Demand for Information and Stock Returns: Evidence from EDGAR

Pingle Wang*

February 28, 2024

Abstract

This paper empirically shows that information acquisition affects stock returns by reducing firm-level information asymmetry. When firms disclose material information known by insiders, information acquisition reduces asymmetric information and lowers stock returns. The effect is stronger for both unexpectedly good and bad news relative to anticipated news and when investors' cost of information processing is lower. Using the Northeast Blackout of 2003 as a natural experiment, I explore an exogenous shock in information acquisition and show causal evidence that information acquisition reduces information asymmetry.

*Jindal School of Management, University of Texas at Dallas. E-mail: pingle.wang@utdallas.edu. I am extremely grateful to Ron Kaniel for his relentless support and guidance. I am also grateful to my dissertation committee members: Alan Moreira and Jerold Warner for their invaluable support. I would like to thank Shuaiyu Chen, Ramona Dagostino, Antonio Gargano, Michael Gofman, Ryan Israelsen, Xiaoxia Lou, Robert Novy-Marx, Christian Opp, Dmitry Orlov, Mykola Pinchuk, Bill Schwert, Clifford Smith, Giulio Trigilia, Pavel Zryumov, and seminar participants at Chinese University of Hong Kong, Singapore Management University, Southern Methodist University, Texas A&M University, University of Connecticut, University of Delaware, University of Rochester, University of Texas at Dallas, and University of Wisconsin-Madison for their helpful discussions and comments. All errors are my own.

1 Introduction

Information plays a central role in determining asset prices. A large and extensive empirical literature has investigated how information disclosure affects information asymmetry between investors and insiders. As pointed out in Huberman and Regev (2001), the supplied information may not be fully absorbed and sometimes may even be neglected by investors, especially in an era of information explosion. Therefore, it is crucially important to study how information acquisition affects asset prices. The empirical literature regarding the effect of information acquisition on stock returns is growing but still small because of data limitations. Empirical studies have used direct and indirect measures of information acquisition, all of which predict positive abnormal returns.[1] Researchers explain the empirical findings with an attention channel: given the limited attention and the short-selling constraint, investors are more likely to buy rather than sell attention-grabbing stocks. Therefore, information acquisition reflects investors' increased attention and recognition of the firm, which in turn predicts positive abnormal returns in the future.[2]

While the empirical findings generally suggest a positive relation between information acquisition and subsequent stock returns, it is different from the prediction of an information asymmetry channel. Theoretical papers show that acquiring material information reduces payoff uncertainty and asymmetric information between insiders and investors (*i.e., the information asymmetry channel*).[3] Therefore, stocks become less risky to hold, and information acquisition should predict lower rather than higher subsequent returns. To the best of my knowledge, no empirical studies have disentangled the information asymmetry channel from the documented attention channel, and my paper attempts to fill this gap.

The existing proxy for information acquisition is limited to aggregated measures, such

[1]Gervais, Kaniel, and Mingelgrin (2001) and Barber and Odean (2007) indirectly capture investor attention with abnormal trading volume or extreme stock returns. Recently Da, Engelberg, and Gao (2011) and Ben-Rephael, Carlin, Da, and Israelsen (2021) use Google Trends and Bloomberg Search Index to directly measure information acquisition of retail and institutional investors.

[2]Ben-Rephael et al. (2021) study the contemporaneous relation between stock returns and Bloomberg Search Index through a risk channel, and they do not examine future returns.

[3]See Grossman and Stiglitz (1980),Verrecchia (1982),Wang (1993), and Easley and O'Hara (2004).

as Google Trends or Bloomberg Search Index. These measures capture search intensity but fail to distinguish types of information collected by investors. On the one hand, if investors acquire *material* information, it can potentially reduce information asymmetry and lower the future required rate of return. On the other hand, if investors are attracted by attention-grabbing events and collect *stale* information to put events into context, it will lead to a positive return in the future because of the short selling constraint. Given the opposite predictions of returns under each channel, it is critical to distinguish types of information acquisition.

My paper proposes a novel empirical setting through which stale and material information acquisition can be cleanly distinguished, which allows me to test the theoretical predictions of information acquisition under the information asymmetry channel. Based on the filing downloads from the EDGAR log data between 2003 and 2016, I show that acquiring historical filings that contain *stale* information predicts positive abnormal returns that gradually decay over time, consistent with the documented attention channel. However, when investors collect unanticipated *material* information that was privately known by insiders, stock returns are reduced in the future. To directly test the underlying mechanism, I show that acquiring material information reduces the information asymmetry between uninformed investors and firm insiders. The pattern is more pronounced when investors are geographically close to firm headquarters or have prior experience in collecting information, as their cost of information processing tends to be low, consistent with the prediction of Verrecchia (1982). Furthermore, to alleviate the concern about endogenous attention allocation and information asymmetry, I use the Northeast Blackout event as an exogenous shock to investors' information acquisition and provide causal evidence that firms with a large reduction in information acquisition exhibit an increase in firm-level information asymmetry. My finding provides the first empirical evidence that supports the information asymmetry channel and highlights the important distinction between *material* and *stale* information acquisition.

Testing whether information acquisition reduces information asymmetry is challenging, as firm managers may have unobserved incentives to disclose information, which can correlate with both subsequent information acquisition and the information asymmetry. To alleviate the concern from the supply side of information, I focus on the mandatory disclosure, Form 8-K, which is required by the SEC and is filed by firms within four business days after a triggering event. The filings typically contain event-specific material information that was privately known by firm insiders. Compared with other types of information disclosure, such as management forecasts or annual reports, Form 8-K leaves firms with little room to manipulate regarding *when* and *what* to disclose. In my setting, firm insiders are defined as the management teams (issuers) and the selected institutional investors who receive warnings and tipping from issuers. The passing of Regulation FD regulates issuers so that the same information must be shared with the public through filings, but uninformed investors still need to acquire such information by downloading the filing. Accordingly, I use the number of 8-K downloads within a short time window following the filing to proxy for material information acquisition through the information asymmetry channel.

I find that material information acquisition predicts lower stock returns in the future, which is consistent with the information asymmetry channel. Specifically, demand for 8-K filings predicts a significant and negative abnormal return of -18 basis points (bps) for the following month using the Fama-MacBeth regression and controlling for media coverage, the firm's information environment, and a rich set of well-known risk factors that can explain cross-sectional stock returns. Moreover, taking advantage of the timestamp for each filing download, I show that the negative return is solely driven by unscheduled 8-K downloads within the first week following the disclosure, thus supporting the idea of acquiring material and new information. The long/short portfolio sorted on 8-K downloads also yields negative alphas in the subsequent month under various asset pricing models, consistent with the Fama-MacBeth result. More importantly, the long leg (high 8-K downloads) portfolio yields zero alphas, whereas the short leg (low 8-K downloads) portfolio generates substantial and positive alphas. This result lends further support to the information asymmetry channel.

Conditional on a set of firms with material information disclosure, the information risk is resolved for firms with a high number of downloads, as investors timely collect, incorporate, and trade on such information. As a result, these firms can be better priced by the traditional asset pricing models and have zero alphas in the subsequent period. For firms with little downloading activity, information asymmetry remains high despite the disclosure, and such information risk is priced positively in the data (Kelly and Ljungqvist (2012)). Therefore, firms with low 8-K downloads show positive alpha with respect to the traditional asset pricing models in subsequent periods.

One potential explanation for the negative return pattern could be that information content drives both returns and the 8-K downloads. For example, investors may overreact to good news so that future returns are lower. Alternatively, bad news may attract more downloads relative to good news, and firms with bad (good) news disclosure are likely to release bad (good) news in the near future, which can also predict subsequent negative returns. To address these concerns, I calculate three-day cumulative abnormal returns (CAR) around event dates extracted from 8-K filings. Ben-Rephael, Da, Easton, and Israelsen (2022) find that some institutional investors receive leaked information around the event date prior to 8-K filings. Therefore, the CAR measure is a good proxy for sophisticated investors' reaction to the news.[4] The CAR has a mean of 0.4% and a standard deviation of 12%, and its histogram suggests that good news and bad news are equally likely among 8-K filings. Conditional on a set of firms with either a good or bad news release through 8-K filings, high 8-K downloads predict low subsequent returns. When there is little market reaction around the event, 8-K downloads do not predict returns. This finding supports the hypothesis that the information asymmetry channel depends on unanticipated information.

To lend further support to the theoretical predictions under the information asymmetry channel, I explore the heterogeneity in investor types, and examine their differential effects of material information acquisition on stock returns. I find that the effect of material

[4] I do not use textual analysis to determine whether the news is good or bad, as that depends on investors' expectation regarding the event.

information acquisition is stronger for firms with a higher fraction of *local* investors who are geographically close to firm headquarters, and with a higher fraction of *recurring* investors who have prior experience in collecting information. These investors typically incur a lower cost when acquiring and processing information, and the effect of their information acquisition should be stronger (Verrecchia (1982)).

Having demonstrated the effect of material information acquisition on returns, I then directly test its effect on information asymmetry proxied by the price impact measure.[5] In a weekly panel regression setting, I show that 8-K downloads significantly reduce the price impact measure, and the effect is magnified when a new 8-K filing becomes available. A 10% increase in 8-K downloads reduces a firm's price impact measure by 8 bps in the subsequent week, corresponding to a 20% reduction in the price impact measure around the mean.

Since information acquisition and information asymmetry can be endogenously determined, this poses an empirical challenge for causal interpretation based on the OLS analysis. An ideal setup would be to assign investors' filing downloading activities randomly when there are filing disclosures. Utilizing the geographical information of each IP visit, I employ the Northeast Blackout of 2003 as a natural experiment and identify the causal effect of information acquisition on firm-level information asymmetry in a difference-in-differences setting. On August 14, 2003, a widespread unanticipated power outage occurred throughout parts of the northeastern and midwestern United States and the Canadian province of Ontario, beginning just after 4:10 p.m. Although Wall Street got its power back at 6 a.m. on August 15, most traders were unable to commute because of the train system shutdown, and the infrastructure suffered from the power outage. As a result, investors from the affected regions had very limited access to firm filings during the blackout.[6] Even though the geographical distribution of information acquisition before the shock was endogenous, the power outage

[5]The results are robust to other measures, such as the bid-ask spread, PINs, and Amihud illiquidity measure. These measures typically have a large liquidity component, whereas the price impact measure captures the information component of the trade.

[6]"Nobody is here. It's dead. It's a complete ghost town. It really is like coming in on the weekend," said Richard Koss, bond portfolio manager at Brown Brothers Harriman. From CNN news.

created an exogenous variation in downloading activity across firms after the blackout. Ex ante, firms with a larger fraction of historical downloads from the affected regions should incur a larger reduction in downloads during the blackout. Using a difference-in-differences estimation on a sample of firms with material information disclosure one day before the blackout, I show that firms with an additional 1% increase in historical 8-K downloads from the affected regions before the shock suffer a 3% increase in information asymmetry as a result of the blackout.[7] Moreover, I find that such an effect is not driven by either different trading volumes between two types of firms or different sophistication levels of investors due to the blackout.

Information acquisition occurs not only when investors trade on time-sensitive material information, but also when investors are attracted by attention-grabbing events. In the latter case, investors are more likely to buy rather than sell attention-grabbing stocks, given the short-selling constraint faced by certain investors. As a result, information acquisition captures increased attention and predicts positive abnormal returns in the future. I use the number of 10-K downloads as a proxy for *stale* information acquisition under the attention channel for the following reasons: First, Form 10-K provides investors with comprehensive financial and operational statements, which are useful for fundamental investment.[8] Second, unlike 8-K filings, 10-K filings have a significant reporting lag between the fiscal-end date and disclosure date.[9] The long reporting lag of 10-Ks discourages investors who trade on time-sensitive information since most information in 10-Ks is released through either earnings announcements or previous 8-K filings. Empirically, the 10-K downloads are spread out throughout the year after their release, thereby validating the argument that 10-K downloads capture stale information acquisition. As a result, demand for 10-K filings is expected to

[7] The rationale of selecting firms with disclosure one day prior to the blackout is the following. All investors had access to all firms' filings for a whole trading day on August 14, and exogenous variation will be evident in downloads on August 15 because of the power outage. Extensive accounting and finance studies on price drifts have shown that the price discovery takes several days.

[8] For example, Deaves, Dine, and Horton (2006) survey 1,600 retail investors and find that the majority of shareholders read and use financial statements to make investment decisions.

[9] Starting in 2003, the filing deadline for 10-Ks is 75 days for accelerated filers and 90 days for non-accelerated filers. The average report lag in my sample is 81 days.

affect stock returns through the attention channel, rather than through the information asymmetry channel, which requires the timely processing of material information.

I show that, in stark contrast to material information acquisition, stale information acquisition predicts higher stock returns, consistent with the attention channel. Specifically, the long/short portfolio sorted on 10-K downloads earns an alpha of 40 bps in the first holding month but reverses to zero over the next year. The alpha decay pattern aligns with previous findings in the literature.[10]

This is not to say, however, that 10-K downloads do not reduce information asymmetry. I further show that, it is not 10-K versus 8-K downloads per se driving the distinctive predictions, but rather the timeliness of the information acquisition. Given the low disclosure frequency of 10-K filings, over 80% (60%) of 10-K downloads happen after the first month (quarter) of release. As a result, demand for 10-K filings is generally dominated by the attention channel. When I limit the sample to firms with the 10-K release, demand for newly filed 10-K in the given week predicts a lower return going forward, similar to the prediction of 8-K downloads. In other words, the effect of 10-K downloads switches from the attention channel to the information asymmetry channel if the acquired information is relatively new to the market.

The paper proceeds as follows. Section 2 discusses the related literature. Section 3 describes the sample selection and data collection. Section 4 jointly tests the information asymmetry channel and the attention channel. Section 5 discusses the mechanisms of material information acquisition. Section 6 shows the differential effects of material information acquisition conditional on ex-ante information asymmetry, cost of information acquisition, and information content. Section 7 concludes.

[10]See Gervais et al. (2001), Chen, Hong, and Stein (2002), and Gargano and Rossi (2018), for example.

