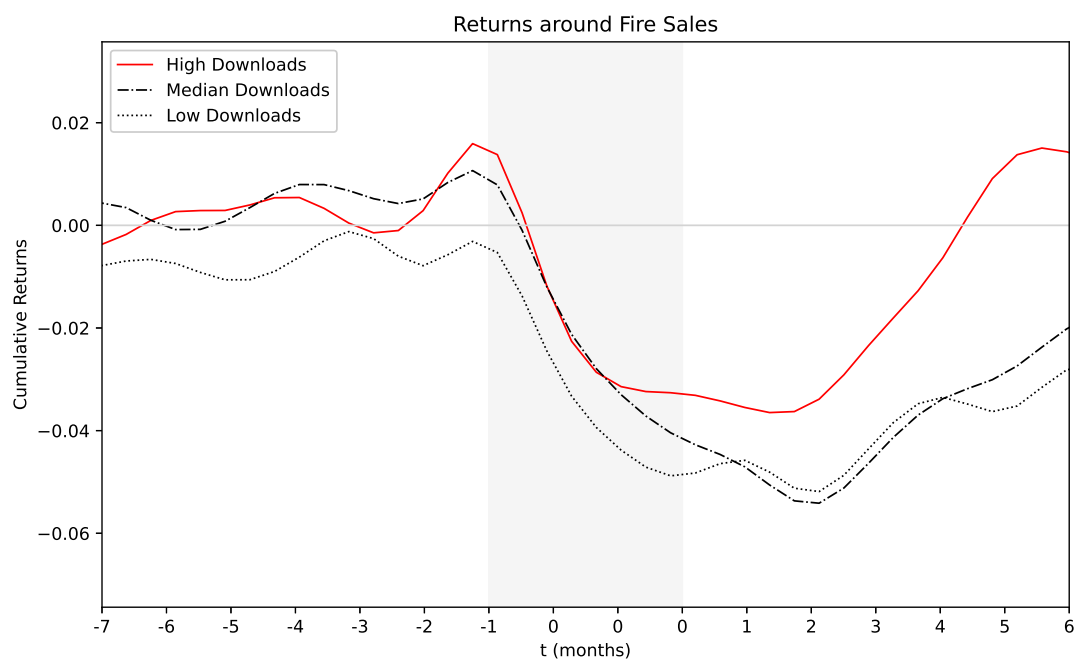


Figure 3

Table 1

Descriptive Statistics. This table presents descriptive statistics for each of the variables used in my analyses. My final sample represents 215,219 firm-quarter observations, representing 8,647 unique firms over the period starting in January of 2003 and extending to July of 2017. All variables are described in further detail in [Appendix A](#). All count and other variables exhibiting excessive skewness are log-transformed to more closely approximate a normal distribution. All ratios and other continuous variables are winsorized at the 1st and 99th percentiles to reduce the impact of small denominators and extreme values.

	N	Mean	Median	Min	Max	Std. Dev.
<i>TotAtt</i>	215,219	1,793.7888	901.00	1.00	2.7e+06	9,245.3417
<i>RetailAtt</i>	215,219	5.9162	2.00	0.00	1,114.00	14.2912
<i>InstAtt</i>	215,219	25.3533	8.00	0.00	6,272.00	86.8786
<i>LnTotAtt</i>	215,219	6.7359	6.80	0.69	14.82	1.2720
<i>LnRetailAtt</i>	215,219	1.2090	1.10	0.00	7.02	1.1104
<i>LnInstAtt</i>	215,219	2.2079	2.20	0.00	8.74	1.4185
$\Delta \text{LnTotAtt}$	194,008	0.0425	0.02	-5.43	7.39	0.4575
$\Delta \text{LnRetailAtt}$	194,008	0.0413	0.00	-4.92	4.73	0.9297
$\Delta \text{LnInstAtt}$	194,008	0.0413	0.00	-4.66	5.79	0.9471
<i>MFFlow</i>	215,219	0.5442	0.20	0.00	6.51	0.9901
<i>FireSale</i>	215,219	0.1037	0.00	0.00	1.00	0.3048
ΔMFFlow	194,008	0.0081	0.00	-6.51	6.51	1.0816
$\Delta \text{FireSale}$	194,008	0.0016	0.00	-1.00	1.00	0.3608
<i>ScreenBTM</i>	216,350	0.0492	0.00	0.00	1.00	0.2163
<i>ScreenCFP</i>	216,350	0.0576	0.00	0.00	1.00	0.2330
<i>TotAttHistorical</i>	215,219	681.3447	232.00	0.00	1.5e+06	4,197.2855
<i>TotAttCurrent</i>	215,219	1,112.4441	629.00	0.00	2.3e+06	6,044.5257
<i>TotAttRecent</i>	215,219	668.4907	377.00	0.00	3.6e+05	1,676.1484
<i>RecoverDays</i>	15,082	215.7134	221.00	7.00	365.00	140.0846
<i>Recover</i>	15,082	0.6151	1.00	0.00	1.00	0.4866
<i>BHMAR</i> _[m-3,m]	216,350	0.0113	-0.01	-1.04	18.15	0.2588
<i>BHMAR</i> _[m-6,m]	214,654	0.0305	-0.01	-1.23	64.99	0.4438
<i>BHMAR</i> _[m,m+3]	214,781	0.0109	-0.01	-1.04	18.15	0.2562
<i>BHMAR</i> _[m,m+6]	213,346	0.0232	-0.01	-1.23	64.99	0.4253
<i>Volatility</i>	187,251	0.2243	0.19	0.01	19.12	0.1636
<i>LnTurnover</i>	215,219	0.9677	0.82	0.00	13.87	0.8414
<i>DecileBTM</i>	207,333	5.5138	6.00	1.00	10.00	2.8842
<i>DecileSize</i>	207,333	5.5399	6.00	1.00	10.00	2.8711
<i>LnNumAnalyst</i>	215,219	1.2025	1.10	0.00	3.99	0.9408
<i>InstOwn</i>	186,751	0.5679	0.61	0.00	15.45	0.3410
<i>Leverage</i>	214,681	2.5942	1.09	-10.51	26.88	4.7185
<i>FSCORE</i>	207,333	5.2714	5.00	0.00	9.00	1.6609
<i>ROA</i>	214,995	-0.0023	0.01	-0.28	0.09	0.0508
<i>SUE</i>	206,713	-0.0008	0.00	-0.35	0.31	0.0649
<i>LossFirm</i>	215,219	0.2725	0.00	0.00	1.00	0.4452
<i>Qtr4</i>	215,219	0.2508	0.00	0.00	1.00	0.4334
<i>Persistence</i>	177,486	0.2171	0.13	-2.21	2.88	0.6532
<i>EarningsVol</i>	213,161	0.5193	0.22	0.01	6.11	0.9089

Table 2

Mispricing and the Consumption of Accounting Information. This table presents the results of estimating the regression $\Delta LnAtt_{iq+1} = \beta_1 \Delta Mispricing_{iq} + \sum_k \beta_k Controls_{iq+1} + \sum_j \beta_j \Delta Controls_{iq+1} + \gamma_t + \varepsilon_{iq+1}$ where $\Delta LnAtt_{iq+1}$ represents the change in information consumption in the quarter following a fire sale and γ_t represents year-quarter fixed effects. Controls include returns, volatility, turnover, earnings surprise, analyst following, institutional ownership, size, BTM, FSCORE, leverage, ROA, and indicators for loss firms and fiscal quarter. I proxy for mispricing using mutual fund flows and fire sales, the results of which are displayed in columns (1) and (2) respectively. Refer to [Appendix A](#) for a complete description of all control variables. All continuous variables are winsorized at the 1st and 99th percentiles to limit the effect of outliers and small denominators. I estimate t-statistics using two-way cluster robust standard errors, clustered by firm and year with *, **, and *** indicating statistical significance at less than 10%, 5%, and 1% levels, respectively.

	(1) $\Delta LnAtt_{q+1}$	(2) $\Delta LnAtt_{q+1}$
$\Delta MFFlow_q$	0.00500*** (3.26)	
$\Delta FireSale_q$		0.0136*** (5.43)
$BHMAR_{[m-3,m+3]}$	0.0361*** (3.86)	0.0360*** (3.85)
$Volatility_{[m,m+3]}$	-0.0971*** (-4.35)	-0.0973*** (-4.36)
$LogTurnover_{[m-9,m+3]}$	0.00120 (0.30)	0.00120 (0.30)
SUE_{q+1}	-0.00679 (-0.34)	-0.00656 (-0.33)
$\Delta LnNumAnalyst_{q+1}$	-0.00417** (-2.22)	-0.00417** (-2.23)
$\Delta DecileSize_{q+1}$	-0.0129*** (-5.73)	-0.0129*** (-5.83)
$\Delta DecileBTM_{q+1}$	-0.00427** (-2.41)	-0.00422** (-2.35)
$\Delta FSCORE_{q+1}$	0.00221*** (4.86)	0.00220*** (4.83)
$\Delta Leverage_{q+1}$	0.00482*** (5.11)	0.00482*** (5.11)
$\Delta InstOwn_{q+1}$	-0.0289** (-2.38)	-0.0289** (-2.38)
ΔROA_{q+1}	-0.195*** (-5.98)	-0.194*** (-5.93)
$\Delta Loss_{q+1}$	0.00333 (0.98)	0.00331 (0.98)
$\Delta Qtr4_{q+1}$	0.0575*** (11.13)	0.0575*** (11.14)
R^2	0.769	0.769
N	167,645	167,645
Fixed Effects	Year-Qtr	Year-Qtr
Cluster by	Firm, Year	Firm, Year

Table 3

Fire Sales and Information Consumption Across Investor Types. This table presents the results of estimating the regression $\Delta LnAtt_{iq+1} = \beta_1 \Delta Mispricing_{iq} + \sum_k \beta_k Controls_{iq+1} + \sum_j \beta_j \Delta Controls_{iq+1} + \gamma_t + \varepsilon_{iq+1}$ where γ_t represents year-quarter fixed effects. I separate EDGAR activity into retail and institutional investors and measure information consumption across each type as $LnRetailAtt_{iq+1}$ and $LnInstAtt_{iq+1}$, respectively. Controls include returns, volatility, turnover, earnings surprise, analyst following, institutional ownership, size, BTM, FSCORE, leverage, ROA, and indicators for loss firms and fiscal quarter. Refer to [Appendix A](#) for a complete description of all control variables. All continuous variables are winsorized at the 1st and 99th percentiles to limit the effect of outliers and small denominators. I estimate t-statistics using two-way cluster robust standard errors, clustered by firm and year with *, **, and *** indicating statistical significance at less than 10%, 5%, and 1% levels, respectively.