2 Related Literature

Whether information acquisition reduces information asymmetry is not a trivial question. There is a large theoretical literature on information acquisition and information asymmetry. One stream of the literature argues that information acquisition reduces information asymmetry (see, e.g., Grossman and Stiglitz (1980), Verrecchia (1982), Wang (1993), Easley, Hvidkjaer, and O'Hara (2002)), while another stream argues the opposite (see, e.g., Barlevy and Veronesi (2007),Veldkamp (2006),Mele and Sangiorgi (2015)). The key ingredient driving the opposite predictions is how the signal of acquired information is correlated with the noise. My paper contributes to the literature by offering the first piece of empirical evidence on information acquisition reducing information asymmetry in the context of unscheduled corporate filings. Prior literature has empirically tested whether information asymmetry affects stock returns (Easley et al. (2002), Kelly and Ljungqvist (2012)). My paper takes a step further by directly testing how information acquisition affects returns through the information asymmetry channel and by exploring its heterogeneous effect through the cost of information acquisition channel.

My paper contributes to the literature on investor attention by highlighting the different aspects of information acquisition on stock returns. Prior literature has focused on the attention channel by examining abnormal trading volume (Gervais et al. (2001)), Google Trends (Da et al. (2011)) and Bloomberg Search Index (Ben-Rephael et al. (2021)). These aggregated measures cannot differentiate sources of information. I show that the timely acquisition of unanticipated material information reduces information asymmetry and lowers stock returns, whereas collecting stale information reflects investors' increased attention and predicts higher abnormal returns. Only by taking different types of information acquisition into account can we distinguish between the attention channel and the information asymmetry channel.

My paper also contributes to the growing literature that use EDGAR log data to examine investor attention. For example, Drake, Roulstone, and Thornock (2016) study downloads of historical accounting reports around earnings announcement; Li and Sun (2022) find positive relation between 10-K downloads and subsequent stock returns; Chen, Cohen, Gurun, Lou, and Malloy (2020) examine informed information acquisition of insider trading filings by institutions ; Gibbons, Iliev, and Kalodimos (2021) study analysts' downloading patterns; Iliev, Kalodimos, and Lowry (2021) investigate investors' attention to governance research; and Cao, Kilic, and Wang (2020) examine investor attention and divergence of opinions. My paper differs from the existing literature by examining the difference in return predictability of material and stale information acquisition. Moreover, utilizing geographical information from IP visits and a natural experiment of the Northeast Blackout, my paper establishes a causal link between material information acquisition and information asymmetry.

3 Data and Sample Selection

The paper uses data from several sources. I use the CRSP, Compustat, I/B/E/S, and TAQ databases to obtain stock related information, the Thomson Reuters database to obtain institutional ownership data; the EDGAR server log to obtain daily logs of page requests for SEC filings;[11] and the EDGAR Master File to obtain filing type and date. To control for media coverage, I use RavenPack news data. RavenPack news data provide news coverage for a large sample of public companies[12]. I also control for Google Trends and Bloomberg News Heat Index. Google Trends data provide an index of the volume of within-firm daily Google searches and are often used to capture retail investors' attention. Bloomberg News Heat Index captures the news search volume by Bloomberg users and is used to capture institutional investors' attention, which is available starting from February 17, 2010.

The sample starts in 2003 when the EDGAR log data became available and ends in 2016. I use all domestic equity stocks with share code 10 or 11. I require stocks with a valid market value at month-end in the CRSP, valid financial statement data in Compustat, and

[11] I use the link file provided by WRDS to link stock identifiers "permno" in CRSP and "cik" in SEC.

[12] I match RavenPack data with the CRSP data using 8-digit CUSIP, ticker symbol, and company names.

valid earnings announcement data in I/B/E/S. I also require that stocks in the CRSP have matched identifiers in the SEC EDGAR database. The matched sample has 5,989 unique stocks. After merging with RavenPack and Google Trends data, the sample reduces to 4,106 unique stocks, where most of the sample loss occurs for micro-cap stocks. For the main analysis, I use the full sample. All my results are robust when using the smaller sample.

3.1 The EDGAR Server Log

The EDGAR log is publicly available and can be obtained from its website. The data contain daily log files from 2003 forward. The log file contains the timestamps of page requests, the firm identifier, the filing accession number, the IP address of the request[13], the index page flag[14], server status code[15], the crawler flag, and so on. Log files between September 24, 2005, and May 10, 2006, were labeled by the SEC as "lost or damaged" and are excluded from the empirical analysis. Some users employ automated programs to crawl SEC filings, but not all crawling activities are flagged by the EDGAR log. Following Lee, Ma, and Wang (2015), I label an IP address as a crawler if it is associated with more than 50 daily requests.

The sample starts with over 21.89 billion records. Following Lee et al. (2015), I first reduce the sample to 9.84 billion records by excluding requests with the index page flag or failed connection, since these requests do not capture any information acquisition activities. I then link the Central Key Index (CIK) provided by EDGAR with the stock identifier in CRSP. After the merge, the sample reduces to 3.36 billion records. I further reduce the sample by focusing on filings of the following three types—Forms 10-K, 10-Q, and 8-K—which leaves me with 1.36 billion records.

Finally, I get the physical locations and service providers of each IP address in EDGAR through several "IP Address Lookup" services.[16] Since EDGAR hide the last octet of an IP

[13]Only the first three octets of the IP address are available. The last octet is replaced with random characters so that the IP address is uniquely identifiable.

[14]There is an index page containing all documents for a filing. The index page flag indicates that the user visits the index page without downloading any documents.

[15]The server status code, which is typically below 300, indicates whether the request is successful.

[16]See https://ip-lookup.net, https://www.iplocation.net, and https://iplocation.io.

address, I search a wide range of IP addresses that share the same first three octets and get a majority vote on the location and provider. Similar procedure is also used in Chen et al. (2020). Lastly, for each IP address, I use the returned longitude and latitude coordinates to pin down its location, and manually identify service providers that belong to financial institutions.

3.2 Overview of the EDGAR Downloads

Figure A1 shows the monthly aggregated downloads in my sample. I separate crawling activities ("robots") from human viewing activities ("human"). Figure A1a shows the plot for all filing types. The trend for viewing activities on EDGAR has been increasing. The number of human downloads starts at 0.25 million in 2003 and ends at 1.5 million in 2016. The number of crawling requests is about 15 times greater than the number of human downloads. Figures A1b to A1d show the monthly aggregated plots by file type. The 10-K and 10-Q downloads exhibit strong seasonality driven by the filing cycles. Conversely, 8-K downloads exhibit weak seasonality because 8-K filings are triggered by unanticipated material events. In terms of the aggregated magnitude, 10-K downloads account for around half of all downloads, with the remaining half split by 8-K and 10-Q downloads.

4 Two Channels of Information Acquisition

In this section, I use three approaches to test the effects of information acquisition on returns through the attention channel and the information asymmetry channel. First, I run monthly Fama-MacBeth (1973) cross-sectional regression of returns on measures of information acquisition. Second, I use a non-parametric approach by forming long/short portfolios and regressing portfolio returns on factors. Third, I use weekly level data to provide high frequency and long-term results.

4.1 Fama-MacBeth (1973) Approach

I first study the relation between future stock returns and the overall demand for information. I estimate a Fama-MacBeth (1973) regression of monthly individual stock returns from month $t+1$ on information acquisition variables from month t.

All regressions control for the following characteristics. For firms' fundamental variables, I include *Asset Growth*, *log(BM)*, *log(ME)*, and *Operating Profit*. *Asset Growth* is the annual growth rate of assets; *log(BM)* is the natural logarithm of the book-to-market ratio; *log(ME)* is the natural logarithm of the firm market capitalization; and *Operating Profit* is the ratio of operating profits to book equity. I include the current month stock return $r_{1,0}$ and the past-12 month stock return $r_{12,2}$ to control for firms' past performance, which may drive both investor demand and future returns. Gervais et al. (2001) and Barber and Odean (2007) document that abnormal trading volume increases a firm's visibility, which could affect both demand and future stock returns. Therefore, I include *Abnormal Trading Volume*, which is the difference between monthly trading volume and past 12-month average trading volume, scaled by the standard deviation of the past 12-month trading volume. Since many of my information acquisition variables capture information acquisition of firms' annual and quarterly filings, I include earnings surprise and earnings drift from the most recent earnings announcement. *SUE* is the unexpected quarterly earnings scaled by market cap; *Earnings Drift* is the sum of daily returns in three days around an earnings announcement. To control for the firm's information environment, I include the number of analysts, the firm's past 8-K filing frequency, media coverage using RavenPack news count, and change in Google Trends. Lastly, I control for firm disclosure. The variables *file 8K*, *file 10K*, and *file 10Q* are the number of Forms 8-K, 10-K, and 10-Q filed in the EDGAR system in the given month, respectively. Column (1) of Table 1 shows the baseline result. Consistent with the previous literature, asset growth, firm size, operating profit, unexpected earnings, abnormal trading volume, and abnormal earnings announcement returns can explain the cross-section of stock returns.

Column (2) of Table 1 shows the effect of aggregated information acquisition on subsequent stock returns. The variable *log* $views_{all}$ is the natural logarithm of all filing downloads of the firm in the current month. The coefficient estimate of *log* $views_{all}$ is positive and significant. Firms with high filing downloads earn a premium of roughly 20 bps per month (2.4% per year), which is consistent with prior studies that use direct measures of information acquisition, such as Google Trends, or indirect measures, such as abnormal trading volume and extreme stock returns.

To disentangle the information asymmetry channel from the attention channel, I split the aggregated downloads by filing types. Demand for 10-K and 10-Q filings is more likely to capture the general demand for assets, as Forms 10-K and 10-Q provide investors with a comprehensive overview of the firm. Information on the firm's balance sheet is also widely used to make fundamental investing decisions. Moreover, 10-K/Qs are often filed with significant delays, so the demand for these filings responds to information that is not time sensitive. Note that it is highly likely that some information contained in 10-K/Qs is new to investors, and downloading this information can reduce information asymmetry between insiders and investors. However, given that around 80% of downloading activities happen one month after the disclosure, the attention channel dominates the information asymmetry channel in terms of the return patterns. Filings of Form 8-K, on the other hand, are generally unanticipated, and the forms contain information that is privately known by insiders. Under the SEC disclosure regulation, the material information needs to be disclosed within four business days. Therefore, demand for 8-K filings is more likely to transform the disclosed information into public information and reduce the information asymmetry. Moreover, the timely downloading pattern of 8-K filings supports this conjecture. In contrast to 10-K downloads, 41% (67%) of 8-K downloads happen within the first week (month) of disclosure.

Column (3) of Table 1 shows the effects of information acquisition through two channels. The variable *log* $views_j$ is the monthly natural logarithm of Form j's downloads. On the one hand, the coefficient estimate of *log* $views_{8K}$ is significantly negative (-11.5 bps per month), consistent with the prediction of the information asymmetry channel. Demand for 8-K filings reduces the information asymmetry between insiders and investors, decreases the payoff uncertainty, and leads to lower subsequent returns.[17] On the other hand, the coefficient estimate of $\textit{log views}_{10K}$ is significantly positive (38.8 bps per month), which is consistent with the prediction of the attention channel. Moreover, the effect of 8-K downloads on stock returns is empirically smaller in magnitude than the effect of 10-K downloads. This result explains why using aggregated measures of information acquisition only finds evidence of the attention channel but overlooks the information asymmetry channel. In fact, what we have seen in the previous literature is the overall effect of information acquisition on stock returns, and my paper is the first to document each channel separately. Lastly, the coefficient of $\textit{log views}_{10Q}$ is insignificant, and its magnitude is relatively small. Two reasons account for this result. The correlation between 10-K downloads and 10-Q downloads is 0.91 over the full panel. As a result, 10-Q downloads do not provide additional variation beyond 10-K downloads in explaining stock returns. Moreover, the substance and quality of Forms 10-K and 10-Q differ. Filings of Form 10-K are required to be audited, whereas filings of Form 10-Q are not. In addition, the MD&A section in Form 10-K is much more detailed than in Form 10-Q.[18] As a result, Form 10-K is a more reliable source as an investment reference than Form 10-Q.

To sharpen the idea of acquiring unanticipated and material information through the information asymmetry channel, I decompose demand for 8-K filings into two parts: demand for scheduled and unscheduled 8-K filings. The scheduled 8-K filing includes pre-scheduled events, such as earnings announcements and annual shareholder meetings. These scheduled reports typically contain information that is well anticipated by the market. Therefore, demand for scheduled 8-K filings should have no predictability under the information asymmetry channel. Following the literature, scheduled 8-K filings are categorized as the ones with

[17] I postpone the analysis for the direct effect of 8-K downloads on information asymmetry to Section 5. Here, I mainly focus on the effect of information acquisition on stock returns.

[18] For example, the MD&A section of IBM Form 10-K spans 50 pages in 2018 and only 20 pages in 2019Q1.

Item 2.02 or contain keywords such as "*scheduled*". Unscheduled filings disclose material information that was known only by insiders. It requires investors to timely collect, process, and incorporate the information into the market. Therefore, acquiring unanticipated and material information can reduce the information asymmetry between insiders and investors, leading to lower subsequent returns. In column (4) of Table 1, I show that the return predictability of demand for 8-K filings is entirely driven by unscheduled 8-K downloads, which further supports the information asymmetry channel.

To further sharpen the identification of new versus stale information acquisition, I calculate the time difference between each filing download and the disclosure. I find that 41% of 8-K downloads happen within the first week of disclosure, and the speed gradually slows down, reaching 67% by the end of the first month. I then split the unscheduled 8-K views by whether the download happens within the first week of disclosure.[19] As shown in column (5) of Table 1, the effect of unscheduled 8-K downloads comes entirely from new information acquisition, consistent with the information asymmetry channel. In column (6), I control for the change in Google Trends and media coverage, and the result is robust. The sample size becomes smaller than the ones in previous columns, as some micro-cap firms are not covered by the RavenPack dataset.

4.2 Portfolio Sort Approach

The previous section demonstrates the effects of information acquisition through the attention channel and the information asymmetry channel using Fama-MacBeth regressions. In this section, I provide additional supporting evidence using a portfolio sort approach. This approach not only is less parametric than the Fama-MacBeth regression but also has the flexibility to show the effects over a longer period. I show that 8-K and 10-K downloads predict not only opposite returns over the subsequent month, but also yield different return paths over the long term. For tests related to the information asymmetry channel, I focus

[19] The result is robust if I use the first month as the cutoff.

on a set of firms with unscheduled 8-K filings.

Since large firms naturally receive more downloads than small firms, it is important to control for firm size when sorting on downloads, especially for the 10-K filings.[20] Each month, I sort stocks into quintiles using NYSE breakpoints. Conditional on each NYSE quintile, I then sort stocks by the number of downloads into quintiles. Finally, I form a long/short portfolio by buying the top quintile stocks and selling the bottom quintile stocks and regress the monthly portfolio returns on benchmark factors. The factor models include CAPM, Fama-French three-factor (FF3), Fama-French-Carhart (FFC), Fama-French five-factor plus momentum (FF5+UMD), and an eight-factor model by including betting-against-beta and liquidity factors.

Table 2 shows the univariate portfolio sort result for 10-K and 8-K downloads. The results are consistent with the ones using the Fama-MacBeth regressions. In the short term, 10-K downloads predict positive abnormal returns, whereas 8-K downloads predict negative abnormal returns. After controlling for common pricing factors, the average alpha of the long/short 10-K portfolio is around 0.42% per month. Moreover, the effect of 10-K demand is short-lived, which can be seen from the insignificant alphas with 3 or 12 holding months. As shown in Panel C, the alpha of the long/short 10-K portfolio is mainly driven by its long leg, stocks with high size-adjusted 10-K views, consistent with the attention channel.

Conversely, the long/short 8-K portfolio earns a monthly alpha of -0.48%, and the alpha comes from its short leg, consistent with the short-term prediction of the information asymmetry channel. Stocks with a small number of 8-K views still face a high level of information asymmetry, and the positive alpha reflects the excess return to traditional asset pricing models. Moreover, the effect of 8-K demand is long-lasting, averaging -0.5% per month for the next 12 months.[21]

[20]The correlation between the log of 10-K downloads and firm size is 0.56, and the correlation between the log of 8-K downloads and firm size is 0.34.

[21]Table A3 separate the robot crawling activity into retail and institutional crawling activities. Institutional crawling for unscheduled 8-K filings predicts lower subsequent returns, whereas retail crawling has no predictability. The result is consistent with the increased popularity of algorithm trading employed by institutional investors.

The results are robust using alternative approaches. In appendix Table A2, I first run a cross-sectional regression of the natural logarithm of downloads on a set of control variables and their quadratic terms.[22] Control variables include the natural log of firm size, the natural log of the number of analysts, abnormal trading volume, and idiosyncratic volatility in the past 12 months. I then form portfolios based on the regression residuals. The result is very close to the baseline result in Table 2, because firm size itself explains more than 50% of the cross-sectional variation in filing downloads, and the rest of the characteristics explain an additional 10%.

Since I use size-adjusted downloads as the sorting variables, it is interesting to see how 10-K and 8-K portfolios perform under different size groups. Each month, I first sort stocks into quintiles by their previous month market capitalization. Conditional on each size quintile, I then sort stocks into quintiles by the 10-K (8-K) downloads and form the long/short portfolio by buying the top quintile stocks and selling the bottom quintile stocks. Figure A3 shows the result. Portfolios sorted on size-adjusted 10-K downloads yield positive and significant alphas across all size quintiles. The result is the strongest in the small size quintile, yielding a 1.2% alpha per month. The magnitude of the alpha decreases with firm size. Both liquidity and the short-selling constraint contribute to the result. Small firms are more illiquid than large firms. When facing a demand shock, small stocks face higher trading pressure than large stocks. Moreover, small stocks have a tighter short-selling constraint than large stocks, which limits the potential arbitrage opportunities and results in a large price increase.

Portfolios sorted on size-adjusted 8-K downloads yield negative alphas across all size quintiles, but alphas are significant for the bottom three size quintiles and insignificant for large stocks. For example, the 8-K portfolio yields an average alpha of -40 bps per month for the bottom three size quintiles, and -10 bps for the top two quintiles. This result is consistent with the information asymmetry hypothesis. Small firms have less media/analyst coverage and institutional holdings than large firms. Investors of small firms face a higher degree of

[22] I add square terms to account for potential non-linear relationship between filing downloads and firm characteristics. The results are very similar without quadratic terms.

information asymmetry and rely more on themselves in processing and incorporating the disclosed information. Moreover, large firms have more channels to disseminate information, so the downloading activities on the EDGAR may not necessarily represent a significant portion of information acquisition. Therefore, demand for 8-K filings has a stronger effect on small firms than large ones.

4.3 Weekly Frequency Result

To zoom in on the supply and demand side of information, I then conduct analyses at a weekly frequency. Filings of Form 10-K generally occur once a year. Filings of Form 8-K, however, occur irregularly, but happen once a month on average. Therefore, the weekly-level analysis allows me to study the interaction between the supply and the demand for information and their effects on prices.

I aggregate the daily stock returns and daily downloads to a weekly frequency (Friday close to Friday close). My main variable of interest is $\log(views_t^k)$, which is the natural logarithm of total downloads of filing type k in week t. I then create a set of dummies to capture the information supply. The dummy variable $Filing\ k_t$ is equal to one if the firm has issued filings of type k in week t. The dummy variable $News_t$ is equal to one if the firm appears in the RavenPack news database in week t. The dummy variable $Earnings\ Release_t$ is equal to one if the firm releases its earnings in week t. I also control for firm size, book-to-market, trading volume, number of analysts, and institutional ownership. For a subset of the analysis,[23] I also control for the Bloomberg Search Index and Google Trends, which capture the institutional and retail demand studied in the previous literature (Ben-Rephael et al. (2021), Da et al. (2011)). The dummy variable AIA_t is equal to one if the Bloomberg News Heat daily index has a maximum of 3 or above in week t. The dummy variable $DADSVI_t$ is equal to one if the Google Trends daily index on any day of the week is above its 90th percentile in the past month.

[23] Bloomberg News Heat index is only available after 2010/02/17.

Table 3 shows the weekly regression result. I regress the weekly stock returns on demand for filings, controlling for the supply of firm filings, media coverage, earnings announcements, the amount of analyst coverage, institutional ownership ratio, firm size, book-to-market ratio, trading volume, lag returns, and time fixed effects. To capture the interaction between the supply and demand of information under the information asymmetry channel, I add an interaction term between the supply and demand of 8-K filings. Columns (1) and (3) study the contemporaneous relation between stock returns and demand for information, where the dependent variable is the current week stock returns. The dependent variables in columns (2) and (4) are stock returns in the subsequent week.

As shown in columns (1) and (2) of Table 3, the coefficient estimates of $\log(views_t^{10K})$ are all positive and significant, consistent with the attention channel that is well documented in the existing literature. The coefficient estimate of the interaction term between 8-K supply and demand is positive and significant in column (1), and negative and significant in column (2). The results have not been documented empirically, and they support the information asymmetry channel. When firms release new information through 8-K filings, acquiring information reduces the information asymmetry between insiders and investors, and stocks become less risky to hold. Therefore, the contemporaneous price increases, and the future return decreases. Both the attention channel and the information asymmetry channel are robust after controlling for the Bloomberg and Google Trends search indexes, as shown in columns (3) and (4).

The two channels of information acquisition not only have opposite predictions in the short term, but also suggest distinct patterns in the long term. For the attention channel, we should see the strongest evidence of a positive contemporaneous return spread, followed by an alpha decay pattern. The speed of the alpha decay process relies on the liquidity of the underlying asset and the time lag between information acquisition and the investment decision. For the information asymmetry channel, although the contemporaneous price also increases, the underlying mechanism is completely different. The reduced returns in the future drive up the contemporaneous price, which is followed by a negative return spread in the future.

To test the long-term return prediction pattern, at the end of each week, I first sort stocks by size into five groups using NYSE breakpoints. Conditional on each NYSE size-group, I sort stocks into quintiles by the weekly 10-K and 8-K downloads and form long/short portfolios. Portfolios are held throughout the next 24 weeks. The alphas of portfolios at each holding week are plotted. For 8-K portfolios, I limit the set of stocks that issue unscheduled 8-K filings in the week, as the evidence suggested in Table 3 shows that the effect of 8-K demand is stronger, conditional on the supply of information. The 8-K portfolios are sorted by the size-adjusted views in the first week after the filing. Figure 1 shows the long-term return patterns of both channels. Consistent with my hypothesis, 8-K downloads capture material information acquisition and predict an increase in the contemporaneous price, followed by lower returns in the future, supporting the information asymmetry channel. Interestingly, the negative return spread is persistent over time. Two reasons account for this persistence. First, note that only stocks with unscheduled 8-K filings are included in the portfolio formation at each given week, and the probabilities of subsequent unscheduled filings do not vary much across high and low groups. Specifically, the probability of filing an unscheduled 8-K for the high (low) group in any subsequent weeks is around 24.7% (23.1%). Second, the unscheduled 8-K downloads are not highly persistent over time. Conditional on being in the top (bottom) quintile, the probability of being in the top (bottom) quintile next time is 34% (33%). In appendix Figure A7, I also conduct robustness checks using an event-study approach and residual 8-K views. The results are very similar.

Conversely, 10-K downloads capture stale information acquisition and predict an alpha-decaying pattern, supporting the attention channel. These results do not imply that 10-K downloads do not affect information asymmetry. Rather, it merely states that the attention channel dominates for the annual filing of Form 10-K, and it is hard to empirically disentangle the two channels because of the low disclosure frequency. To test whether demand for newly disclosed 10-K filings reduces information asymmetry, I limit my sample to a set of stocks that just disclosed Form 10-K in a week. I then sort these stocks into quintiles based on the size-adjusted downloads of the newly issued 10-K. The result is plotted in Figure 2. Conditional on firms just issuing 10-Ks in week 0, firms with higher downloads of newly released 10-Ks yield higher returns in the contemporaneous week, and lower returns in the upcoming weeks. The effect of 10-K downloads switches to the information asymmetry channel in this small subset. The long-term return pattern is comparable to the one found in 8-K filings. However, the alpha is noisily estimated, since only a small portion of firms file 10-Ks in each week.

4.4 Evidence from Volatility

Having demonstrated that 8-K downloads lead to lower stock returns in the future, I then test whether volatility changes following a spike in information acquisition. For each stock in each month, I calculate the volatility of daily returns within the month. I also decompose the volatility into systematic volatility and idiosyncratic volatility using the Fama-French five factors and momentum factor loadings estimated with a three-year rolling window. For a stock with unscheduled 8-K filings in a given month, I then calculate the abnormal volatility by subtracting the volatility from its one-year volatility average. Lastly, I sort stocks into quintiles by size-adjusted 8-K downloads and plot the difference in abnormal volatility between the top and bottom quintiles.

Figure 3 shows the result. Compared to stocks with low 8-K downloads, stocks with high 8-K downloads show a spike in volatility in the 8-K disclosure month, and persistently lower volatility in subsequent months. The initial spike in volatility suggests that investors are learning through downloading the 8-K filings, which resolves uncertainty in subsequent periods. This pattern is also consistent with predictions documented in Andrei and Hasler (2015). Moreover, it is the idiosyncratic volatility driving the result. Note that most of the 8-K filings are event-based and highly idiosyncratic in nature, which explains why the reduction in volatility comes from the idiosyncratic component. Meanwhile, the information risk is not captured by the Fama-French factor models and enters the idiosyncratic component when regressing stock returns on these factors. Furthermore, the prediction power of 8-K downloads on returns is mainly from small stocks (the bottom three size quintiles shown in Figure A3). These stocks are widely held by individuals and households so that the idiosyncratic risk is not easily diversified away.

To sum up, the effects of material information acquisition on stock returns and volatility are consistent with the information asymmetry channel. Upon releasing new information through the 8-K filings, investors learn more about the firm by downloading the filings, which reduces the asymmetric information between uninformed investors and insiders. Therefore, contemporaneous volatility goes up, followed by a persistent reduction in future volatility.

5 Mechanisms of 8-K Demand on Stock Returns

Easley and O'Hara (2004) document that investors demand higher returns for stocks with more private information. Boot and Thakor (2001) suggest that disclosing information that is only known to informed investors decreases the information advantage that informed investors have over the uninformed. However, little research has shown the effect of information acquisition on information asymmetry, as the past literature almost exclusively focuses on the supply side. In this section, I show that investors' 8-K demand decreases the firm-level information asymmetry. As a result, the stock becomes less risky for uninformed investors to hold, and the subsequent return decreases.

5.1 Panel Regressions of Information Asymmetry

To test whether demand for information reduces the information asymmetry, I estimate a weekly panel regression of future information asymmetry on information acquisition, controlling for firm characteristics and information disclosure. The result is shown in Table 4.

I use the price impact measure estimated following Holden and Jacobsen (2014) as a proxy for the information asymmetry of the firm.[24] For a given stock, the price impact on the k^{th} trade is defined as

$$Price\ Impact_k = \frac{2D_k(M_{k+5} - M_k)}{M_k}, \tag{1}$$

where M_{k+5} is the midpoint five minutes after the midpoint M_k, and D_k is the buy-sell indicator of the trade. The price impact measure captures the permanent component of the effective spread and is widely used in the microstructure literature to proxy for firm-level information asymmetry. In column (1), the coefficient estimate of $log\ views_{8K}$ is negative and significant, suggesting that a higher number of 8-K downloads is associated with lower information asymmetry in the next week. The supply of 8-K filings also reduces information asymmetry, as can be seen by the negative coefficient estimate of *Filing 8K*. However, once we interact the demand and supply of 8-K filing, the supply of 8-K does not have any significance, as shown in column (2). The interaction term between 8-K demand and supply is negative and significant, showing that the demand for 8-K filings has a stronger effect on reducing information asymmetry, conditional on new information arrivals. The economic magnitude is also large. A 10% increase in 8-K downloads leads to a 9 bps reduction in the price impact measure, which has an average of 45 bps.

Consistent with the mechanism, the effect of 8-K downloads on stock returns should be larger when the ex-ante information asymmetry is higher. The previous section (Figure A3) shows that the effect is indeed larger for small stocks than for large stocks, as market capitalization is one of the most important proxies for information asymmetry. As additional robustness checks, I also use the Amihud illiquidity measure and previous quarter analyst forecast dispersion measure to proxy for ex ante information asymmetry. I first sort stocks by these measures into terciles. Conditional on each tercile, I sort stocks by 8-K downloads within the NYSE size-group into quintiles. Table 5 shows the portfolio double-sort results

[24] I have also used alternative information asymmetry measures, such as Amihud illiquidity measure, bid-ask spread, or PINs. The results are similar and available upon request.

for 8-K downloads and the ex-ante information asymmetry. When the Amihud measure is low, the alpha of the long/short 8-K downloads portfolio is -11 bps per month. When the Amihud measure is high, the magnitude of alpha increases to -53 bps per month. The results are similar using the analyst forecast dispersion measure.