EDGAR activity across retail and institutional investors		
	(1) $\Delta LnRetailAtt_{q+1}$	(2) $\Delta LnInstAtt_{q+1}$
$\Delta FireSale_q$	0.0069 (0.70)	0.0220*** (3.03)
R^2	0.0612	0.104
N	167,645	167,645
Controls	Y	Y
Fixed Effects	Year-Qtr	Year-Qtr
Cluster by	Firm, Year	Firm, Year

Table 4

Fire Sales and Institutional Consumption - Investor Type. This table presents the results of estimating the regression $\Delta \text{LnAtt}_{it+1} = \beta_1 \Delta \text{FireSale}_{it} + \sum_j \beta_j \Delta \text{Controls}_{it+1} + \gamma_t + \varepsilon_{it+1}$ where $\Delta \text{LnAtt}_{it+1}$ is measured using consumption from institutional investors and γ_t represents year-quarter fixed effects. I classify new investors as users who search for a firm in the current, but not in the previous quarter (extensive margin) and existing investors as users who search for a firm in both the current and previous quarter (intensive margin). I use 13F filings to classify institutions as small and large depending upon the number of unique holdings (bottom tercile and top tercile, respectively). Controls include returns, volatility, turnover, earnings surprise, analyst following, institutional ownership, size, BTM, FSCORE, leverage, ROA, and indicators for loss firms and fiscal quarter. Refer to [Appendix A](#) for a complete description of all control variables. All continuous variables are winsorized at the 1st and 99th percentiles to limit the effect of outliers and small denominators. I estimate t-statistics using two-way cluster robust standard errors, clustered by firm and year with *, **, and *** indicating statistical significance at less than 10%, 5%, and 1% levels, respectively.

Panel A: Investors' Prior Search History				
	New Investors		Existing Investors	
	$\Delta \text{LnInst}_{q+1}$		$\Delta \text{LnInst}_{q+1}$	
$\Delta \text{FireSale}_q$	0.0195*	(2.06)	0.0006	(0.07)
N	167,614		167,614	
R^2	0.090		0.048	
Controls	Y		Y	
Fixed Effects	Year-Qtr		Year-Qtr	
Cluster by	Firm, Year		Firm, Year	
Panel B: Investor Size				
	Small		Large	
	$\Delta \text{LnInst}_{q+1}$		$\Delta \text{LnInst}_{q+1}$	
$\Delta \text{FireSale}_q$	0.0235*	(2.05)	0.00852	(0.89)
N	167,614		167,614	
R^2	0.208		0.401	
Controls	Y		Y	
Fixed Effects	Year-Qtr		Year-Qtr	
Cluster by	Firm, Year		Firm, Year	
Panel C: Investors' Search History & Size				
	New \times Small	New \times Large	Existing \times Small	Existing \times Large
	$\Delta \text{LnInst}_{q+1}$	$\Delta \text{LnInst}_{q+1}$	$\Delta \text{LnInst}_{q+1}$	$\Delta \text{LnInst}_{q+1}$
$\Delta \text{FireSale}_q$	0.0222*	0.0184	0.0120	-0.00528
	(1.87)	(1.68)	(1.64)	(-1.09)
N	167,614	167,614	167,614	167,614
R^2	0.360	0.039	0.116	0.041
Controls	Y	Y	Y	Y
Fixed Effects	Year-Qtr	Year-Qtr	Year-Qtr	Year-Qtr
Cluster by	Firm, Year	Firm, Year	Firm, Year	Firm, Year

Table 5

Fire Sales and Investor Holdings - Evidence from 13F Filings. This table presents the results of estimating the regression $\Delta \text{LnInvest}_{iq+1} = \beta_1 \Delta \text{FireSale}_{iq} + \sum_j \beta_j \Delta \text{Controls}_{iq+1} + \gamma_t + \varepsilon_{iq+1}$ where $\Delta \text{LnInvest}_{iq+1}$ represents the change in the number of investors in the quarter following a fire sale and γ_t represents year-quarter fixed effects. I measure LnNewInvest_{iq} as the number of investors who report holdings for a firm in the current, but not the previous, period (extensive margin) as well as the dollar volume of purchases on the extensive margin in columns (1) and (2), respectively. Controls include returns, volatility, turnover, earnings surprise, analyst following, institutional ownership, size, BTM, FSCORE, leverage, ROA, and indicators for loss firms and fiscal quarter. Refer to [Appendix A](#) for a complete description of all control variables. All continuous variables are winsorized at the 1st and 99th percentiles to limit the effect of outliers and small denominators. I estimate t-statistics using two-way cluster robust standard errors, clustered by firm and year with *, **, and *** indicating statistical significance at less than 10%, 5%, and 1% levels, respectively.

	(1) $\Delta \text{LnNewInvest}_{q+1}$	(2) $\Delta \text{LnNewVolume}_{q+1}$
$\Delta \text{FireSale}_q$	0.0391*** (3.91)	0.159** (2.43)
R^2	0.0971	0.0459
N	159,158	159,158
Controls	Y	Y
Fixed Effects	Year-Qtr	Year-Qtr
Cluster by	Firm, Year	Firm, Year

Table 6

Fire Sales, Information Consumption, and the Magnitude of Return Reversals. This table presents the results of estimating the regression $Ret_{i\tau} = \beta_1 FireSale_{iq} + \beta_2 PostAtt_{iq+1} + \beta_3 PostAtt_{iq+1} \times FireSale_{iq} + \sum_k \beta_k Controls_{iq+1} + \gamma_t + \delta_i + \varepsilon_{iq+1}$ where $PostAtt_{iq+1}$ represents information consumption in the quarter following a fire sale, $Ret_{i\tau}$ represents buy-and-hold market-adjusted returns over the next τ months and γ_t and δ_i represent year-quarter and firm fixed effects, respectively. Controls include lagged attention, prior period returns, volatility, analyst following, institutional ownership, size, BTM, FSCORE, leverage, ROA, and indicators for loss firms and fiscal quarter. Columns (1) through (4) measure returns over the 3, 6, 9, and 12 month periods following a fire sale, respectively. Panel B repeats this analysis for retail and institutional information consumption. Refer to [Appendix A](#) for a complete description of all control variables. All continuous variables are winsorized at the 1st and 99th percentiles to limit the effect of outliers and small denominators. The sample is limited to fire sale firms. I estimate t-statistics using two-way cluster robust standard errors, clustered by firm and year with *, **, and *** indicating statistical significance at less than 10%, 5%, and 1% levels, respectively.

Panel A: All EDGAR Activity								
	(1) <i>Ret</i> ₃	(2) <i>Ret</i> ₆	(3) <i>Ret</i> ₉	(4) <i>Ret</i> ₁₂				
<i>FireSale</i> _q	0.0015 (0.61)	0.0109 (1.57)	0.0237* (1.93)	0.0295** (2.70)				
<i>PostAtt</i> _{q+1}	0.0325*** (4.55)	0.0113 (1.10)	-0.0030 (-0.21)	-0.0186 (-1.23)				
<i>PostAtt</i> _{q+1} × <i>FireSale</i> _q	0.0007 (0.25)	0.0230** (2.16)	0.0396 (1.74)	0.0451* (1.93)				
<i>R</i> ²	0.128	0.156	0.195	0.213				
N	172,826	172,388	171,174	169,289				
Panel B: Retail and Institutional EDGAR Activity								
	Retail				Institutional			
	(1) <i>Ret</i> ₃	(2) <i>Ret</i> ₆	(3) <i>Ret</i> ₉	(4) <i>Ret</i> ₁₂	(5) <i>Ret</i> ₃	(6) <i>Ret</i> ₆	(7) <i>Ret</i> ₉	(8) <i>Ret</i> ₁₂
<i>FireSale</i> _q	0.00171 (0.71)	0.0105 (1.72)	0.0211** (2.23)	0.0263*** (3.15)	0.0025 (0.82)	0.0105 (1.50)	0.0226* (1.91)	0.0280** (2.72)
<i>PostAtt</i> _{q+1}	0.0033 (1.63)	-0.0001 (-0.03)	-0.0042 (-1.13)	-0.0047 (-1.06)	0.0070** (2.86)	0.0007 (0.24)	-0.0043 (-1.02)	-0.012* (-2.13)
<i>PostAtt</i> _{q+1} × <i>Fire</i> _q	0.0020 (0.68)	0.0166 (1.56)	0.0190 (1.76)	0.0205 (1.63)	0.0065* (1.89)	0.0188* (1.85)	0.0289* (1.80)	0.0327* (1.89)
<i>R</i> ²	0.127	0.156	0.194	0.213	0.127	0.156	0.195	0.213
N	172,826	172,388	171,174	169,289	172,826	172,388	171,174	169,289

Table 7

Fire Sales, Information Consumption, and the Likelihood of Recovery. This table presents the results of estimating the hazard model $Recover_{i\tau} = \beta_1 PostAtt_{iq+1} + \sum_k \beta_k Controls_{iq+1} + \varepsilon_{iq+1}$ where $PostAtt_{iq+1}$ represents information consumption in the quarter following a fire sale and . To estimate the hazard model, I count the number of days until a firms' buy-and-hold market-adjusted return, including the loss from the fire sale, is greater than 0 for at least 5 consecutive days. $Recover_{i\tau}$ is an indicator variable equal to one when the previous condition is satisfied. Controls include lagged attention, prior period returns, volatility, analyst following, institutional ownership, size, BTM, FSCORE, leverage, ROA, and indicators for loss firms and fiscal quarter. I estimate the hazard model over various time horizons. Columns (1) through (3), (4) through (6), and (7) through (9) show the results over a 90-day, 180-day, and 365-day horizon for all, retail, and institutional activity, respectively. Refer to [Appendix A](#) for a complete description of all control variables. All continuous variables are winsorized at the 1st and 99th percentiles to limit the effect of outliers and small denominators. The sample is limited to fire sale firms. Coefficients are reported as likelihood ratios. I estimate t-statistics using robust standard errors with *, **, and *** indicating statistical significance at less than 10%, 5%, and 1% levels, respectively.