5.2 Natural Experiment: Northeast Blackout of 2003

Given the endogenous nature of information acquisition and information asymmetry, the panel regression result documented in the previous section could be driven by reverse causality. For example, investors may be aware of firms with high information asymmetry. Upon releasing new information, firms with higher information asymmetry attract more downloads, and the supply of information reduces the information asymmetry. Then we may observe a spurious relation between information acquisition and information asymmetry. To rule out such endogeneity concerns, we need a shock to information acquisition, which is orthogonal to the firm characteristics and the supply of information. Therefore, in this section, I test how demand for 8-K filings causally reduces information asymmetry using the Northeast Blackout of 2003 as a natural experiment. On Thursday, August 14, 2003, a widespread power outage occurred throughout parts of the northeastern and midwestern United States, beginning just after 4:10 p.m. EDT. For the next 30 minutes, outages were reported in parts of Ohio, New York, and New Jersey, including in the major cities of New York City, Toronto, Baltimore, and Detroit. Manhattan, including Wall Street, was completely shut down. Although Wall Street got its power back at 6 a.m. on August 15, most traders were not able to commute because of the shutdown of the train system, and the infrastructure suffered from the power outage. Shares were lightly traded on Friday, and the NYSE ended up with just under 624 million shares traded. Figure 4 shows the EDGAR 8-K hourly downloading traffic by investors from the affected and non-affected regions. The number of hourly downloads on August 15 fell by a lot, especially for the affected regions.

The Northeast Blackout provides an opportunity for a natural experiment. Since demand for 8-K filings is an endogenous choice, the ideal experiment would be to have an exogenous shock to the endogenous choice. In a simplistic world, suppose we have only two types of firms—type A and type B—that just disclosed unanticipated information. Before the shock arrives, type A firms have a higher proportion of 8-K downloads from the affected regions than type B firms. The geographical distribution of attention allocation before the shock can be completely endogenous. When the shock arrives, it would (ideally) shut down all information acquisition from the affected regions, and investors from non-affected regions would be free to acquire information. Such a shock affects the information acquisition of investors for the two firms differently. Investors who would have acquired information on type A firms during the blackout cannot do so and stay uninformed during the blackout, which leaves the firm-level information asymmetry at a relatively high level.

The difference-in-differences estimation is a natural fit for this problem. In this setting, I have two periods: August 14, 2003, and August 15, 2003. I first calculate the fraction of historical 8-K downloads from the affected regions, denoted as *frac*, for each firm before August 13, 2003. I then limit my sample to firms that just disclosed material information on August 13, so that information acquisition is crucially important for both types of firms. I also limit my sample to firms with headquarters outside the affected regions, which alleviates the concern that firms from the affected regions may suffer additional economic disruption relative to firms outside the affected regions. I define a firm as treated if the fraction *frac* is above the median. The regression specifications are the following:

$$Price\ Impact_{i,t} = \beta_0 + \beta_1 treated_i + \beta_2 post_t + \beta_3 treated_i \times post_t + \epsilon_{i,t},$$

$$Price\ Impact_{i,t} = \beta_0 + \beta_1 frac_i + \beta_2 post_t + \beta_3 frac_i \times post_t + \epsilon_{i,t},$$

where dummy variable *post* is equal to one if it is the period of August 15, 2003, and zero otherwise. The first regression specification considers the binary case, whereas the second specification considers the continuous treatment effect. The dependent variable is the daily *Price Impact* measure. The regression results are presented in Table 6. The coefficient on the interaction term is positive and significant, suggesting that the reduction in information acquisition increases the firm-level information asymmetry. In terms of the magnitude, firms with 1% more downloads from the affected regions before the shock suffer a 3% increase in information asymmetry after the blackout. To alleviate the concern that headquarter locations may not perfectly control for business activities, I get the frequency of state name appearances from the firm 10-K filings as a proxy for state-level operation intensity, and exclude firms with top state name appearances in the affected regions. The results do not change in the smaller sample.

I also test and rule out the following alternative explanations. First, it could be that traders of firms with a higher fraction of historical downloads from affected regions cannot trade as freely as traders of firms with a lower fraction. Second, it could be that the blackout creates a selection among traders, so that the level of sophistication between two types of firms has changed. In Table 7, I use trading volumes normalized by the number of shares outstanding to rule out the first explanation, since there is no differential effect in trading volumes among firms due to the blackout. I also use the number of trades after the market close to capture trader's sophistication level. There is no evidence to suggest that the composition of sophisticated traders has changed.

Moreover, the result in Table 6 is not driven by firm characteristics that correlate with the fraction of downloads from affected regions. For each trading day t between July 13 and August 12, I repeat the analysis by selecting firms with unscheduled 8-K filings and running a placebo difference-in-differences regression with the same procedure as in Table 6, where $t + 1$ and $t + 2$ are the pseudo pre- and post-treatment days, respectively. None of the interaction terms is statistically significant, suggesting that the blackout result is not simply driven by the fraction of downloads from affected regions per se. Instead, the blackout enabled an exogenous variation in downloads across firms.

6 Heterogeneous Effect of 8-K Demand

This section shows the heterogeneous effects of information acquisition through the information asymmetry channel. Specifically, I examine how the effect varies with the cost of information acquisition and the information content.

6.1 Cost of Information Acquisition

The cost of information acquisition plays an important role in reducing information asymmetry. In Verrecchia (1982) Corollary 4, the informativeness of price increases when the information acquisition cost is reduced. Although I do not directly observe the cost of information acquisition for each investor, an investor's past information acquisition history and his/her geographical location are observed in the data. Moreover, institutional investors tend to have a lower cost of information acquisition than retail investors. I use the firm-level fraction of local demand, the fraction of recurring viewers, and abnormal institutional investors' attention to capture the cost of information acquisition.

Local investors have an information advantage over non-local investors in collecting and processing information. Therefore, holding the level of information acquisition fixed, firms with more local demand for information have a lower cost of information acquisition. Moreover, I make the explicit assumption that the cost of information acquisition is lower for an investor who acquired information on the firm in the past quarter than one who did not. Therefore, the recurring visitor ratio defined in Figure A6 can be used as a proxy for the cost of information acquisition. The higher the recurring visitor ratio is, the lower the cost of information acquisition.

I also take advantage of the names of IP service providers (ISP) and classify investors' downloads by whether the ISP is from an institution, such as banks and investment companies. I then calculate the fraction of 8-K downloads from institutions. To capture the spike in institutional attention, I create an abnormal institutional attention variable, ab_inst^{8K}, which is the fraction of institutional 8-K downloads normalized by the past 12-month average and standard deviation. The higher the abnormal institutional attention measure is, the lower the cost of information acquisition.

Panel A of Table 8 shows the portfolio double-sort results by 8-K demand and the average distance of viewer location to firms' headquarters. For each stock in each month, I calculate the average distance between IP addresses and firm headquarters for each filing type. I then double sort stocks by the average distance into terciles and by the downloads within the NYSE size-group into quintiles. The effect of 8-K demand is mainly concentrated in the low (-61 bps/month) and medium (-41 bps/month) distance tercile. Moreover, the difference between high and low terciles is statistically significant.

Panel B of Table 8 studies the effect of 8-K demand on prices, conditional on visitors' past visiting patterns. For each firm-month, I calculate the proportion of recurring visitors. I then double sort stocks by the frequency ratio into terciles and by 8-K downloads within the NYSE size-group into quintiles. Portfolios sorted by 8-K demand show significant and negative alphas when downloads are from recurring visitors (-55 bps/month). When the recurring ratio is low, however, the 8-K portfolio yields an insignificant alpha. These results suggest that the effect of information acquisition on stock returns is magnified when investors have a low cost of information acquisition.

Panel C of Table 8 studies the effect of 8-K demand on prices, conditional on the abnormal institution attention. Portfolios sorted by 8-K demand show more negative alphas when there is a spike in institutional investors' attention. For example, conditional on high abnormal institutional attention, the 8-K portfolio yields an alpha of -70 bps in the subsequent month, compared to -49 bps when the abnormal institutional attention is low. The difference is significant at a 10% level.

6.2 Information Content

Moreover, the effect of 8-K demand should be a function of the information content provided in the filings. The demand for information reduces the information asymmetry only if the information provided by the firm was previously private. Some filings, such as reports about the pre-scheduled meetings, do not convey any private information. Others, such as a material agreement and change of officers, require investors' attention to interpret the information. Therefore, it is important to see how the effect of information acquisition interacts with the information content provided in the filings.

I extract the "event date" and "post date" for each filing and calculate the three-day market excess abnormal return of the firm around both dates.[25] Two measures are then used to quantify the importance of each filing. The first measure is simply the maximum of absolute abnormal returns around event and post dates. This measure captures the market response to the information provided in the filing. If the new information is good (bad) news, the measure is high (low). If the information conveyed in the filing is already anticipated or even well understood by the market, the measure should be small in absolute terms. In my sample, the measure has a mean of 0.4% and a standard deviation of 12%. Figure 5 shows the histogram of the measure.

The second measure is constructed using textual analysis and machine learning. For each filing i, I build a document classifier based on the past one-year 8-K filings of all firms in my sample. I then compute the document similarity vector between the filing i and all past year filings. The similarity vector represents how similar the pair of documents is. I calculate the expected market response to filing i as the weighted average of three-day abnormal returns of filings in the past year, with the weight determined by the similarity vector. The expected market response captures what the abnormal return level should be, given the similarity of information content between the filing i and past filings. Lastly, I calculate the

[25] Starting in 2004, the SEC requires firms to disclose any material information within four days of the event. In practice, however, the lag can be more than four days, as firms can ask for some additional grace periods.

difference between the realized market response and the expected market response, and use this "unexpected market response" as a proxy for information importance. The measure has a mean of 0.1% and a standard deviation of 10%. The difference between the two measures is that the second measure more clearly captures the shock in information content beyond the part expected by the market.

To see how the effect of 8-K demand varies with the importance of information content of the filing, I double sort stocks by the 8-K downloads and the above two measures. The result is shown in Table 9.[26] In Panel A, the information importance measure is the raw abnormal cumulative return around the event. In Panel B, the information importance measure is the unexpected abnormal return. Both panels yield a similar result. The relation between the effect of 8-K downloads on returns and the cumulative abnormal return around the event date exhibits a "V-shape". The effect of 8-K demand is concentrated in the low and high abnormal return terciles, and relatively weak in the middle tercile, where the average abnormal return is around zero. When abnormal returns are high (low), firms are likely to have disclosed good (bad) private information. The demand for 8-K filings then plays an important role in interpreting the piece of information and reducing information asymmetry, which leads to a negative spread in future returns, regardless of whether the information itself is good or bad. However, when there is little market reaction around the event/post date, it is likely that the market has already taken the information content into account, which leaves investors with little to learn. As a result, the spread in 8-K demand does not predict future returns.

7 Conclusion

In this paper, I empirically test two channels where the demand for information affects asset prices. On the one hand, acquiring unanticipated material information reduces information

[26]In untabulated results, I also limit the filings to be in the same category. The sample size is about 23% smaller because of the matching. The results are quantitatively similar and available upon request.

asymmetry, which in turn drives up contemporaneous prices, followed by lower subsequent returns. On the other hand, investors acquire stale information because of attention-grabbing events. Such information acquisition reflects investors' increased attention and predicts positive abnormal returns that quickly decay over time. Two channels of information acquisition have opposite predictions for subsequent stock returns, yet the empirical literature so far has only provided evidence for the latter channel. My findings point out that it crucially depends on what information is acquired in order to disentangle the two channels.

Empirically documenting how information acquisition reduces information asymmetry has important implications. First, this is the first paper to empirically show that the demand side of information reduces firm-level information asymmetry. Extensive literature has focused almost solely on the supply side of information. Using the Northeast Blackout event as an exogenous shock to investors' information acquisition capacity, I show that the demand side of information also plays an important role.

Second, information acquisition has a larger effect through the information asymmetry channel for small firms, which typically have lower levels of institutional holdings, analyst coverage, and media exposure than large firms. Therefore, communication through EDGAR plays a more important role for management teams of small firms to deliver their messages to investors. Investors of small firms also rely more heavily on EDGAR to gain insights into the firm's operations. Timely processing the disclosed information reduces the information asymmetry between insiders and investors, leading to a reduction in the cost of capital. Such a reduction is valuable for small firms, as they are the ones facing high financing costs.

Third, my paper has important implications for firms' information dissemination. I find that the effect of acquiring unanticipated information is higher when the cost of information acquisition is lower. The cost of information acquisition is not just about collecting information, but also about processing and interpreting that information. Local investors and recurring investors have such information advantages, and their information acquisition has a higher effect on stock returns than non-local and inexperienced investors. Although firms cannot choose the composition of their investors, they have control over how to efficiently disclose the information and lower the cost of information acquisition by their investors.

References

Andrei, Daniel, and Michael Hasler, 2015, Investor attention and stock market volatility, *The Review of Financial Studies* 28, 33–72.

Barber, Brad M, and Terrance Odean, 2007, All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors, *The Review of Financial Studies* 21, 785–818.

Barlevy, Gadi, and Pietro Veronesi, 2007, Information acquisition in financial markets: a correction, Federal Reserve Bank of Chicago.

Ben-Rephael, Azi, Bruce I Carlin, Zhi Da, and Ryan D Israelsen, 2021, Information consumption and asset pricing, *The Journal of Finance* 76, 357–394.

Ben-Rephael, Azi, Zhi Da, Peter D Easton, and Ryan D Israelsen, 2022, Who pays attention to SEC Form 8-K? *The Accounting Review* 97, 59–88.

Boot, Arnoud WA, and Anjan V Thakor, 2001, The many faces of information disclosure, *The Review of Financial Studies* 14, 1021–1057.

Cao, Zhen, Osman Kilic, and Xuewu Wang, 2020, Investor attention, divergence of opinions, and stock returns, *Journal of Behavioral Finance* 1–15.

Chen, Huaizhi, Lauren Cohen, Umit Gurun, Dong Lou, and Christopher Malloy, 2020, Iq from ip: Simplifying search in portfolio choice, *Journal of Financial Economics* 138, 118–137.

Chen, Joseph, Harrison Hong, and Jeremy C Stein, 2002, Breadth of ownership and stock returns, *Journal of Financial Economics* 66, 171–205.

Da, Zhi, Joseph Engelberg, and Pengjie Gao, 2011, In search of attention, *The Journal of Finance* 66, 1461–1499.

Deaves, Richard, Catherine Dine, and William Horton, 2006, How are investment decisions made, *Task Force to Modernize Securities Legislation in Canada* .

Drake, Michael S, Darren T Roulstone, and Jacob R Thornock, 2016, The usefulness of historical accounting reports, *Journal of Accounting and Economics* 61, 448–464.

Easley, David, Soeren Hvidkjaer, and Maureen O'Hara, 2002, Is information risk a determinant of asset returns? *The Journal of Finance* 57, 2185–2221.

Easley, David, and Maureen O'Hara, 2004, Information and the cost of capital, *The Journal of Finance* 59, 1553–1583.

Gargano, Antonio, and Alberto G Rossi, 2018, Does it pay to pay attention? *The Review of Financial Studies* 31, 4595–4649.

Gervais, Simon, Ron Kaniel, and Dan H Mingelgrin, 2001, The high-volume return premium, *The Journal of Finance* 56, 877–919.

Gibbons, Brian, Peter Iliev, and Jonathan Kalodimos, 2021, Analyst information acquisition via EDGAR, *Management Science* 67, 769–793.

Grossman, Sanford J, and Joseph E Stiglitz, 1980, On the impossibility of informationally efficient markets, *The American Economic Review* 70, 393–408.

Holden, Craig W, and Stacey Jacobsen, 2014, Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions, *The Journal of Finance* 69, 1747–1785.

Huberman, Gur, and Tomer Regev, 2001, Contagious speculation and a cure for cancer: A nonevent that made stock prices soar, *The Journal of Finance* 56, 387–396.

Iliev, Peter, Jonathan Kalodimos, and Michelle Lowry, 2021, Investors' attention to corporate governance, *The Review of Financial Studies* 34, 5581–5628.

Kelly, Bryan, and Alexander Ljungqvist, 2012, Testing asymmetric-information asset pricing models, *The Review of Financial Studies* 25, 1366–1413.

Lee, Charles MC, Paul Ma, and Charles CY Wang, 2015, Search-based peer firms: Aggregating investor perceptions through internet co-searches, *Journal of Financial Economics* 116, 410–431.

Li, Frank Weikai, and Chengzhu Sun, 2022, Information acquisition and expected returns: Evidence from EDGAR search traffic, *Journal of Economic Dynamics and Control* 141, 104384.

Mele, Antonio, and Francesco Sangiorgi, 2015, Uncertainty, information acquisition, and price swings in asset markets, *The Review of Economic Studies* 82, 1533–1567.

Veldkamp, Laura L., 2006, Information Markets and the Comovement of Asset Prices, *The Review of Economic Studies* 73, 823–845.

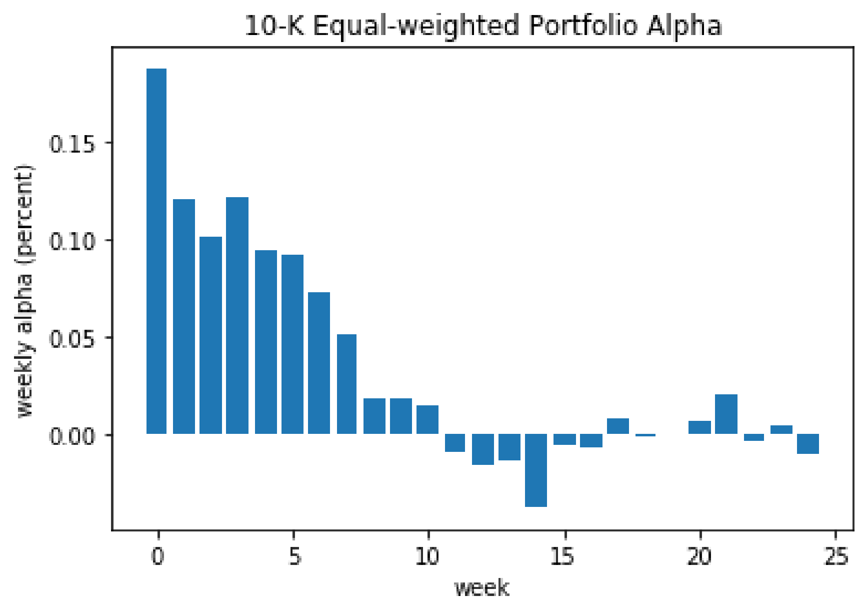
Verrecchia, Robert E, 1982, Information acquisition in a noisy rational expectations economy, *Econometrica: Journal of the Econometric Society* 1415–1430.

Wang, Jiang, 1993, A model of intertemporal asset prices under asymmetric information, *The Review of Economic Studies* 60, 249–282.

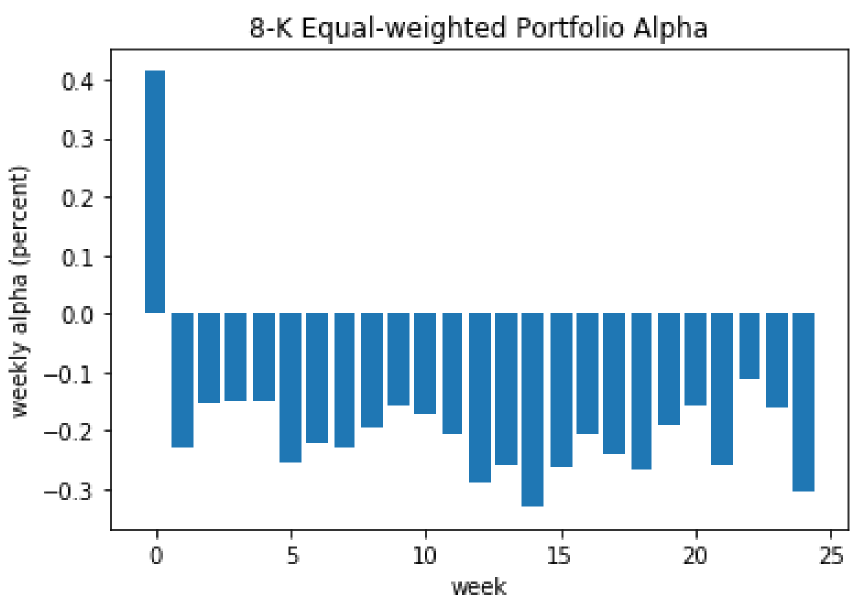
Figure 1

Long/Short 10-K (8-K) Demand Portfolio - Weekly Returns

The figure shows the weekly Fama French 5-factor alphas of 10-K (Panel a) and 8-K (Panel b) portfolios. Stocks are sorted by the weekly downloads at the end of Friday within the NYSE size-group. Long/short portfolios are held throughout the next 24 weeks. For 8-K portfolios, I limit the set of stocks that issue 8-K filings in the week, and focus on the number of newly issued 8-K downloads.



(a)



(b)

Figure 2
Newly Disclosed 10-K Portfolios
The figure shows the weekly Fama French 5-factor alphas of 10-K portfolios, conditional on a set of firms just disclosed 10-K in a week. At each week, I limit the sample to firms just disclosed 10-K in the week. Stocks are then sorted within the NYSE size-groups by the downloads of the newly disclosed 10-K filing into quintiles. Long/short portfolios are held throughout the next 24 weeks.

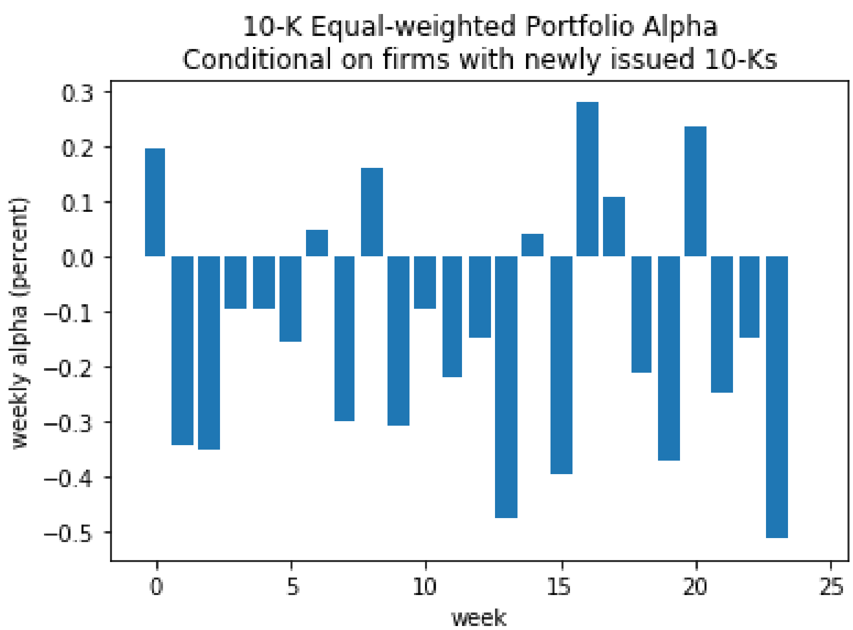
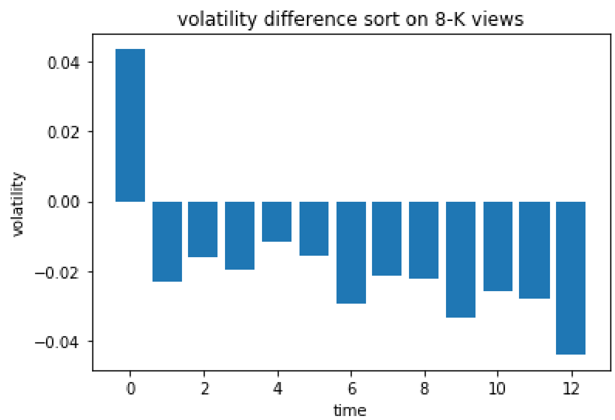
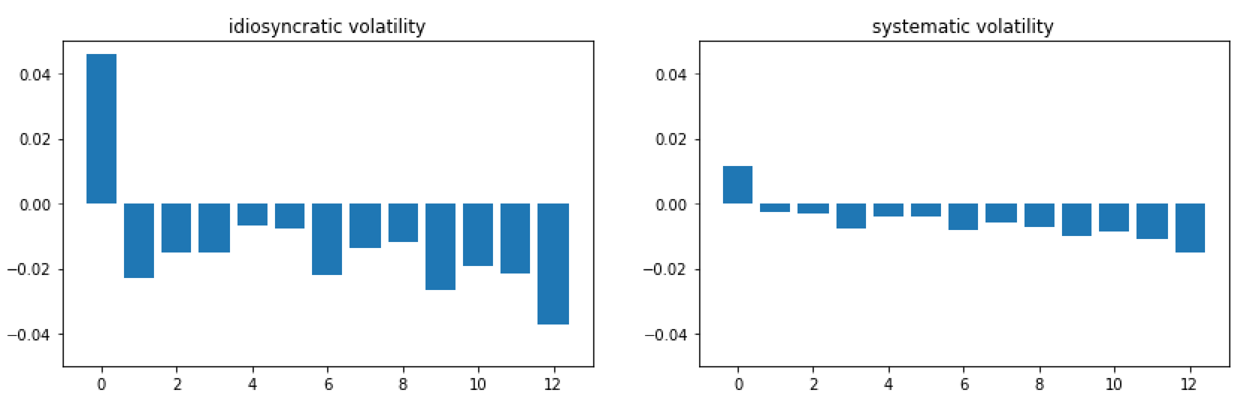


Figure 3
Changes in Volatility
This figure shows the changes in volatility following 8-K downloads. For each stock in each month, I calculate the volatility of daily returns within the month. I also decompose the volatility into systematic volatility and idiosyncratic volatility by using the Fama-French five factors and momentum factor loadings estimated with a three-year rolling window. For a stock with unscheduled 8-K filings in a given month, I then calculate the abnormal volatility by subtracting the volatility from its one-year volatility average. I then sort stocks by size-adjusted 8-K downloads into quintiles and plot the difference in abnormal volatilities between the top and the bottom quintiles. Panel (a) shows the difference in abnormal volatility, and Panel (b) shows the result with the decomposition.