	DV: 90 Day Recovery			DV: 180 Day Recovery			DV: 365 Day Recovery		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Total	Retail	Inst	Total	Retail	Inst	Total	Retail	Inst
$PostAtt_{q+1}$	1.177** (2.20)	0.944* (-1.78)	1.090* (1.92)	1.211*** (3.29)	1.001 (0.02)	1.075** (2.05)	1.103** (1.97)	1.008 (0.35)	1.057* (1.86)
N	10,328	10,328	10,328	10,328	10,328	10,328	10,328	10,328	10,328
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y
SEs	Robust	Robust	Robust	Robust	Robust	Robust	Robust	Robust	Robust

Table 8

Fire Sales, Information Consumption, and the Speed of Return Reversals. This table presents the results of estimating the regression $AdjIPT_{i[m,m+12]} = \beta_1 FireSale_{iq} + \beta_2 PostAtt_{iq+1} + \beta_3 PostAtt_{iq+1} \times FireSale_{iq} + \sum_k \beta_k Controls_{iq+1} + \gamma_t + \delta_i + \varepsilon_{iq+1}$ where $PostAtt_{iq+1}$ represents information consumption in the quarter following a fire sale and γ_t and δ_i represent year-quarter and firm fixed effects, respectively. IPT is measured as $IPT = \sum_{m=1}^{11} (BH_{1,m}/BH_{1,12}) + 0.5$ where BH is the firm's market-adjusted buy-and-hold return. I follow Blankespoor et al. (2018) in constructing $AdjIPT$. Refer to Section 4.2 for more details. Controls include lagged attention, prior period returns, volatility, analyst following, institutional ownership, size, BTM, FSCORE, leverage, ROA, and indicators for loss firms and fiscal quarter. Columns (1) through (3) measure information consumption using all EDGAR activity, retail investor activity, and institutional investor activity, respectively. Refer to Appendix A for a complete description of all control variables. All continuous variables are winsorized at the 1st and 99th percentiles to limit the effect of outliers and small denominators. The sample is limited to fire sale firms. I estimate t-statistics using two-way cluster robust standard errors, clustered by firm and year with *, **, and *** indicating statistical significance at less than 10%, 5%, and 1% levels, respectively.

	(1) Total	(2) Retail	(3) Institutional
	$AdjIPT_{[m,m+12]}$	$AdjIPT_{[m,m+12]}$	$AdjIPT_{[m,m+12]}$
$FireSale_q$	0.0735* (1.85)	0.0429 (1.11)	0.0697* (1.89)
$PostAtt_{q+1}$	-0.00801 (-0.25)	0.0214 (1.11)	-0.0249 (-0.92)
$PostAtt_{q+1} \times FireSale_q$	0.172** (2.23)	0.0263 (0.49)	0.137* (1.83)
R^2	0.0505	0.0505	0.0505
N	145,961	145,961	145,961
Controls	Y	Y	Y
Fixed Effects	Firm, Year-Qtr	Firm, Year-Qtr	Firm, Year-Qtr
Cluster by	Firm, Year	Firm, Year	Firm, Year

Table 9

Fire Sales, Information Consumption, and Earnings Events. This table presents the results of estimating the regression $BHMAR_{iq+1} = \beta_1 SUE_{iq+1} + \beta_2 PreAtt_{iq+1} + \beta_3 PreAtt_{iq+1} \times SUE_{iq+1} + \sum_k \beta_k Controls_{iq+1} + \gamma_t + \sigma_{ind} + \varepsilon_{iq+1}$ where $PreAtt_{iq+1}$ represents information consumption in the 10 day period $[d-12, d-2]$ prior to an earnings event and γ_t and σ_{ind} represent year-quarter and industry fixed effects, respectively. Controls include lagged attention, prior period returns, volatility, analyst following, institutional ownership, size, BTM, FSCORE, leverage, ROA, earnings volatility, earnings persistence, and indicators for loss firms and fourth quarter reports. Column (1) measures information consumption using all EDGAR activity and columns (2) and (3) measure information consumption using institutional investors' and retail investors' EDGAR activity, respectively. Refer to [Appendix A](#) for a complete description of all control variables. All continuous variables are winsorized at the 1st and 99th percentiles to limit the effect of outliers and small denominators. The sample is limited to fire sale firms. I estimate t-statistics using two-way cluster robust standard errors, clustered by firm and year with *, **, and *** indicating statistical significance at less than 10%, 5%, and 1% levels, respectively.

Panel A: Prices Lead Earnings			
	(1) Total	(2) Retail	(3) Inst
	$BHMAR_{[d-12,d-2]}$	$BHMAR_{[d-12,d-2]}$	$BHMAR_{[d-12,d-2]}$
SUE_{q+1}	0.0327 (1.72)	0.0555* (2.07)	0.0646** (2.27)
$PreAtt_{q+1}$	0.00138 (0.91)	0.00195 (1.51)	0.000570 (0.46)
$PreAtt_{q+1} \times SUE_{q+1}$	0.100* (1.82)	0.0226 (0.63)	0.0676* (1.95)
R^2	0.252	0.250	0.251
N	15,962	15,962	15,962
Controls	Y	Y	Y
Fixed Effects	Firm, Year-Qtr	Firm, Year-Qtr	Firm, Year-Qtr
Cluster by	Firm, Year	Firm, Year	Firm, Year
Panel B: Earnings Response Coefficients			
	(1) Total	(2) Retail	(3) Inst
	$BHMAR_{[d-1,d+1]}$	$BHMAR_{[d-1,d+1]}$	$BHMAR_{[d-1,d+1]}$
SUE_{q+1}	0.152*** (8.43)	0.155*** (5.12)	0.124*** (4.82)
$PreAtt_{q+1}$	-0.000527 (-0.20)	0.000324 (0.31)	0.00112 (0.88)
$PreAtt_{q+1} \times SUE_{q+1}$	-0.0657 (-1.57)	0.0201 (0.68)	-0.0673* (-2.01)
R^2	0.249	0.249	0.250
N	15,962	15,962	15,962
Controls	Y	Y	Y
Fixed Effects	Firm, Year-Qtr	Firm, Year-Qtr	Firm, Year-Qtr
Cluster by	Firm, Year	Firm, Year	Firm, Year

6 Appendix A

Variable Description

Table A1. Description of Variables used in this Study

Variable	Definition
Att	Total EDGAR activity for a given firm-quarter. Obtained from the EDGAR server logs which report each server request made within the EDGAR online system. Computed as $\ln(1 + Requests)$
MFFlow	Mutual fund flow-induced price pressure computed following Edmans et al. (2012). Measured as $MFFlow_{i,t} = \left(\sum_{j=1}^m \frac{F_{j,t} s_{i,j,t-1}}{VOL_{i,t}} \mid Outflow_{j,t} \geq 5\% \right)$ where $i = (1...n)$ represents individual firms and $VOL_{i,t}$ is the dollar trading volume of stock i in quarter t and $s_{i,j,t-1} = \frac{SHARES_{i,j,t-1} \times PRC_{i,t-1}}{TA_{i,t-1}}$
FireSale	An indicator equal to 1 when MFFlows are in the top decile of all MFFlows.
ScreenBTM	An indicator equal to 1 when a firm moves from any quantile to the highest quantile of BTM.
ScreenCFP	An indicator equal to 1 when a firm moves from any quantile to the highest quantile of CFP.
ScreenPE	An indicator equal to 1 when a firm moves from any quantile to the lowest quantile of PE.
AttNew	Total number of unique EDGAR requests for a given firm-quarter for IP addresses that did not search for the firm in the prior quarter. Computed as $\ln(1 + NewRequests)$.
RetailAttNew	Total number of unique EDGAR requests for a given firm-quarter for IP addresses belonging to retail investors that did not search for the firm in the prior quarter. Computed as $\ln(1 + NewRetailRequests)$.
InstAttNew	Total number of unique EDGAR requests for a given firm-quarter for IP addresses belonging to institutional investors that did not search for the firm in the prior quarter. Computed as $\ln(1 + NewInstRequests)$.
AttOld	Total number of unique EDGAR requests for a given firm-quarter for IP addresses that searched for the firm in the prior quarter. Computed as $\ln(1 + OldRequests)$.

Variable	Definition
RetailAttOld	Total number of unique EDGAR requests for a given firm-quarter for IP addresses belonging to retail investors that searched for the firm in the prior quarter. Computed as $\ln(1 + OldRetailRequests)$.
InstAttOld	Total number of unique EDGAR requests for a given firm-quarter for IP addresses belonging to institutional investors that searched for the firm in the prior quarter. Computed as $\ln(1 + OldInstRequests)$.
NewInvest	The number of filers that report holdings of a given firm in the current period, but not in the prior period. Holdings obtained via 13F filings.
OldInvest	The number of filers that increase their holdings in an existing position. Holdings obtained via 13F filings.
RetailAtt	EDGAR activity for retail investors in a given firm-quarter. Retail investors classified using WhoIs data and home router IP address ranges (see Appendix for a more detailed description of this classification process). Computed as $\ln(1 + RetailRequests)$
InstAtt	EDGAR activity for retail investors in a given firm-quarter. Retail investors classified using WhoIs data and fuzzy matching (see Appendix for a more detailed description of this classification process). Computed as $\ln(1 + InstRequests)$.
10K	An indicator equal to 1 for form 10K.
10Q	An indicator equal to 1 for form 10Q.
8K	An indicator equal to 1 for form 8K.
Form4	An indicator equal to 1 for form 4.
Proxy	An indicator equal to 1 for form DEF14A (Proxy Statements).
RecentAtt	Total EDGAR downloads of forms filed within 90 days from the time of the download request. Computed as $\ln(1 + RecentRequests)$
CurrentAtt	Total EDGAR downloads of forms filed within 365 days from the time of the download request. Computed as $\ln(1 + CurrentRequests)$.
HistoricalAtt	Total EDGAR downloads of forms filed more than 365 days prior to the time of the download request. Computed as $\ln(1 + HistoricalRequests)$.
PostAtt	Total EDGAR activity in the quarter following a fire sale. Computed as $\ln(1 + Att_{q+1})$, ranked into deciles each quarter and transformed to range between $[-1, 1]$.

Variable	Definition
PostRetail	EDGAR activity for retail investors in the quarter following a fire sale. Computed as $\ln(1 + RetailAtt_{q+1})$, ranked into deciles each quarter and transformed to range between $[-1, 1]$.
PostInst	EDGAR activity for institutional investors in the quarter following a fire sale. Computed as $\ln(1 + InstAtt_{q+1})$, ranked into deciles each quarter and transformed to range between $[-1, 1]$.
Returns	Buy-and-hold market adjusted returns computed via CRSP. Measured over various return windows.
AdjIPT	A measure of area under the curve for cumulative returns. Computed as $IPT = \sum_{m=1}^{11} (BH_{1,m}/BH_{1,12}) + 0.5$ where BH is the firm's market-adjusted buy-and-hold return. I follow Blankespoor et al. (2018) in constructing <i>AdjIPT</i> .
DaysRecovery	The number of days until a firm experiences buy-and-hold market-adjusted returns, including losses from fire sales, greater than 0 for at least 5 consecutive days.
Recover	An indicator equal to 1 for days \geq DaysRecovery following a fire sale.
PreAtt	Total EDGAR activity in the 10 day period $[-12, -2]$ prior to an earnings event. Computed as $\ln(1 + Att_{[d-12, d-2]})$, ranked into deciles each quarter and transformed to range between $[-1, 1]$.
PreRetail	EDGAR activity for retail investors in the 10 day period $[-12, -2]$ prior to an earnings event. Computed as $\ln(1 + RetailAtt_{[d-12, d-2]})$, ranked into deciles each quarter and transformed to range between $[-1, 1]$.
PreInst	EDGAR activity for institutional investors in the 10 day period $[-12, -2]$ prior to an earnings event. Computed as $\ln(1 + InstAtt_{[d-12, d-2]})$, ranked into deciles each quarter and transformed to range between $[-1, 1]$.
SUE	Unexpected earnings following a seasonal random walk, computed $(EPS_t - EPS_{t-4})/Price_t$.
Volatility	Volatility of stock returns computed as the standard deviation of daily returns over the past 90 days scaled by $\sqrt{90}$.
Analyst Following	Computed as $\ln(1 + analysts)$ where analysts is number of analysts following the firm obtained from IBES.
Size	Computed as $\ln(1 + mve)$ where mve is the market value of equity obtained from CRSP.

Variable	Definition
BTM	Computed as $\ln(1 + btm)$ where btm is the book to market obtained from Compustat and CRSP.
FSCORE	A measure of fundamental performance developed by Piotroski (2000) intended to capture firms' financial health across various dimensions which sums across 9 indicator variables.
Leverage	Computed as $\ln(1 + TA/TE)$ where TA and TE are total assets and total equity, respectively.
Loss	An indicator equal to 1 for firm-quarter observations with net income less than 0.
Qtr4	An indicator equal to 1 for firm-quarter observations in the fourth fiscal quarter.
Institutional Ownership	Computed as $\frac{\sum_j Shares_{ijt} Prc_{ijt}}{Shares_{iq} * Prc_{it}}$ where the numerator represents the aggregate market cap of all institutional holdings of a given firm and the denominator represents the aggregate market cap of the firm.
ROA	Return on assets, computed as net income divided by total assets.
Turnover	Share turnover computed as the mean of the ratio of trading volume to shares outstanding over the previous 12 month period.
Earnings Persistence	The mean of $\hat{\beta}_1$ in the regression $EPS_{iq} = \alpha + \beta_1 EPS_{iq-4} + \varepsilon_{iq}$ estimated over the previous 4 years.
Earnings Volatility	The standard deviation of $EPS_{iq} - EPS_{iq-4}$ estimated over the previous 4 years.

Figure 4 provides a summary of all the steps taken to arrive at the final data sample. Intermediate steps taken to clean and compile many of these datasets are omitted. An example of ARIN IP registration information can be found in Figure 5. Using this IP information I am able to match EDGAR users to specific organizations.

Steps	Source / Data / Methodology
1. Download EDGAR logs	SEC.gov
2. Clean EDGAR logs and remove bots	Ryans (2017)
3. De-anonymize IP addresses	Chan et al. (2020)
4. Match IP addresses to organization names	ARIN WhoIs
5. Match downloads to specific filings	WRDS
6. Match organization names to CRSP IDs	Fuzzy matching – Levenshtein distance
7. Compile full download panel	
8. Compute mutual fund flows	CRSP, Compustat, Reuters; Edmans et al. (2012)
9. Merge datasets to form final panel	

Figure 4

<div> <div> <div>ARIN</div> <div>American Registry for Internet Numbers</div> </div> <div> <div>SEARCH WhoisRWS</div> <div>all requests subject to terms of use</div> <div>advanced search</div> </div> </div>	
<div> <div> <div>ARIN Online</div> <div>enter</div> </div> <div>WHOIS-RWS</div> </div>	
<div>You searched for: 17.145.55.67</div>	
Network	
Net Range	17.0.0.0 - 17.255.255.255
CIDR	17.0.0.0/8
Name	APPLE-WWNET
Handle	NET-17-0-0-1
Parent	
Net Type	Direct Allocation
Origin AS	
Organization	Apple Inc. (APPLEC-1-Z)
Registration Date	1990-04-16
Last Updated	2021-12-14
Comments	
RESTful Link	https://whois.arin.net/rest/net/NET-17-0-0-1

RELEVANT LINKS

- ARIN Whois/Whois-RWS Terms of Service
- Report Whois Inaccuracy
- Search ARIN Whois with RDAP

Figure 5

Figure 6.1 illustrates the fuzzy matching procedure for a hypothetical example using Panera Bread Co. To improve matching results, I match along 3 scores: Standard, Clean, and Non-Dictionary. The Standard score strips out articles and common words and standardizes suffixes such as co, co., comp, comp., company. The Clean score strips out all suffixes. The Non-Dictionary score strips out any words that are in the english dictionary and attempts to match on non-standard words. I compute a composite score using a weighted average of these scores. Figure 6.2 provides an example of the fuzzy matching results from actual output for Talbots Inc.

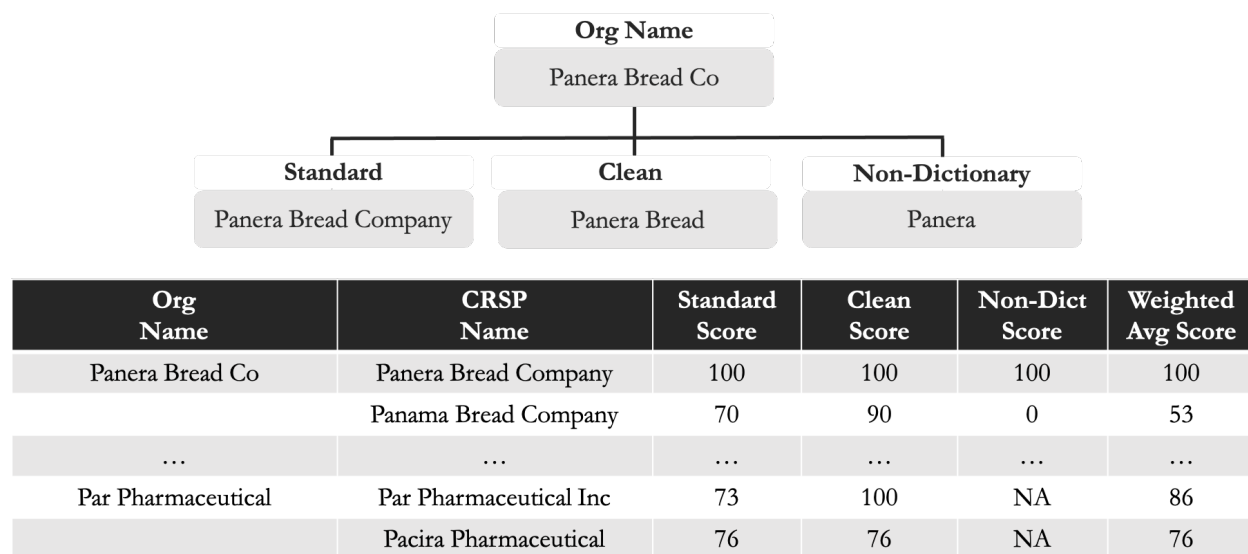


Figure 6.1

company_log	company_crsp	standard_log	standard_crsp	clean_log	clean_crsp	non_dict_log	non_dict_crsp	standard_score	clean_score	non_dict_score
the talbots inc	tel inc	Talbots Incorporated	Tel Incorporated	Talbots	Tel	Talbots	NaN	83	40	0
the talbots inc	tago inc	Talbots Incorporated	Tago Incorporated	Talbots	Tago	Talbots	Tago	86	55	55
the talbots inc	talbots inc	Talbots Incorporated	Talbots Incorporated	Talbots	Talbots	Talbots	Talbots	100	100	100
the talbots inc	talcott james inc	Talbots Incorporated	Talcott James Incorporated	Talbots	Talcott James	Talbots	Talcott	83	60	71
the talbots inc	talk city inc	Talbots Incorporated	Talk City Incorporated	Talbots	Talk City	Talbots	NaN	81	50	0
the talbots inc	talk inc	Talbots Incorporated	Talk Incorporated	Talbots	Talk	Talbots	NaN	86	55	0
the talbots inc	talon inc	Talbots Incorporated	Talon Incorporated	Talbots	Talon	Talbots	NaN	89	67	0
the talbots inc	talos energy inc	Talbots Incorporated	Talos Energy Incorporated	Talbots	Talos Energy	Talbots	Talos	80	53	83
the talbots inc	tambrands inc	Talbots Incorporated	Tambrands Incorporated	Talbots	Tambrands	Talbots	Tambrands	81	50	50
the talbots inc	tangoe inc	Talbots Incorporated	Tangoe Incorporated	Talbots	Tangoe	Talbots	Tangoe	82	46	46

Figure 6.2

EDGAR log cleaning

Altogether the EDGAR server logs contain more than 2 billion user requests, however, many of these requests are either (1) uninformative or (2) performed by web scrapers and other bots. Because I am most interested in targeted, human-initiated requests, I follow Ryans (2017) which incorporates filters from two prior studies (Drake et al., 2015; Loughran and McDonald, 2017) to screen out irrelevant and bot-generated requests and eliminate noise from my sample.

I classify a request as uninformative if any of the following conditions are met:

1. **File requested is an index page (Index = 1)** These pages are akin to a menu or home screen. Although they can be linked to a specific company, they provide no evidence that investors have accessed any filings and are not indicative of consumption.
2. **File request failed (Code = 300)** This represents a failed download request. This could occur for many reasons, but is indicative of a lack of consumption.

I classify a user as a bot if any of the following conditions are met:

1. **User self-identifies as a web crawler (Crawler = 1)**
2. **User requests ≥ 500 files in a day**
3. **User requests ≥ 25 files in a minute**
4. **User requests files from ≥ 3 distinct firms in a minute**
5. **User requests a file every hour for 24 hours**

I eliminate all uninformative and bot-generated requests from my sample. I further group requests by IP, filing, date such that if the same user requests the same filing multiple times in a day this will only be reflected as one request. After applying these filtering and grouping processes, I am left with a sample of approximately 800 million unique EDGAR requests for the period from 2003 - 2017.