(a)

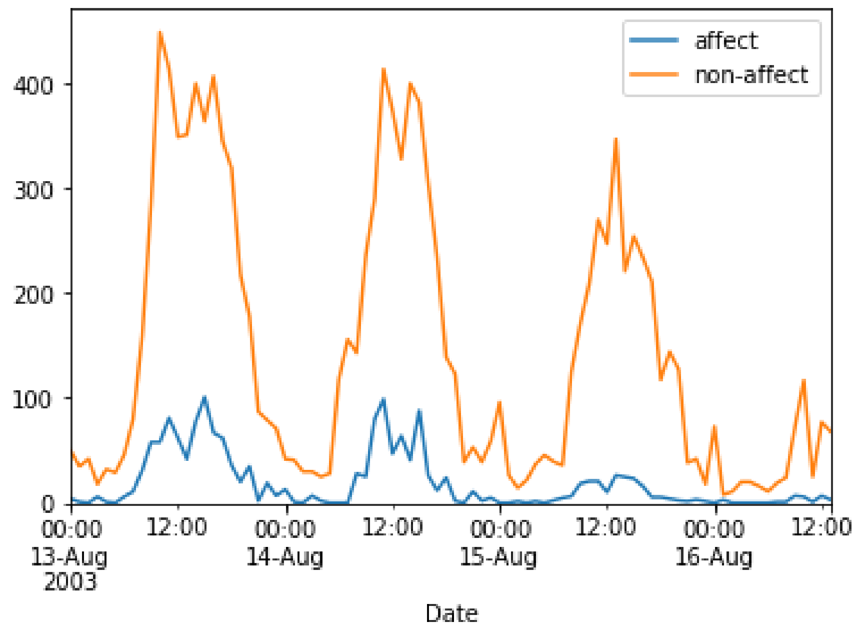


(b)

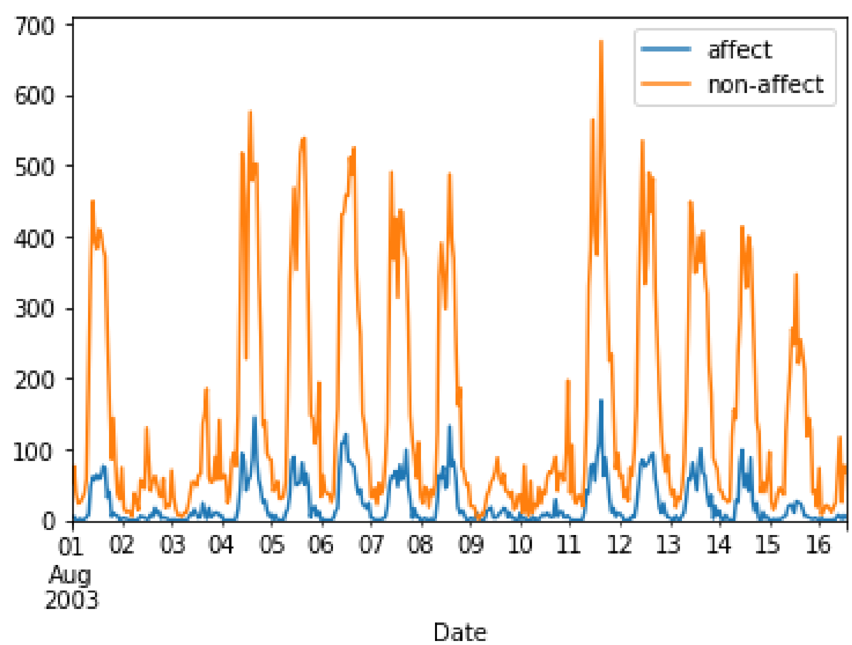
Figure 4

Viewing Activity during Northeast Blackout of 2003

The figure shows the hourly 8-K download pattern by investors in affected and non-affected regions during the blackout event. Panel (a) shows the downloading activity between August 13 and August 16, 2003. Panel (b) shows the downloading activity between August 1 and August 16, 2003.



(a) Downloading Activity between August 13 and August 16, 2003



(b) Downloading Activity between August 1 and August 16, 2003

Figure 5

Histogram of Cumulative Abnormal Return around Events

The figure shows the histogram of the cumulative abnormal return around the 8-K event date. For each unscheduled 8-K filings, I calculate the 3-day cumulative abnormal return relative to the market around the event.

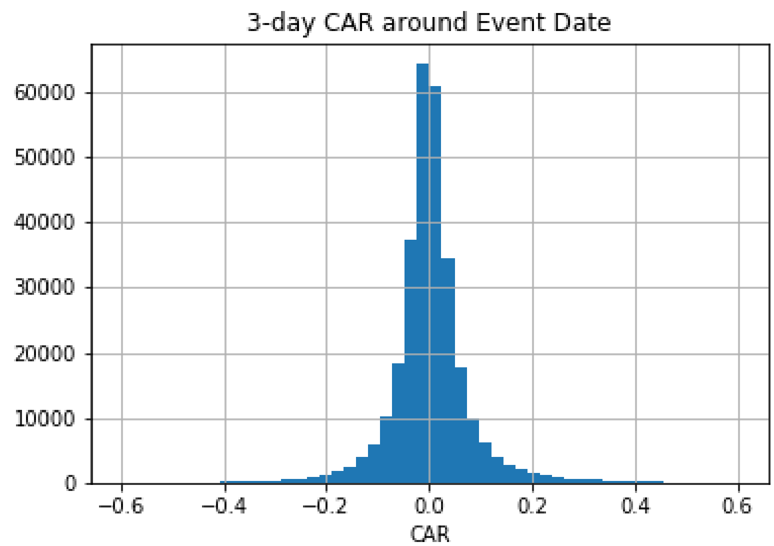


Figure 6

Demand for 8-K and Abnormal Return around Events

The figure studies the long/short portfolio of 8-K demand and abnormal returns around 8-K filing and event date. For each unscheduled 8-K filings, I calculate the cumulative abnormal return relative to the market around event and filing date. I then double sort stocks by the 8-K demand within the NYSE size-groups and the cumulative abnormal return into 5-by-5 blocks. Conditional on each abnormal return quintile, I regress the long/short 8-K portfolio return on the Fama-French five-factor and the UMD factor, and plot the alphas and 95% confidence intervals. For stocks with multiple unscheduled filings in a month, I choose the one with the highest absolute abnormal return.

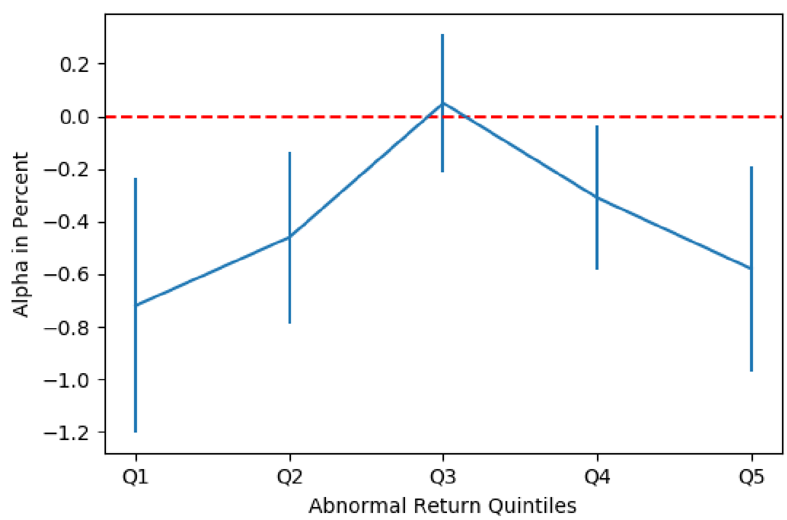


Table 1

Fama-Macbeth Regression on EDGAR Demand for Filings

The table shows results from Fama-Macbeth regressions of monthly individual stock returns on EDGAR downloads. The variable *log views*$_k$ is the natural logarithm of human downloads of the firm for filing type k. For unscheduled 8-K filings, I also split the downloading activity by whether the filing was disclosed within a week. Regressions include controls for other variables that are known to predict cross-section variation in returns. Independent variables are winsorized at 1% and 99% levels. The sample covers from 2003 to 2016, with the dates determined by the availability of EDGAR Log data. Asset Growth is the annual percentage change in total assets. log(BM) is the natural logarithm of the book-to-market ratio. log(ME) is the natural logarithm of market capitalization. Operating Profit is the revenue minus cost of goods sold, SG&A expenses, and interest expense, divided by lagged common shareholders' equity. Abnormal Trading Volume is the difference between trading volume and previous 12-month average trading volume, scaled by the standard deviation of the previous 12-month trading volume. SUE is the unexpected quarterly earnings (adjusted by median forecast earnings) divided by fiscal-quarter-end market capitalization. Earnings Drift is the sum of daily returns in three days around the earnings announcement. Media Coverage is the total number of news covered by Ravenpack. Count Variables file 10K/10Q/8K are the number of 10-K/10-Q/8-K filings in the month. Past 12-month filings are the number of 8-K filings in the past 12 months.

	(1) Ret_{t+1}	(2) Ret_{t+1}	(3) Ret_{t+1}	(4) Ret_{t+1}	(5) Ret_{t+1}	(6) Ret_{t+1}
$log\ views_{all}$		0.199^{**}				
		(2.07)				
$log\ views_{10K}$			0.388^{***}	0.386^{***}	0.385^{***}	0.340^{***}
			(7.43)	(7.35)	(7.45)	(5.82)
$log\ views_{10Q}$			-0.0663	-0.0669	-0.0619	-0.0614
			(-1.14)	(-1.14)	(-1.06)	(-1.10)
$log\ views_{8K}$			-0.115^{**}			
			(-2.30)			
$log\ views_{8K}^{scheduled}$				0.0240	0.00626	0.00173
				(0.73)	(0.19)	(0.04)
$log\ views_{8K}^{unscheduled}$				-0.112^{**}		
				(-2.39)		
$log\ views_{8K}^{new\ unscheduled}$					-0.197^{***}	-0.185^{***}
					(-5.10)	(-4.66)
$log\ views_{8K}^{old\ unscheduled}$					0.0850	0.00830
					(1.03)	(0.18)
file 10K	0.225^{*}	0.139	-0.0641	-0.0680	-0.0579	-0.0624
	(1.90)	(1.15)	(-0.51)	(-0.54)	(-0.46)	(-0.41)
file 8K	-0.0653^{**}	-0.102^{***}	-0.0507^{*}	-0.0509^{*}	-0.00841	-0.00676
	(-2.56)	(-3.65)	(-1.71)	(-1.75)	(-0.28)	(-0.20)
file 10Q	-0.178^{**}	-0.247^{**}	-0.223^{**}	-0.223^{**}	-0.209^{**}	-0.173
	(-1.98)	(-2.46)	(-2.18)	(-2.18)	(-2.06)	(-1.56)
Past 12-month Filings	0.00442	-0.00445	0.000228	0.00129	-0.00134	-0.00531
	(0.48)	(-0.64)	(0.03)	(0.19)	(-0.21)	(-0.89)
Asset Growth	-0.724^{***}	-0.680^{***}	-0.628^{***}	-0.630^{***}	-0.639^{***}	-0.541^{***}
	(-4.74)	(-4.67)	(-4.31)	(-4.32)	(-4.39)	(-3.64)
log(BM)	0.120	0.105	0.0921	0.0913	0.0933	0.0538
	(0.79)	(0.69)	(0.61)	(0.60)	(0.62)	(0.35)
log(ME)	-0.0950	-0.143	-0.183^{**}	-0.184^{**}	-0.192^{**}	-0.0851
	(-1.23)	(-1.54)	(-2.01)	(-2.02)	(-2.10)	(-0.93)
Operating Profit	0.0808^{**}	0.0625^{*}	0.0467	0.0463	0.0475	0.0465
	(2.23)	(1.68)	(1.27)	(1.26)	(1.29)	(1.35)
$r_{1,0}$	-2.357^{***}	-2.410^{***}	-2.414^{***}	-2.409^{***}	-2.414^{***}	-2.232^{***}
	(-3.59)	(-3.75)	(-3.76)	(-3.75)	(-3.77)	(-3.10)
$r_{12,2}$	-0.564	-0.475	-0.473	-0.471	-0.466	-0.377
	(-1.39)	(-1.28)	(-1.27)	(-1.27)	(-1.26)	(-0.97)
Abnormal Trading Volume	0.138^{***}	0.129^{***}	0.134^{***}	0.134^{***}	0.138^{***}	0.129^{***}
	(4.00)	(3.80)	(3.98)	(3.98)	(4.10)	(3.37)
SUE	3.933^{***}	3.909^{***}	3.874^{***}	3.873^{***}	3.883^{***}	3.990^{***}
	(4.93)	(4.94)	(4.90)	(4.90)	(4.91)	(4.31)
Earnings Drift	1.241^{***}	1.250^{***}	1.230^{***}	1.225^{***}	1.212^{***}	1.095^{***}
	(3.28)	(3.35)	(3.30)	(3.29)	(3.26)	(2.68)
No. of Analysts	0.0141	0.00718	0.00583	0.00588	0.00610	0.00880
	(1.07)	(0.62)	(0.50)	(0.51)	(0.53)	(0.71)
Change in Google Trend						-0.110
						(-0.83)
Media Coverage						0.00364
						(0.83)
Constant	2.046^{*}	1.893^{*}	2.451^{**}	2.450^{**}	2.485^{**}	1.184
	(1.87)	(1.79)	(2.21)	(2.20)	(2.19)	(1.05)
N	502662	502662	502662	502662	502662	347381
Average R^2	0.0384	0.0410	0.0425	0.0427	0.0433	0.0464
F	10.04	10.06	12.74	11.96	12.85	7.568

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2

Long/Short Portfolio by 10-K and 8-K Demand

The table shows monthly alphas and factor loadings of portfolios sorted by the 10-K/8-K viewing activity. To control for firm sizes, I first sort stocks by size into five groups using NYSE breakpoints. Conditional on each NYSE size-group, I then sort stocks by 10-K (8-K) downloads into quintiles and form equal-weighted portfolios. Panel A and B show the long/short portfolio returns and alphas with one, three, and twelve holding months for the 10-K and 8-K portfolios. Panel C and D show the factor loadings of 10-K and 8-K portfolios. For 8-K downloads, I only focus on the downloads on unscheduled filings.

Panel A: Size-adjusted 10-K Views Equal Weighted L/S Alpha

Holding Months	$Alpha^{CAPM}$	$Alpha^{FF3}$	$Alpha^{FFC}$	$Alpha^{FF5+UMD}$	$Alpha^{8-factor}$
1	0.39*	0.39**	0.47***	0.37**	0.46***
	(1.89)	(1.99)	(3.31)	(2.6)	(3.14)
3	0.11	0.11	0.19	0.07	0.16
	(0.57)	(0.62)	(1.48)	(0.55)	(1.22)
12	-0.03	-0.05	0.01	-0.12	-0.02
	(-0.22)	(-0.33)	(0.12)	(-1.03)	(-0.21)

Panel B: Size-adjusted 8-K Views Equal Weighted L/S Alpha

Holding Months	$Alpha^{CAPM}$	$Alpha^{FF3}$	$Alpha^{FFC}$	$Alpha^{FF5+UMD}$	$Alpha^{8-factor}$
1	-0.54***	-0.55***	-0.47***	-0.43***	-0.41***
	(-2.82)	(-2.96)	(-3.52)	(-3.11)	(-2.94)
3	-0.57***	-0.58***	-0.49***	-0.45***	-0.45***
	(-3.13)	(-3.29)	(-3.98)	(-3.5)	(-3.41)
12	-0.55***	-0.56***	-0.49***	-0.44***	-0.46***
	(-3.34)	(-3.65)	(-4.17)	(-3.65)	(-3.69)

Panel C: Factor Loadings of 10-K Portfolio

Quintiles	Alpha	Market	SMB	HML	UMD	RMW	CMA
Low	0.08	0.779***	0.525***	0.148***	-0.029	-0.325***	-0.173**
	(0.85)	(28.23)	(11.53)	(3.3)	(-1.34)	(-5.34)	(-2.32)
2	0.07	0.897***	0.697***	0.088**	-0.067***	-0.328***	-0.116*
	(0.9)	(38.33)	(18.05)	(2.3)	(-3.6)	(-6.35)	(-1.85)
3	0.22***	0.994***	0.733***	0.022	-0.129***	-0.292***	-0.001
	(2.71)	(41.34)	(18.49)	(0.57)	(-6.72)	(-5.5)	(-0.02)
4	0.32***	1.018***	0.796***	0.003	-0.212***	-0.204***	0.029
	(3.38)	(36.56)	(17.32)	(0.07)	(-9.54)	(-3.34)	(0.38)
High	0.46***	1.097***	0.887***	0.107*	-0.441***	-0.117	0.106
	(3.42)	(27.91)	(13.68)	(1.68)	(-14.05)	(-1.35)	(1.01)
H-L	0.37**	0.32***	0.366***	-0.042	-0.41***	0.21**	0.277**
	(2.6)	(7.62)	(5.28)	(-0.62)	(-12.25)	(2.27)	(2.45)

Panel D: Factor Loadings of 8-K Portfolio

Quintiles	Alpha	Market	SMB	HML	UMD	RMW	CMA
Low	0.31***	0.8***	0.554***	0.106***	-0.039**	-0.252***	-0.0
	(4.2)	(37.69)	(15.83)	(3.09)	(-2.3)	(-5.38)	(-0.01)
2	0.19***	0.899***	0.656***	0.091***	-0.106***	-0.165***	-0.143**
	(2.67)	(42.27)	(18.71)	(2.63)	(-6.23)	(-3.51)	(-2.5)
3	0.28***	0.954***	0.734***	0.038	-0.159***	-0.234***	-0.051
	(2.96)	(34.91)	(16.29)	(0.85)	(-7.27)	(-3.88)	(-0.69)
4	0.13*	1.019***	0.761***	0.072*	-0.166***	-0.282***	0.01
	(1.87)	(40.01)	(18.1)	(1.74)	(-8.18)	(-5.03)	(0.14)
High	-0.11	1.097***	0.854***	0.056	-0.423***	-0.354***	-0.051
	(-0.77)	(25.98)	(12.25)	(0.82)	(-12.56)	(-3.81)	(-0.45)
H-L	-0.43***	0.3***	0.304***	-0.053	-0.383***	-0.101	-0.053
	(-3.11)	(7.52)	(4.62)	(-0.82)	(-12.02)	(-1.15)	(-0.5)

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3
Weekly Regression of Stock Returns on EDGAR Demand for Filings

The table shows results from regressions of weekly individual stock returns on EDGAR downloads. The dependent variable in columns (1) and (2) is the current week stock returns in basis points. The dependent variable in columns (3) and (4) is the next week's stock returns in basis points. $views_t^k$ is the cumulative downloads of filing type k at week t. *Filing* k_t is a dummy variable, which is equal to one if the firm issued any filings with type k at week t. $News_t$ is a dummy variable, which is equal to one if there is any news coverage of the firm in Ravenpack at week t. *Earnings Release*$_t$ is a dummy variable, which is equal to one if the firm releases its earnings at week t. AIA_t is a dummy variable, which is equal to one if the Bloomberg News Heat daily index has a maximum of 3 or above in week t. $DADSVI_t$ is a dummy variable, which is equal to one if the Google Trends daily index in any day of the week is above its 90 percentile in the past month. Firm controls include the log of firm market capitalization, book-to-market ratio, trading volume scaled by shares outstanding, number of analysts, and institutional ownership ratio. Time fixed effects are included, and standard errors are clustered by week.

	(1) ret_t	(2) ret_{t+1}	(3) ret_t	(4) ret_{t+1}
$\log(views_t^{10K})$	14.29***	11.42***	7.233***	4.526***
	(9.19)	(8.01)	(4.52)	(2.89)
$\log(views_t^{8K})$	-0.308	-1.724	-2.125	-1.503
	(-0.20)	(-1.23)	(-1.55)	(-1.22)
Filing 10K$_t$	-28.02	-27.72	-4.143	-5.806
	(-1.10)	(-1.35)	(-0.16)	(-0.23)
Filing 8K$_t$	-7.618	7.131	-43.74***	9.576
	(-1.21)	(1.63)	(-3.24)	(1.02)
Filing 10K$_t \times \log(views_t^{10K})$	-1.504	6.902	-0.946	0.895
	(-0.22)	(1.14)	(-0.16)	(0.16)
Filing 8K$_t \times \log(views_t^{8K})$	12.38***	-3.841**	19.09***	-4.437*
	(4.88)	(-1.96)	(4.43)	(-1.82)
Media Coverage$_t$	35.03***	5.664***	25.13***	3.729
	(17.11)	(3.27)	(7.59)	(1.14)
Earnings Release$_t$	39.22***	14.10***	1.576	17.93***
	(6.80)	(3.36)	(0.20)	(3.27)
AIA_t			49.66***	1.479
			(11.35)	(0.53)
$DADSVI_t$			21.78***	2.189
			(11.13)	(1.28)
Lag Returns	Yes	Yes	Yes	Yes
Firm Controls	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
N	2172842	2169992	504523	503489
Adjusted R^2	0.118	0.116	0.159	0.178
F	52.72	17.33	24.27	3.177

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4
Panel Regression of Information Asymmetry

The table shows the weekly panel regression of next-month information asymmetry proxy on current month investor demand for filings. The dependent variable is the price impact measure estimated following Holden and Jacobsen (2014). $views_t^k$ is the cumulative downloads of filing type k at week t. $Filing\ k_t$ is a dummy variable, which is equal to one if the firm issued any filings with type k at week t. $News_t$ is a dummy variable, which is equal to one if there is any news coverage of the firm in Ravenpack at week t. $Earnings\ Release_t$ is a dummy variable, which is equal to one if the firm releases its earnings at week t. Firm controls include the log of firm market capitalization, book-to-market ratio, trading volume scaled by shares outstanding, number of analysts, and institutional ownership ratio. Time and firm fixed effects are included. Standard errors are two-way clustered by time and firm.

	(1) $Price\ Impact_{t+1}$	(2) $Price\ Impact_{t+1}$
$\log(views_t^{10K})$	-0.00144*	-0.00126***
	(-1.89)	(-1.83)
$\log(views_t^{8K})$	-0.00359**	-0.00338**
	(-2.48)	(-2.31)
$Filing\ 10K_t$	0.00679	0.0389**
	(0.68)	(2.09)
$Filing\ 8K_t$	-0.00277*	0.00174
	(-1.69)	(0.53)
$Filing\ 10K_t \times \log(views_t^{10K})$		-0.00357*
		(-1.85)
$Filing\ 8K_t \times \log(views_t^{8K})$		-0.00585**
		(-2.11)
$Media\ Coverage_t$	0.00702***	0.00696***
	(2.90)	(2.88)
$Earnings\ Release_t$	-0.0102**	-0.0103**
	(-2.16)	(-2.20)
Firm and Time FEs	Yes	Yes
Firm Controls	Yes	Yes
N	1902133	1902133
Adjusted R^2	0.491	0.491
F	218.8	172.1

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5
8-K Demand and Information Asymmetry
The table shows monthly alphas of portfolios sorted by 8-K downloads and information asymmetry. I use Amihud illiquidity measure and previous quarter earnings forecast dispersion to measure ex-ante information asymmetry. For each portfolio, I regress the portfolio returns on the Fama-French five-factor and the UMD factor. For 8-K downloads, I only focus on the downloads on unscheduled filings.

Panel A: Double Sort by Size-adjusted 8-K Views and Amihud

Amihud/Views	Low	2	3	4	High	H-L
Low	0.12	0.06	0.01	0.07	0.02	-0.11
	(1.58)	(0.84)	(0.21)	(0.82)	(0.17)	(-0.93)
Medium	0.20**	0.13	0.26**	0.13	-0.12	-0.33*
	(2.09)	(1.57)	(2.4)	(1.22)	(-0.74)	(-1.67)
High	0.45***	0.48**	0.54**	0.57**	-0.10	-0.53**
	(2.75)	(2.55)	(2.14)	(2.32)	(-0.27)	(-2.27)

Panel B: Double Sort by Size-adjusted 8-K Views and Past Forecast Dispersion

Forecast Dispersion/Views	Low	2	3	4	High	H-L
Low	0.28***	0.12	0.15*	0.20**	0.33***	0.05
	(3.56)	(1.58)	(1.82)	(2.13)	(2.7)	(0.34)
Medium	0.17**	-0.03	0.26***	0.13	-0.05	-0.22
	(2.09)	(-0.37)	(2.68)	(1.29)	(-0.34)	(-1.34)
High	0.17	0.02	-0.05	-0.15	-0.38**	-0.56***
	(1.19)	(0.13)	(-0.33)	(-1.16)	(-2.03)	(-2.85)

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6

Difference-in-differences Estimation on Blackout Event

The table shows the estimation result of diff-in-diff regressions. The dependent variable is the daily price impact measure to proxy for firm-level information asymmetry. The sample includes two days of data, from August 14 to August 15, 2003. Firms in the sample just disclosed material information on August 13 and have headquarters outside the affected regions. The variable *Frac* is the fraction of historical 8-K download from the affected regions prior to the shock. The variable *Treated* is equal to one if the firms have above-median *Frac*. The dummy variable *Post* is equal to one if it is the second day of the sample. The last two columns further exclude firms with top business activity from affected regions, where the level of business activity is measured by the frequency of state name appeared in 10-K filings.

	(1) Price Impact	(2) Price Impact	(3) Price Impact	(4) Price Impact
Post	-0.0611**	-0.0499	-0.0574*	-0.0445
	(-2.00)	(-1.33)	(-1.83)	(-1.14)
Treated	-0.0395		-0.0212	
	(-1.25)		(-0.63)	
Treated × Post	0.167**		0.169*	
	(2.00)		(1.81)	
Frac		-0.409*		-0.284
		(-1.89)		(-1.21)
Frac × Post		0.855**		0.834**
		(2.34)		(2.02)
Constant	0.234***	0.246***	0.228***	0.239***
	(10.10)	(9.60)	(9.67)	(9.05)
Exclude Plans from Affected Regions	No	No	Yes	Yes
N	640	640	580	580
Adjusted R^2	0.00968	0.00427	0.0112	0.00409
F	3.960	3.995	3.816	3.487

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7
Alternative Explanations on Blackout Event
The table shows the estimation result of diff-in-diff regressions. The dependent variable in columns (1) and (2) is the daily trading volume normalized by the number of shares outstanding. The dependent variable in columns (3) and (4) is the daily number of trades after the market close. The sample includes two days of data, from August 14 to August 15, 2003. Firms in the sample just disclosed material information on August 13. The variable *Frac* is the fraction of historical 8-K download from the affected regions prior to the shock. The variable *Treated* is equal to one if the firms have above-median *Frac*. The dummy variable *Post* is equal to one if it is the second day of the sample.

	(1) Volume	(2) Volume	(3) N_Trades$^{after\ market}$	(4) N_Trades$^{after\ market}$
Post	-0.00395***	-0.00347***	0.101	0.227
	(-3.67)	(-3.03)	(0.14)	(0.29)
Treated	0.000266		1.534*	
	(0.23)		(1.90)	
Treated × Post	-0.000137		-0.167	
	(-0.08)		(-0.15)	
Frac		0.0118		14.55***
		(1.45)		(2.60)
Frac × Post		-0.00764		-2.815
		(-0.66)		(-0.36)
Constant	0.00696***	0.00625***	5.548***	5.180***
	(9.16)	(7.72)	(10.61)	(9.32)
N	580	580	580	580
Adjusted R^2	0.0401	0.0439	0.0112	0.0190
F	8.017	8.817	2.170	3.715

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8
8-K Demand and the Cost of Information Acquisition

The table shows monthly alphas of equal-weighted portfolios sorted by 8-K downloads, conditional on geographical distance distribution to headquarters and recurring viewer ratios. Geographical distance is the value-weighted distance between the location of viewing IP and the firm headquarter. I classify a view as a recurring view if the IP address visited any firm filings in the past three months. Recurring visitor ratio is the ratio between the numbers of recurring and non-recurring downloads. Abnormal institutional attention is the fraction of institutional 8-K downloads normalized by the past 12-month average and standard deviation. For each stock at each month, I first sort stocks by geographical distance (recurring visitor ratio, abnormal institutional attention) into terciles. Conditional on each tercile, I then sort stocks by 8-K downloads within the NYSE size-group into quintiles. For each portfolio, I regress portfolio returns on the Fama-French five-factor and the UMD factor, and report the alphas.

Panel A: Double Sort by Size-adjusted 8-K Views and Distance

Distance/Views	Low	2	3	4	High	H-L
Low	0.38***	0.16*	0.28**	0.19**	-0.21	-0.61***
	(4.67)	(1.73)	(2.61)	(2.02)	(-1.35)	(-3.94)
Medium	0.23**	0.31***	0.22**	0.09	-0.19	-0.41**
	(2.32)	(3.45)	(2.01)	(0.8)	(-1.15)	(-2.43)
High	0.28**	0.12	0.28**	0.33**	0.05	-0.23
	(2.53)	(1.18)	(2.22)	(2.54)	(0.23)	(-1.13)

Panel B: Double Sort by Size-adjusted 8-K Views and 8-K Recurring Ratio

$Freq^{8K}$/Views	Low	2	3	4	High	H-L
Low	0.29***	0.20**	0.27**	0.36***	0.16	-0.15
	(3.0)	(2.0)	(2.38)	(2.74)	(0.95)	(-0.81)
Medium	0.26***	0.26***	0.31***	0.23**	-0.15	-0.41**
	(2.62)	(3.1)	(2.8)	(2.15)	(-0.86)	(-2.25)
High	0.25***	0.11	0.20*	0.05	-0.30	-0.55***
	(2.81)	(1.14)	(1.74)	(0.45)	(-1.62)	(-2.95)

Panel C: Double Sort by Size-adjusted 8-K Views and Abnormal Institution Attention

$Abnormal\ Inst\ Attention^{8K}$/Views	Low	2	3	4	High	H-L
Low	0.35**	0.44***	0.11	0.21	-0.14	-0.49**
	(2.35)	(2.72)	(0.69)	(1.04)	(-0.76)	(-2.51)
Medium	0.37**	0.13	0.22	0.26	-0.1	-0.47*
	(2.07)	(0.81)	(1.47)	(1.37)	(-0.47)	(-1.76)
High	0.57***	0.37**	0.29*	0.06	-0.13	-0.7***
	(3.82)	(2.15)	(1.7)	(0.31)	(-0.65)	(-3.01)

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 9
8-K Demand and Information Content

The table shows monthly alphas of portfolios sorted by 8-K downloads and cumulative abnormal returns around filing and event date of unscheduled 8-K filings. For each unscheduled 8-K filing, I calculate the cumulative abnormal return relative to the market around event and filing dates and select the one with the highest absolute abnormal return. I also calculate a context-based abnormal return for each filing by subtracting the weighted abnormal returns of all filings across firms in the past year, weighted by the text similarity with the filing. I then double sort stocks by the 8-K downloads within each NYSE size-group and abnormal returns into 5-by-3 blocks. Conditional on each abnormal return tercile, I regress the long/short 8-K portfolio returns on the Fama-French five-factor and the UMD factor. For stocks with multiple unscheduled filings in a month, I choose the one with the highest absolute abnormal return.