IP Matching and Investor Classification

One of the unique features of the EDGAR server logs is that all download requests are matched to unique IP addresses. As a result, I obtain additional details about *who* consumes accounting information by mapping IP addresses to specific users and organizations. In order to protect users' privacy, each download in the EDGAR server log is associated with an anonymized IP address that takes the form `###.###.###.xxx` where xxx represents a string of characters that masks the full IP address. For instance, I might observe an IP address of 168.11.43.xyz.²⁸

I rely on Chen et al. (2020) to de-anonymize these IP addresses. In their paper, they provide a novel method of de-anonymizing IP addresses by cross-referencing server log activity between that of EDGAR and another widely used related website and examining the frequency of co-search patterns. In so doing, they obtain a mapping that can be used to decipher the EDGAR logs. The authors generously made this cipher table available to me and I use it to de-anonymize the EDGAR server logs.²⁹

After de-ciphering the IP addresses, I use WhoIs data obtained from ARIN to match the IP address to its owner. ARIN's WhoIs data contains a record of all current IP addresses as well as a name of the organization they are registered to and the date of registration.³⁰ Using WhoIs, I am able to match 99.9% of observations. Of these, I am able to match roughly 36.8% to specific organizations and individuals.³¹

Having deciphered the IP addresses, I am also able to classify retail IP addresses. I do this in two ways. First, I exploit a characteristic of the WhoIs data. In their documentation of the WhoIs database, ARIN mentions that all names and addresses for private customers are hidden in order to protect customers' privacy. As such, any IP match which returns a private address or private name I classify as retail. Second, most ISPs and other providers lease only a small range of IP blocks for use in home routers. I classify

²⁸Importantly, these characters have a one to one mapping with numbers over the full history of the EDGAR dataset. In other words, if 'xyz' is masking the numbers '111', this is true across the full sample period from 2003 to 2017.

²⁹I would note that even without the cipher table, most organizations lease large IP blocks, such that deciphering the final octet is not always important in order to identify the end user. For example, many companies rent blocks such as 130.110.110.000 - 130.110.110.255. In this case, no matter what the final octet is, it will belong to this company. Other large companies, such as Apple, often rent very large blocks such as 17.000.000.000 - 17.255.255.255.

³⁰One limitation of using WhoIs data is that organizations may receive new IP addresses over time. I am only able to observe the most recent registrant. This means that if an IP address is registered to XYZ Company on Jan 1, 2015 and I observe that IP address on EDGAR in 2014 and 2016, I can only classify the 2016 observation as belonging to XYZ Company. ARIN does provide a list of historical IP addresses in their WhoWas database, however, access is limited to a handful of requests and is quite cumbersome to obtain. Dyer (2021) is one of the first to obtain and use WhoWas data

³¹IP blocks are often owned by Internet Service Providers (ISPs). These ISPs lease subsections of their IP blocks to other providers, who may, in turn, lease to individual firms. In other words, it is not uncommon for ARIN to own all IP addresses from 1.000.000.000 to 1.255.255.255 and lease blocks 1.000.000.000 to 1.111.111.111 to AT&T and for AT&T to lease blocks 1.000.000.000 to 1.000.000.255 to Ford Motor Company. When indicating that I am able to match 38.6% of observations to organizations and individuals I am indicating that I can identify the end user as a specific organization (i.e. Ford Motor Company) rather than an ISP. I am able to match roughly 99.9% of observations when considering ISPs and intermediaries.

any IP addresses within the ranges 192.168.000.000 - 192.168.255.255, 10.000.000.000 - 10.255.255.255, and 172.16.000.000 - 172.31.255.255 as retail investors.³²

To classify IP matches as institutions, I match the names reported in ARIN's WhoIs data to a list of names obtained from CRSP and 13F filings. Unfortunately, there are often minor differences in how an organization reports its name to ARIN, relative to how it is recorded in CRSP and/or in 13F Filings. To overcome this, I implement a fuzzy matching program using the fuzzywuzzy package in Python and compute matching scores based on Levenshtein distance. For each observation, I compute a score based on the full name, a clean name which strips out articles such as 'and' and 'the', and a non-dictionary name which strips out additional words such as 'incorporated' or 'group'. I then compute a composite score that computes a weighted average of these scores and keep the match with the highest composite score. Following prior literature, I accept matches with a score greater than 80 (100 is the maximum). Upon manual inspection these matches appear to be mostly successful.³³ This matching process is able to match approximately 22% of ARIN organizations to CRSP firms and 13F filers.³⁴ Finally, I classify institutional investors in two ways. First, I classify any organization with SIC code in the financial sector (6000-6999) as an institutional investor. Second, I classify any organization that files schedule 13F as an institutional investor.

³²Although the users of these IP addresses may turnover at a high frequency, it is highly unlikely that these IP addresses would be assigned to anything other than a home router. Since I am not concerned with identifying the behavior of individual retail investors, these changes are not relevant for my classification.

³³In addition, failed matches appear to link companies with very similar names. This suggests that even if a match is incorrect, it will likely match with a business in a similar industry. Again, because I am not tracking specific institutions, this should be sufficient. However, for robustness, I run my tests with a score threshold of 95 and find that my results are qualitatively similar.

³⁴This does not mean that the matching program has a 22% success rate. Rather, many IP addresses pertain to organizations that represent small non-public companies such as law firms that are not listed in CRSP or in 13F filings.

7 Appendix B: Supplementary Analyses

Additional Tests

Table B1

Mispricing and Screening on Price-Fundamental Ratios. This table presents the results of estimating the regression $\Delta \text{LnAtt}_{i,q+1} = \beta_1 \Delta \text{FireSale}_{i,q} + \beta_2 \Delta \text{ScreenRatio}_{i,q} \times \text{FireSale}_{i,q} + \sum_j \beta_j \Delta \text{Controls}_{i,q+1} + \gamma_t + \varepsilon_{i,q+1}$ where $\Delta \text{LnAtt}_{i,q+1}$ represents the change in information consumption in the quarter following a fire sale and γ_t represents year-quarter fixed effects. $\text{ScreenRatio}_{i,q}$ is an indicator for whether a firm enters the extreme quintile of BTM, CFP, and PE, respectively. Controls include returns, volatility, turnover, earnings surprise, analyst following, institutional ownership, size, BTM, FSCORE, leverage, ROA, and indicators for loss firms and fiscal quarter. Columns (1) through (3) present the results for BTM, CFP, and PE separately and Column (4) presents the results when all are included. Refer to [Appendix A](#) for a complete description of all control variables. All continuous variables are winsorized at the 1st and 99th percentiles to limit the effect of outliers and small denominators. I estimate t-statistics using two-way cluster robust standard errors, clustered by firm and year with *, **, and *** indicating statistical significance at less than 10%, 5%, and 1% levels, respectively.

	(1) $\Delta \text{LnAtt}_{q+1}$	(2) $\Delta \text{LnAtt}_{q+1}$	(3) $\Delta \text{LnAtt}_{q+1}$	(4) $\Delta \text{LnAtt}_{q+1}$
$\Delta \text{FireSale}_q$	0.0125*** (5.02)	0.0123*** (4.65)	0.0125*** (5.12)	0.0110*** (4.12)
$\Delta \text{ScreenBTM}_q \times \text{FireSale}_q$	0.0166* (2.00)			0.0165* (2.00)
$\Delta \text{ScreenCFP}_q \times \text{FireSale}_q$		0.0262* (2.05)		0.0260* (2.05)
$\Delta \text{ScreenPE}_q \times \text{FireSale}_q$			0.00835 (1.57)	0.00843 (1.61)
R^2	0.770	0.770	0.770	0.770
N	163,793	163,793	163,793	163,793
Controls	Y	Y	Y	Y
Fixed Effects	Year-Qtr	Year-Qtr	Year-Qtr	Year-Qtr
Cluster by	Firm, Year	Firm, Year	Firm, Year	Firm, Year

Table B2

Mispicing and the Consumption of Accounting Information (10-K and 10-Q Only). This table presents the results of estimating the regression $\Delta \text{LnAtt}_{iq+1} = \beta_1 \Delta \text{FireSale}_{iq} + \sum_j \beta_j \Delta \text{Controls}_{iq+1} + \gamma_t + \varepsilon_{iq+1}$ where $\Delta \text{LnAtt}_{iq+1}$ is measured using total, retail, and institutional downloads and γ_t represents year-quarter fixed effects. In these tests I restrict the sample to only include information consumption related to forms 10-K and 10-Q. Controls include returns, volatility, turnover, earnings surprise, analyst following, institutional ownership, size, BTM, FSCORE, leverage, ROA, and indicators for loss firms and fiscal quarter. Refer to [Appendix A](#) for a complete description of all control variables. All continuous variables are winsorized at the 1st and 99th percentiles to limit the effect of outliers and small denominators. I estimate t-statistics using two-way cluster robust standard errors, clustered by firm and year with *, **, and *** indicating statistical significance at less than 10%, 5%, and 1% levels, respectively.

	(1) $\Delta \text{LnAtt}_{q+1}$	(2) $\Delta \text{LnRetailAtt}_{q+1}$	(3) $\Delta \text{LnInstAtt}_{q+1}$
$\Delta \text{FireSale}_q$	0.0119*** (3.76)	0.00561 (0.83)	0.0173** (2.56)
R^2	0.757	0.0940	0.0846
N	149,709	149,709	149,709
Controls	Y	Y	Y
Fixed Effects	Year-Qtr	Year-Qtr	Year-Qtr
Cluster by	Firm, Year	Firm, Year	Firm, Year

Table B3

Mispricing and Information Consumption - Filings. This table presents the results of estimating the regression $\Delta LnAtt_{iq+1} = \beta_1 \Delta FireSale_{iq} + \sum_j \beta_j \Delta Controls_{iq+1} + \gamma_t + \varepsilon_{iq+1}$ where $\Delta LnAtt_{iq+1}$ represents the change in information consumption in the quarter following a fire sale and γ_t represents year-quarter fixed effects. I separate EDGAR activity by forms and measure information consumption across the 5 most downloaded forms (10K, 10Q, 8K, Form 4, and Proxy Statements). Controls include returns, volatility, turnover, earnings surprise, analyst following, institutional ownership, size, BTM, FSCORE, leverage, ROA, and indicators for loss firms and fiscal quarter. Refer to [Appendix A](#) for a complete description of all control variables. All continuous variables are winsorized at the 1st and 99th percentiles to limit the effect of outliers and small denominators. I estimate t-statistics using two-way cluster robust standard errors, clustered by firm and year with *, **, and *** indicating statistical significance at less than 10%, 5%, and 1% levels, respectively.

	(1) $\Delta LnAtt_{q+1}$	(2) $\Delta LnAtt_{q+1}$	(3) $\Delta LnAtt_{q+1}$	(4) $\Delta LnAtt_{q+1}$	(5) $\Delta LnAtt_{q+1}$
$\Delta FireSale_q \times TenK_q$	0.00854 (1.41)				
$\Delta FireSale_q \times TenQ_q$		0.0135*** (5.05)			
$\Delta FireSale_q \times EightK_q$			0.00912* (2.07)		
$\Delta FireSale_q \times Four_q$				0.0105 (1.14)	
$\Delta FireSale_q \times Proxy_q$					0.00431 (0.66)
R^2	0.655	0.603	0.428	0.234	0.412
N	147,797	147,520	147,368	143,264	143,210
Controls	Y	Y	Y	Y	Y
Fixed Effects	Year-Qtr	Year-Qtr	Year-Qtr	Year-Qtr	Year-Qtr
Cluster by	Firm, Year	Firm, Year	Firm, Year	Firm, Year	Firm, Year

Table B4

Mispricing and Information Consumption - Recent vs. Historical. This table presents the results of estimating the regression $\Delta LnAtt_{iq+1} = \beta_1 \Delta FireSale_{iq} + \sum_j \beta_j \Delta Controls_{iq+1} + \gamma_t + \varepsilon_{iq+1}$ where $\Delta LnAtt_{iq+1}$ represents the change in information consumption in the quarter following a fire sale and γ_t represents year-quarter fixed effects. I separate EDGAR activity based upon how recently the forms have been filed where $LnRecent_{q+1}$ and $LnHistorical_{q+1}$ measure the consumption of forms <90 days old and >365 days old, respectively. Controls include returns, volatility, turnover, earnings surprise, analyst following, institutional ownership, size, BTM, FSCORE, leverage, ROA, and indicators for loss firms and fiscal quarter. Refer to [Appendix A](#) for a complete description of all control variables. All continuous variables are winsorized at the 1st and 99th percentiles to limit the effect of outliers and small denominators. I estimate t-statistics using two-way cluster robust standard errors, clustered by firm and year with *, **, and *** indicating statistical significance at less than 10%, 5%, and 1% levels, respectively.

	(1) $\Delta LnRecent_{q+1}$	(2) $\Delta LnHistorical_{q+1}$
$\Delta FireSale_q$	0.0252*** (7.65)	0.0110** (2.22)
R^2	0.541	0.638
N	167,645	167,645
Controls	Y	Y
Fixed Effects	Year-Qtr	Year-Qtr

Tests using Wardlaw (2020) Flow-to-Volume measure

If you recall from equation (3), the *MFFlow* measure is defined as:

$$MFFlow_{i,t} = \left(\sum_{j=1}^m \frac{F_{j,t} s_{i,j,t-1}}{VOL_{i,t}} \middle| Outflow_{j,t} \geq 5\% \right)$$

If we expand this by plugging in s_{ijt-1} and breaking $VOL_{i,t}$ into its component terms, we get the following:

$$MFFlow_{i,t} = \left(\sum_{j=1}^m \frac{F_{j,t} \times SHARES_{i,j,t-1} \times PRC_{i,t-1}}{TA_{j,t-1} \times PRC_{i,t} \times SHARE_VOL_{i,t}} \right)$$

where I have dropped the conditional component for illustrative purposes. Wardlaw points out that these terms can be grouped into 3 distinct terms.

$$MFFlow_{i,t} = \sum_{j=1}^m \left(\frac{F_{j,t} \times SHARES_{i,j,t-1}}{TA_{j,t-1}} \right) \left(\frac{1}{SHARE_VOL_{i,t}} \right) \left(\frac{PRC_{i,t-1}}{PRC_{i,t}} \right)$$

where the third term represents the inverse of the gross return since

$$\frac{PRC_{t-1}}{PRC_t} = 1 / \left(\frac{PRC_t}{PRC_{t-1}} \right) = \frac{1}{1 + r_t}$$

Wardlaw notes that this term creates a mechanical issue as the fund flow portion of the *MFFlow* measure is multiplied by the inverse of the gross return, such that large negative returns lead to large values of *MFFlow*. To alleviate this potential problem, he proposes an alternative measure, which he calls flow-to-volume, by changing PRC_t in the denominator to PRC_{t-1} , which eliminates this mechanical component.

$$FlowToVolume_{i,t} = \sum_{j=1}^m \left(\frac{F_{j,t} \times SHARES_{i,j,t-1}}{TA_{j,t-1}} \right) \left(\frac{1}{SHARE_VOL_{i,t}} \right)$$

I re-examine my findings using *FlowToVolume* in place of *MFlow*. As before, I classify fire sales as firm-quarter observations in the top decile of *FlowToVolume*. The results of these tests are qualitatively similar to my main results and can be found on the following pages in Tables B1-B6.

Table B5

Mispricing and the Consumption of Accounting Information. This table presents the results of estimating the regression $\Delta LnAtt_{iq+1} = \beta_1 \Delta Mispricing_{iq} + \sum_k \beta_k Controls_{iq+1} + \sum_j \beta_j \Delta Controls_{iq+1} + \gamma_t + \varepsilon_{iq+1}$ where $\Delta LnAtt_{iq+1}$ represents the change in information consumption in the quarter following a fire sale and γ_t represents year-quarter fixed effects. Controls include returns, volatility, turnover, earnings surprise, analyst following, institutional ownership, size, BTM, FSCORE, leverage, ROA, and indicators for loss firms and fiscal quarter. I proxy for mispricing using mutual fund flows and fire sales, the results of which are displayed in columns (1) and (2) respectively. Refer to [Appendix A](#) for a complete description of all control variables. All continuous variables are winsorized at the 1st and 99th percentiles to limit the effect of outliers and small denominators. I estimate t-statistics using two-way cluster robust standard errors, clustered by firm and year with *, **, and *** indicating statistical significance at less than 10%, 5%, and 1% levels, respectively.

	(1) $\Delta LnAtt_{q+1}$	(2) $\Delta LnAtt_{q+1}$	(3) $\Delta LnAtt_{q+1}$	(4) $\Delta LnAtt_{q+1}$
$\Delta MFFlow_q$	0.00513*** (3.37)			
$\Delta FireSale_q$		0.0140*** (5.81)		
$\Delta FlowToVolume_q$			0.00460*** (2.99)	
$\Delta FireSaleFTV_q$				0.00790** (2.53)
R^2	0.769	0.769	0.769	0.769
N	163,210	163,210	163,210	163,210
Controls	Y	Y	Y	Y
Fixed Effects	Year-Qtr	Year-Qtr	Year-Qtr	Year-Qtr
Cluster by	Firm, Year	Firm, Year	Firm, Year	Firm, Year

Table B6

Attributes of Information Consumption Following Mispricing. This table studies the effects of mispricing across different investor types. Specifically, it displays the results of estimating the regression $\Delta LnInvAtt_{iq+1} = \beta_1 \Delta FireSale_{iq} + \sum_j \beta_j \Delta Controls_{iq+1} + \gamma_t + \varepsilon_{iq+1}$ where $\Delta LnInvAtt_{iq+1}$ represents the change in information consumption for retail investors $LnRetailAtt_{q+1}$ and institutional investors $LnInstAtt_{q+1}$ in the quarter following a fire sale and γ_t represents year-quarter fixed effects. Controls include returns, volatility, turnover, earnings surprise, analyst following, institutional ownership, size, BTM, FSCORE, leverage, ROA, and indicators for loss firms and fiscal quarter. Refer to [Appendix A](#) for a complete description of all control variables. All continuous variables are winsorized at the 1st and 99th percentiles to limit the effect of outliers and small denominators. I estimate t-statistics using two-way cluster robust standard errors, clustered by firm and year with *, **, and *** indicating statistical significance at less than 10%, 5%, and 1% levels, respectively.

	(1) $\Delta LnRetailAtt_{q+1}$	(2) $\Delta LnInstAtt_{q+1}$
$\Delta FireSaleFTV_q$	-0.0003 (-0.03)	0.0177* (2.12)
R^2	0.0623	0.104
N	163,210	163,210
Controls	Y	Y
Fixed Effects	Year-Qtr	Year-Qtr
Cluster by	Firm, Year	Firm, Year

Table B7

Fire Sales and Institutional Consumption - Investor Type. This table presents the results of estimating the regression $\Delta \text{LnAtt}_{iq+1} = \beta_1 \Delta \text{FireSaleFTV}_{iq} + \sum_j \beta_j \Delta \text{Controls}_{iq+1} + \gamma_t + \varepsilon_{iq+1}$ where $\Delta \text{LnAtt}_{iq+1}$ is measured using consumption from institutional investors and γ_t represents year-quarter fixed effects. I classify new investors as users who search for a firm in the current, but not in the previous quarter and existing investors as users who search for a firm in both the current and previous quarter. I use 13F filings to classify institutions as small and large depending upon the number of unique holdings (bottom tercile and top tercile, respectively). Controls include returns, volatility, turnover, earnings surprise, analyst following, institutional ownership, size, BTM, FSCORE, leverage, ROA, and indicators for loss firms and fiscal quarter. Refer to [Appendix A](#) for a complete description of all control variables. All continuous variables are winsorized at the 1st and 99th percentiles to limit the effect of outliers and small denominators. I estimate t-statistics using two-way cluster robust standard errors, clustered by firm and year with *, **, and *** indicating statistical significance at less than 10%, 5%, and 1% levels, respectively.

Panel A: Investors' Prior Search History				
	New Investors		Existing Investors	
	$\Delta \text{LnInst}_{q+1}$		$\Delta \text{LnInst}_{q+1}$	
$\Delta \text{FireSaleFTV}_q$	0.0208**		-0.00297	
	(2.58)		(-0.51)	
R^2	0.0937		0.0455	
N	167,835		167,835	
Controls	Y		Y	
Fixed Effects	Year-Qtr		Year-Qtr	
Cluster by	Firm, Year		Firm, Year	
Panel B: Investor Size				
	Small		Large	
	$\Delta \text{LnInst}_{q+1}$		$\Delta \text{LnInst}_{q+1}$	
$\Delta \text{FireSaleFTV}_q$	0.0170		0.00828	
	(1.45)		(0.84)	
	(-0.20)		(-0.52)	
R^2	0.390		0.0428	
N	167,835		167,835	
Controls	Y		Y	
Fixed Effects	Year-Qtr		Year-Qtr	
Cluster by	Firm, Year		Firm, Year	
Panel C: Investors' Search History & Size				
	New \times Small	New \times Large	Existing \times Small	Existing \times Large
	$\Delta \text{LnInst}_{q+1}$	$\Delta \text{LnInst}_{q+1}$	$\Delta \text{LnInst}_{q+1}$	$\Delta \text{LnInst}_{q+1}$
$\Delta \text{FireSaleFTV}_q$	0.0182	0.0197*	0.00314	-0.00414
	(1.42)	(1.90)	(0.47)	(-0.86)
R^2	0.356	0.0341	0.109	0.0327
N	167,835	167,835	167,835	167,835
Controls	Y	Y	Y	Y
Fixed Effects	Year-Qtr	Year-Qtr	Year-Qtr	Year-Qtr
Cluster by	Firm, Year	Firm, Year	Firm, Year	Firm, Year

Table B8

Fire Sales and Investor Holdings - Evidence from 13F Filings. This table presents the results of estimating the regression $\Delta \text{LnInvest}_{iq+1} = \beta_1 \Delta \text{FireSaleFTV}_{iq} + \sum_j \beta_j \Delta \text{Controls}_{iq+1} + \gamma_t + \varepsilon_{iq+1}$ where $\Delta \text{LnInvest}_{iq+1}$ represents the change in the number of investors in the quarter following a fire sale and γ_t represents year-quarter fixed effects. I measure LnNewInvest_{iq} as the number of investors who report holdings for a firm in the current, but not the previous, period (extensive margin) as well as the dollar volume of purchases on the extensive margin in columns (1) and (2), respectively. Controls include returns, volatility, turnover, earnings surprise, analyst following, institutional ownership, size, BTM, FSCORE, leverage, ROA, and indicators for loss firms and fiscal quarter. Refer to [Appendix A](#) for a complete description of all control variables. All continuous variables are winsorized at the 1st and 99th percentiles to limit the effect of outliers and small denominators. I estimate t-statistics using two-way cluster robust standard errors, clustered by firm and year with *, **, and *** indicating statistical significance at less than 10%, 5%, and 1% levels, respectively.

	(1) $\Delta \text{LnNewInvest}_{q+1}$	(2) $\Delta \text{LnNewVolume}_{q+1}$
$\Delta \text{FireSaleFTV}_q$	0.0342*** (4.92)	0.114* (1.97)
R^2	0.0907	0.0388
N	159,664	159,664
Controls	Y	Y
Fixed Effects	Year-Qtr	Year-Qtr
Cluster by	Firm, Year	Firm, Year

Table B9

Fire Sales, Information Consumption, and the Magnitude of Return Reversals. This table presents the results of estimating the regression $Ret_{i\tau} = \beta_1 FireSaleFTV_{iq} + \beta_2 PostAtt_{iq+1} + \beta_3 PostAtt_{iq+1} \times FireSaleFTV_{iq} + \sum_k \beta_k Controls_{iq+1} + \gamma_t + \delta_i + \varepsilon_{iq+1}$ where $PostAtt_{iq+1}$ represents information consumption in the quarter following a fire sale and γ_t and δ_i represent year-quarter and firm fixed effects, respectively. $Ret_{i\tau}$ represents buy-and-hold market-adjusted returns for the period $[t, t + \tau]$ such that columns (1) through (4) measure returns over the 3, 6, 9, and 12 month periods following a fire sale, respectively. Controls include lagged attention, prior period returns, volatility, analyst following, institutional ownership, size, BTM, FSCORE, leverage, ROA, and indicators for loss firms and fiscal quarter. Refer to [Appendix A](#) for a complete description of all control variables. All continuous variables are winsorized at the 1st and 99th percentiles to limit the effect of outliers and small denominators. The sample is limited to fire sale firms. I estimate t-statistics using two-way cluster robust standard errors, clustered by firm and year with *, **, and *** indicating statistical significance at less than 10%, 5%, and 1% levels, respectively.

Panel A: All EDGAR Activity								
	(1) <i>Ret</i> ₃	(2) <i>Ret</i> ₆	(3) <i>Ret</i> ₉	(4) <i>Ret</i> ₁₂				
<i>FireSaleFTV</i> _q	-0.0020 (-0.72)	0.0034 (0.70)	0.0125* (2.08)	0.0201** (2.75)				
<i>PostAtt</i> _{q+1}	0.0317*** (4.30)	0.0121 (1.22)	-0.0005 (-0.04)	-0.0162 (-1.14)				
<i>PostAtt</i> _{q+1} × <i>FireSaleFTV</i> _q	-0.0006 (-0.19)	0.0110 (1.41)	0.0199* (1.94)	0.0268* (1.89)				
<i>R</i> ²	0.129	0.157	0.198	0.212				
N	167,699	167,278	166,098	164,291				
Controls	Y	Y	Y	Y				
Fixed Effects	Firm, Year-Qtr	Firm, Year-Qtr	Firm, Year-Qtr	Firm, Year-Qtr				
Cluster by	Firm, Year	Firm, Year	Firm, Year	Firm, Year				
Panel B: Retail and Institutional EDGAR Activity								
	Retail				Institutional			
	(1) <i>Ret</i> ₃	(2) <i>Ret</i> ₆	(3) <i>Ret</i> ₉	(4) <i>Ret</i> ₁₂	(5) <i>Ret</i> ₃	(6) <i>Ret</i> ₆	(7) <i>Ret</i> ₉	(8) <i>Ret</i> ₁₂
<i>FireSaleFTV</i>	-0.0020 (-0.74)	0.0032 (0.72)	0.0117* (1.95)	0.0184** (2.52)	-0.0012 (-0.47)	0.0036 (0.76)	0.0125* (1.99)	0.0197** (2.68)
<i>PostAtt</i> _{q+1}	0.0079* (1.96)	0.0035 (0.59)	-0.0042 (-0.65)	-0.0058 (-0.70)	0.0139** (2.80)	0.0036 (0.68)	-0.0053 (-0.79)	-0.0187* (-2.13)
<i>PostAtt</i> _{q+1} × <i>FireSaleFTV</i> _q	0.0004 (0.06)	0.0139 (0.96)	0.0202 (1.34)	0.0212 (1.04)	0.0077* (1.77)	0.0190 (1.62)	0.0304* (1.85)	0.0382* (1.80)
<i>R</i> ²	0.128	0.157	0.198	0.211	0.128	0.157	0.198	0.212
N	167,699	167,278	166,098	164,291	167,699	167,278	166,098	164,291
Controls	Y	Y	Y	Y	Y	Y	Y	Y
Fixed Effects	Firm, Year- Qtr	Firm, Year- Qtr	Firm, Year- Qtr	Firm, Year- Qtr	Firm, Year- Qtr	Firm, Year- Qtr	Firm, Year- Qtr	Firm, Year- Qtr
Cluster by	Firm, Year	Firm, Year	Firm, Year	Firm, Year	Firm, Year	Firm, Year	Firm, Year	Firm, Year

Table B10

Fire Sales, Information Consumption, and the Likelihood of Recovery. This table presents the results of estimating the hazard model $Recover_{i\tau} = \beta_1 PostAtt_{iq+1} + \sum_k \beta_k Controls_{iq+1} + \varepsilon_{iq+1}$ where $PostAtt_{iq+1}$ represents information consumption in the quarter following a fire sale and τ is the time horizon. To estimate the hazard model, I count the number of days until a firm's buy-and-hold market-adjusted return, including the loss from the fire sale, is greater than 0 for at least 5 consecutive days. $Recover_{i\tau}$ is an indicator variable equal to one when the previous condition is satisfied. Controls include lagged attention, prior period returns, volatility, analyst following, institutional ownership, size, BTM, FSCORE, leverage, ROA, and indicators for loss firms and fiscal quarter. I estimate the hazard model over various time horizons. Columns (1) through (3), (4) through (6), and (7) through (9) show the results over a 90-day, 180-day, and 365-day horizon for all, retail, and institutional activity, respectively. Refer to [Appendix A](#) for a complete description of all control variables. All continuous variables are winsorized at the 1st and 99th percentiles to limit the effect of outliers and small denominators. The sample is limited to fire sale firms. Coefficients are reported as likelihood ratios. I estimate t-statistics using robust standard errors with *, **, and *** indicating statistical significance at less than 10%, 5%, and 1% levels, respectively.

	DV: 90 Day Recovery			DV: 180 Day Recovery			DV: 365 Day Recovery		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Total	Retail	Inst	Total	Retail	Inst	Total	Retail	Inst
$PostAtt_{q+1}$	0.103 (1.41)	-0.063* (-1.84)	0.061 (1.27)	0.142** (2.43)	-0.011 (-0.40)	0.058 (1.52)	0.062 (1.23)	0.003 (0.14)	0.041 (1.27)
N	8,914	8,914	8,914	8,914	8,914	8,914	8,914	8,914	8,914
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y
Cluster by	Robust	Robust	Robust	Robust	Robust	Robust	Robust	Robust	Robust

Table B11

Fire Sales, Information Consumption, and the Speed of Return Reversals. This table presents the results of estimating the regression $AdjIPT_{i[m,m+12]} = \beta_1 FireSaleFTV_{iq} + \beta_2 PostAtt_{iq+1} + \beta_3 PostAtt_{iq+1} \times FireSaleFTV_{iq} + \sum_k \beta_k Controls_{iq+1} + \gamma_t + \delta_i + \varepsilon_{iq+1}$ where $PostAtt_{iq+1}$ represents information consumption in the quarter following a fire sale and γ_t and δ_i represent year-quarter and firm fixed effects, respectively. IPT is measured as $IPT = \sum_{m=1}^{11} (BH_{1,m}/BH_{1,12}) + 0.5$ where BH is the firm's market-adjusted buy-and-hold return. I follow (Blankespoor et al., 2018) in constructing $AdjIPT$. Refer to Section 4.2 for more details. Controls include lagged attention, prior period returns, volatility, analyst following, institutional ownership, size, BTM, FSCORE, leverage, ROA, and indicators for loss firms and fiscal quarter. Columns (1) through (3) measure information consumption using all EDGAR activity, retail investor activity, and institutional investor activity, respectively. Refer to Appendix A for a complete description of all control variables. All continuous variables are winsorized at the 1st and 99th percentiles to limit the effect of outliers and small denominators. The sample is limited to fire sale firms. I estimate t-statistics using two-way cluster robust standard errors, clustered by firm and year with *, **, and *** indicating statistical significance at less than 10%, 5%, and 1% levels, respectively.

	(1) Total	(2) Retail	(3) Institutional
	$AdjIPT_{[m,m+12]}$	$AdjIPT_{[m,m+12]}$	$AdjIPT_{[m,m+12]}$
<i>FireSaleFTV</i>	0.0120 (0.32)	-0.00492 (-0.11)	0.00882 (0.26)
<i>PostAtt_{q+1}</i>	-0.00205 (-0.32)	0.00376 (0.95)	-0.00373 (-0.75)
<i>PostAtt_{q+1} × FireSaleFTV_q</i>	0.0204** (2.17)	0.00347 (0.42)	0.0156 (1.53)
R^2	0.0498	0.0498	0.0498
N	141,716	141,716	141,716
Controls	Y	Y	Y
Fixed Effects	Firm, Year-Qtr	Firm, Year-Qtr	Firm, Year-Qtr
Cluster by	Firm, Year	Firm, Year	Firm, Year

Table B12

Fire Sales, Information Consumption, and Earnings Events. This table presents the results of estimating the regression $BHMAR_{iq+1} = \beta_1 SUE_{iq+1} + \beta_2 PreAtt_{iq+1} + \beta_3 PreAtt_{iq+1} \times SUE_{iq+1} + \sum_k \beta_k Controls_{iq+1} + \gamma_t + \sigma_{ind} + \varepsilon_{iq+1}$ where $PreAtt_{iq+1}$ represents information consumption in the 10 day period $[d-12, d-2]$ prior to an earnings event and γ_t and σ_{ind} represent year-quarter and industry fixed effects, respectively. Controls include lagged attention, prior period returns, volatility, analyst following, institutional ownership, size, BTM, FSCORE, leverage, ROA, earnings volatility, earnings persistence, and indicators for loss firms and fourth quarter reports. Column (1) measures information consumption using all EDGAR activity and columns (2) and (3) measure information consumption using institutional investors' and retail investors' EDGAR activity, respectively. Refer to [Appendix A](#) for a complete description of all control variables. All continuous variables are winsorized at the 1st and 99th percentiles to limit the effect of outliers and small denominators. The sample is limited to fire sale firms. I estimate t-statistics using two-way cluster robust standard errors, clustered by firm and year with *, **, and *** indicating statistical significance at less than 10%, 5%, and 1% levels, respectively.

Panel A: Prices Lead Earnings			
	(1) Total	(2) Retail	(3) Inst
	$BHMAR_{[d-12,d-2]}$	$BHMAR_{[d-12,d-2]}$	$BHMAR_{[d-12,d-2]}$
SUE_{q+1}	0.0379 (1.71)	0.0518* (2.08)	0.0538* (2.13)
$PreAtt_{q+1}$	0.000710 (0.70)	0.00156** (2.48)	0.000837 (0.82)
$PreAtt_{iq+1} \times SUE_{q+1}$	0.0275 (0.98)	0.0218 (0.78)	0.0429 (1.39)
R^2	0.213	0.213	0.213
N	16,113	16,113	16,113
Controls	Y	Y	Y
Fixed Effects	Firm, Year-Qtr	Firm, Year-Qtr	Firm, Year-Qtr
Cluster by	Firm, Year	Firm, Year	Firm, Year
Panel B: Earnings Response Coefficients			
	(1) Total	(2) Retail	(3) Inst
	$BHMAR_{[d-1,d+1]}$	$BHMAR_{[d-1,d+1]}$	$BHMAR_{[d-1,d+1]}$
SUE_{q+1}	0.154*** (7.07)	0.146*** (5.68)	0.151*** (6.11)
$PreAtt_{q+1}$	0.000131 (0.06)	-0.000364 (-0.47)	0.000822 (0.57)
$PreAtt_{iq+1} \times SUE_{q+1}$	0.00353 (0.12)	-0.0141 (-0.67)	-0.00901 (-0.31)
R^2	0.198	0.198	0.198
N	16,200	16,200	16,200
Controls	Y	Y	Y
Fixed Effects	Firm, Year-Qtr	Firm, Year-Qtr	Firm, Year-Qtr
Cluster by	Firm, Year	Firm, Year	Firm, Year

Table B13

Mispricing and Price-Fundamental Ratios. This table presents the results of estimating the regression $\Delta \text{LnAtt}_{iq+1} = \beta_1 \Delta \text{FireSale}_{iq} + \beta_2 \Delta \text{ScreenRatio}_{iq} \times \text{FireSale}_{iq} + \sum_k \beta_k \text{Controls}_{iq+1} + \sum_j \beta_j \Delta \text{Controls}_{iq+1} + \gamma_t + \varepsilon_{iq+1}$ where $\Delta \text{LnAtt}_{iq+1}$ represents the change in information consumption in the quarter following a fire sale and γ_t represents year-quarter fixed effects. ScreenRatio_{iq} is an indicator for whether a firm enters the extreme quintile of BTM, CFP, and PE, respectively. Controls include returns, volatility, turnover, earnings surprise, analyst following, institutional ownership, size, BTM, FSCORE, leverage, ROA, and indicators for loss firms and fiscal quarter. Columns (1) through (3) present the results for BTM, CFP, and PE separately and Column (4) presents the results when all are included. Refer to [Appendix A](#) for a complete description of all control variables. All continuous variables are winsorized at the 1st and 99th percentiles to limit the effect of outliers and small denominators. I estimate t-statistics using two-way cluster robust standard errors, clustered by firm and year with *, **, and *** indicating statistical significance at less than 10%, 5%, and 1% levels, respectively.

	(1) $\Delta \text{LnAtt}_{q+1}$	(2) $\Delta \text{LnAtt}_{q+1}$	(3) $\Delta \text{LnAtt}_{q+1}$	(4) $\Delta \text{LnAtt}_{q+1}$
$\Delta \text{FireSaleFTV}_q$	0.00718** (2.23)	0.00732** (2.18)	0.00756** (2.43)	0.00612* (1.78)
$\Delta \text{ScreenBTM}_q \times \text{FireSaleFTV}_q$	0.0181** (2.45)			0.0182** (2.46)
$\Delta \text{ScreenCFP}_q \times \text{FireSaleFTV}_q$		0.0221* (1.76)		0.0221* (1.78)
$\Delta \text{ScreenPE}_q \times \text{FireSaleFTV}_q$			0.00586 (1.18)	0.00580 (1.20)
R^2	0.769	0.769	0.769	0.769
N	167,645	167,645	167,645	167,645
Controls	Y	Y	Y	Y
Fixed Effects	Year-Qtr	Year-Qtr	Year-Qtr	Year-Qtr
Cluster by	Firm, Year	Firm, Year	Firm, Year	Firm, Year

Table B14

Mispicing and the Consumption of Accounting Information (10-K and 10-Q Only). This table presents the results of estimating the regression $\Delta LnAtt_{iq+1} = \beta_1 \Delta FireSaleFTV_{iq} + \sum_j \beta_j \Delta Controls_{iq+1} + \gamma_t + \varepsilon_{iq+1}$ where $\Delta LnAtt_{iq+1}$ is measured using total, retail, and institutional downloads and γ_t represents year-quarter fixed effects. In these tests I restrict the sample to only include information consumption related to forms 10-K and 10-Q. Controls include returns, volatility, turnover, earnings surprise, analyst following, institutional ownership, size, BTM, FSCORE, leverage, ROA, and indicators for loss firms and fiscal quarter. Refer to [Appendix A](#) for a complete description of all control variables. All continuous variables are winsorized at the 1st and 99th percentiles to limit the effect of outliers and small denominators. I estimate t-statistics using two-way cluster robust standard errors, clustered by firm and year with *, **, and *** indicating statistical significance at less than 10%, 5%, and 1% levels, respectively.

	(1) $\Delta LnAtt_{q+1}$	(2) $\Delta LnRetailAtt_{q+1}$	(3) $\Delta LnInstAtt_{q+1}$
$\Delta FireSale_FTV_{q+1}$	0.0095** (2.46)	-0.0012 (-0.17)	0.0142** (2.49)
R^2	0.757	0.0940	0.0846
N	149,709	149,709	149,709
Controls	Y	Y	Y
Fixed Effects	Year-Qtr	Year-Qtr	Year-Qtr
Cluster by	Firm, Year	Firm, Year	Firm, Year

Table B15

Mispricing and Information Consumption - Filings. This table presents the results of estimating the regression $\Delta LnAtt_{iq+1} = \beta_1 \Delta FireSaleFTV_{iq} + \sum_j \beta_j \Delta Controls_{iq+1} + \gamma_t + \varepsilon_{iq+1}$ where $\Delta LnAtt_{iq+1}$ represents the change in information consumption in the quarter following a fire sale and γ_t represents year-quarter fixed effects. I separate EDGAR activity by forms and measure information consumption across the 5 most downloaded forms (10K, 10Q, 8K, Form 4, and Proxy Statements). Controls include returns, volatility, turnover, earnings surprise, analyst following, institutional ownership, size, BTM, FSCORE, leverage, ROA, and indicators for loss firms and fiscal quarter. Refer to [Appendix A](#) for a complete description of all control variables. All continuous variables are winsorized at the 1st and 99th percentiles to limit the effect of outliers and small denominators. I estimate t-statistics using two-way cluster robust standard errors, clustered by firm and year with *, **, and *** indicating statistical significance at less than 10%, 5%, and 1% levels, respectively.

	(1) $\Delta LnAtt_{q+1}$	(2) $\Delta LnAtt_{q+1}$	(3) $\Delta LnAtt_{q+1}$	(4) $\Delta LnAtt_{q+1}$	(5) $\Delta LnAtt_{q+1}$
$\Delta FireSaleFTV_q \times TenK_q$	0.00854 (1.41)				
$\Delta FireSaleFTV_q \times TenQ_q$		0.0135*** (5.05)			
$\Delta FireSaleFTV_q \times EightK_q$			0.00912* (2.07)		
$\Delta FireSaleFTV_q \times Four_q$				0.0105 (1.14)	
$\Delta FireSaleFTV_q \times Proxy_q$					0.00431 (0.66)
R^2	0.655	0.603	0.428	0.234	0.412
N	147,797	147,520	147,368	143,264	143,210
Controls	Y	Y	Y	Y	Y
Fixed Effects	Year-Qtr	Year-Qtr	Year-Qtr	Year-Qtr	Year-Qtr
Cluster by	Firm, Year	Firm, Year	Firm, Year	Firm, Year	Firm, Year

Table B16

Mispricing and Information Consumption - Recent vs. Historical. This table presents the results of estimating the regression $\Delta LnAtt_{iq+1} = \beta_1 \Delta FireSaleFTV_{iq} + \sum_j \beta_j \Delta Controls_{iq+1} + \gamma_t + \varepsilon_{iq+1}$ where $\Delta LnAtt_{iq+1}$ represents the change in information consumption in the quarter following a fire sale and γ_t represents year-quarter fixed effects. I separate EDGAR activity based upon how recently the forms have been filed where $LnRecent_{q+1}$ and $LnHistorical_{q+1}$ measure the consumption of forms <90 days old and >365 days old, respectively. Controls include returns, volatility, turnover, earnings surprise, analyst following, institutional ownership, size, BTM, FSCORE, leverage, ROA, and indicators for loss firms and fiscal quarter. Refer to [Appendix A](#) for a complete description of all control variables. All continuous variables are winsorized at the 1st and 99th percentiles to limit the effect of outliers and small denominators. I estimate t-statistics using two-way cluster robust standard errors, clustered by firm and year with *, **, and *** indicating statistical significance at less than 10%, 5%, and 1% levels, respectively.

	(1) $\Delta LnRecent_{q+1}$	(2) $\Delta LnHistorical_{q+1}$
$\Delta FireSaleFTV_q$	0.0216*** (5.42)	0.00830 (1.59)
R^2	0.541	0.638
N	167,645	167,645
Controls	Y	Y
Fixed Effects	Year-Qtr	Year-Qtr
Cluster by	Firm, Year	Firm, Year