Panel A: Double Sort by Size-adjusted 8-K Views and Abnormal Returns

Abnormal Returns/Views	Low	2	3	4	High	H-L
Low	0.42*	0.08	0.33**	-0.05	-0.31	-0.75**
	(1.7)	(0.49)	(2.01)	(-0.29)	(-1.59)	(-2.51)
Medium	0.30**	0.24**	0.09	0.19*	0.16	-0.14
	(2.19)	(2.21)	(0.8)	(1.87)	(1.23)	(-0.72)
High	0.33*	0.45***	0.50***	0.34**	-0.14	-0.48**
	(1.71)	(2.72)	(2.87)	(2.35)	(-0.74)	(-1.99)

Panel B: Double Sort by Size-adjusted 8-K Views and Context-based Abnormal Returns

Context-based Abnormal Returns/Views	Low	2	3	4	High	H-L
Low	0.42*	0.13	0.33**	-0.02	-0.33*	-0.77**
	(1.68)	(0.77)	(2.0)	(-0.11)	(-1.68)	(-2.56)
Medium	0.30**	0.17	0.11	0.17	0.25*	-0.05
	(2.22)	(1.57)	(1.01)	(1.63)	(1.75)	(-0.27)
High	0.32*	0.49***	0.48***	0.33**	-0.18	-0.51*
	(1.69)	(3.05)	(2.79)	(2.28)	(-0.94)	(-1.91)

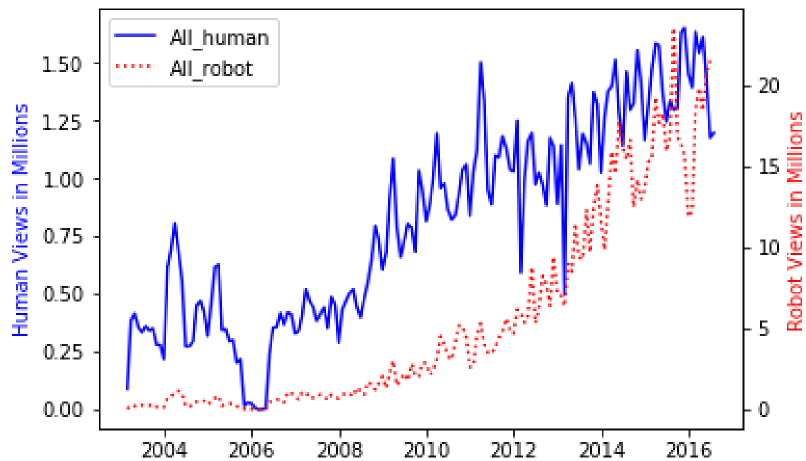
t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

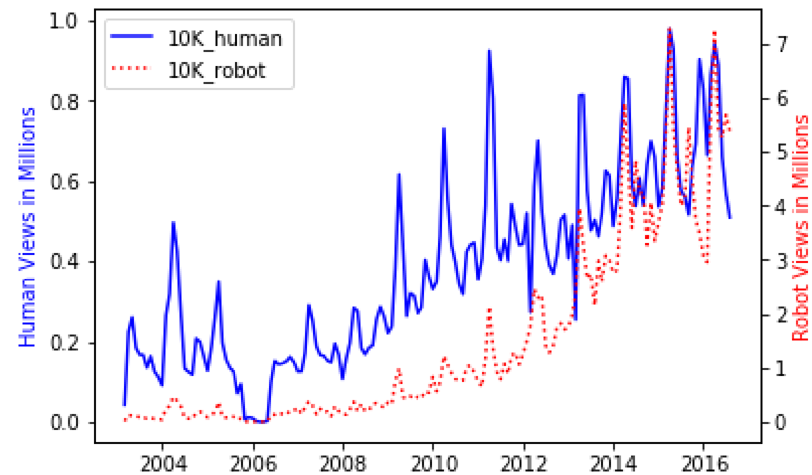
Figure A1

Time-series EDGAR Viewing Activity

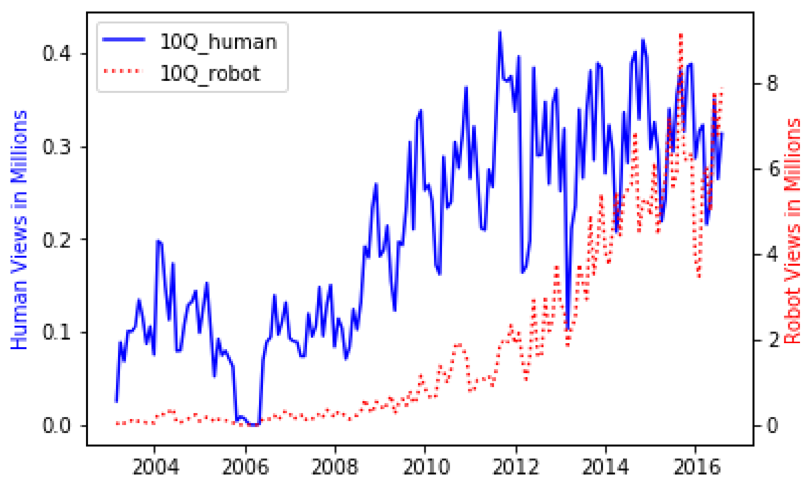
The figure shows the monthly aggregated number of downloads on the EDGAR Log system. Following Lee et al. (2015), I separate crawling activities ("robot") from human viewing activities ("human"). Figures (b) to (d) show the number of downloads for 10-K, 10-Q, and 8-K filings, respectively.



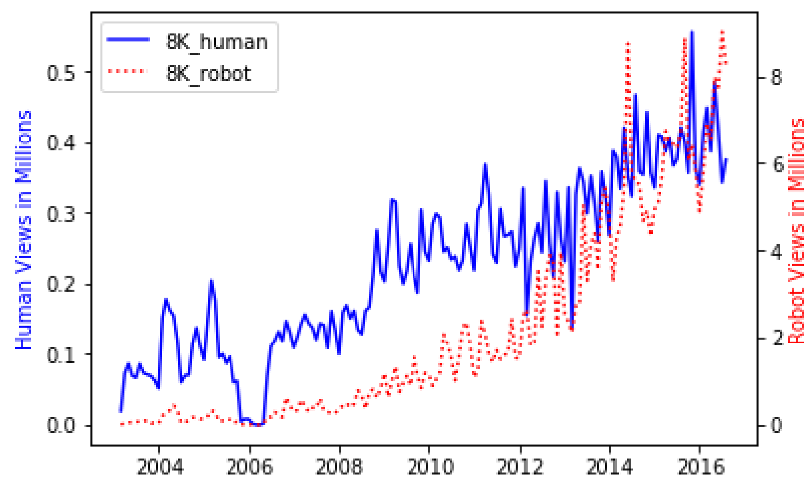
(a) All Filings



(b) 10-K Filings



(c) 10-Q Filings



(d) 8-K Filings

Figure A2

Investors' Demand for Filings Histogram by Firm Sizes

The figure shows the histogram of investor demand for filings on EDGAR, grouped by firm sizes. The horizontal axis is the natural logarithm of monthly filing downloads of a firm. A small firm is defined with a firm market cap below 20% NYSE percentile. A large firm is defined with a firm market cap above 80% NYSE percentile. A medium-sized firm is defined with a firm market cap between 20% and 80% NYSE percentile.

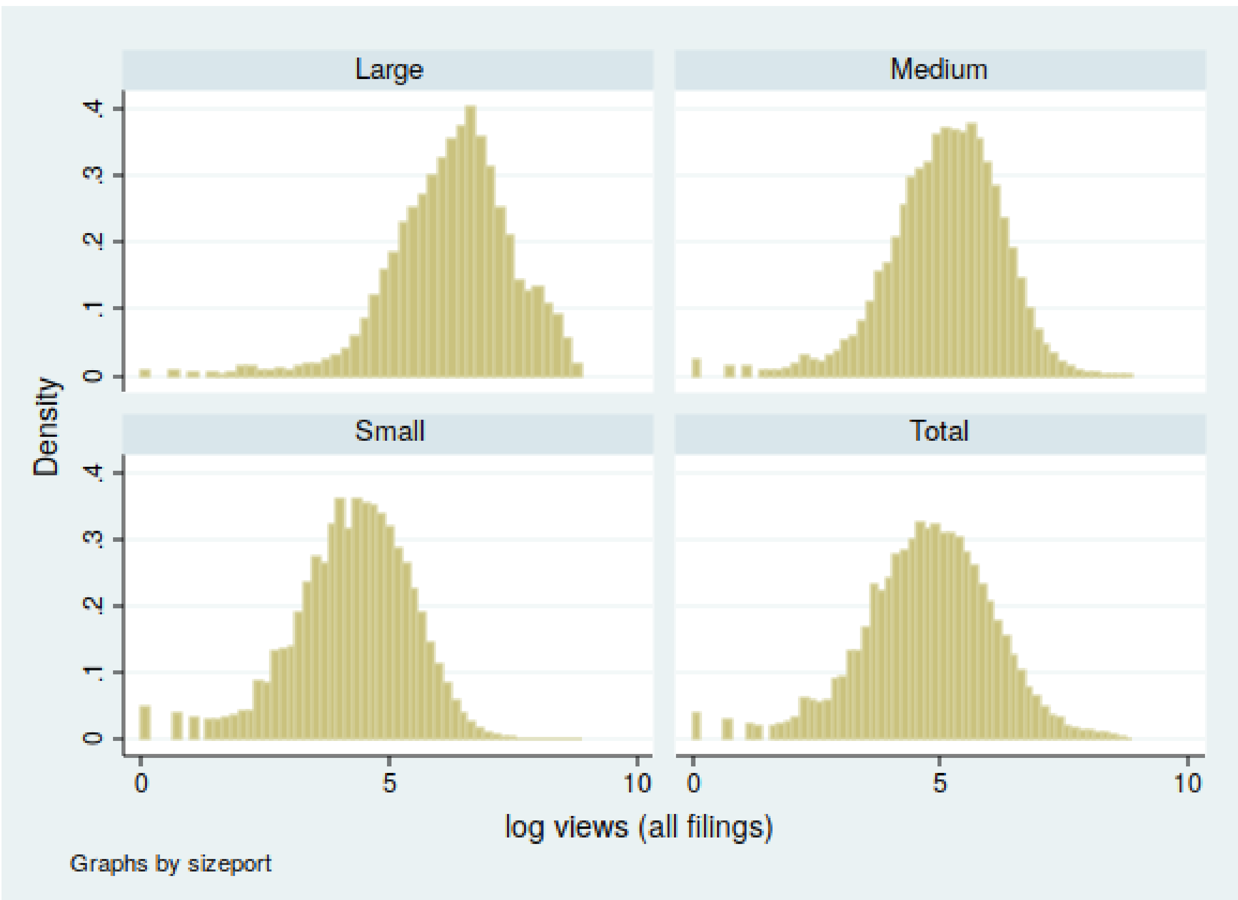
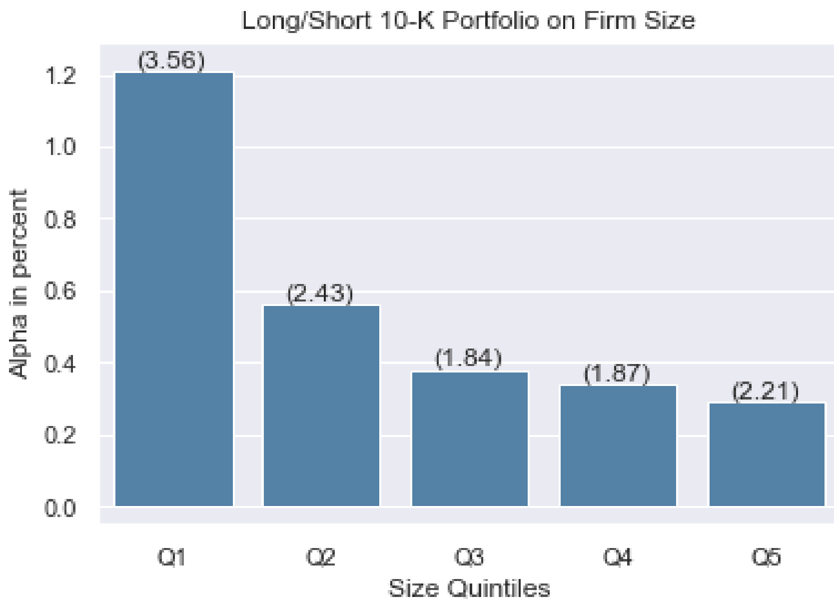


Figure A3

10-K (8-K) Portfolio Alpha and Firm Size

The figure shows the monthly 10-K (8-K) portfolio alpha, conditional on size quintiles. Stocks are sorted by the 10-K (8-K) downloads and the lag firm size into quintiles. Conditional on each size quintile, I form long/short portfolios and regress portfolio return on the Fama-French five-factor and the UMD factor. I then plot the average alpha of long/short portfolio for each size quintile, with t-statistics in parenthesis.



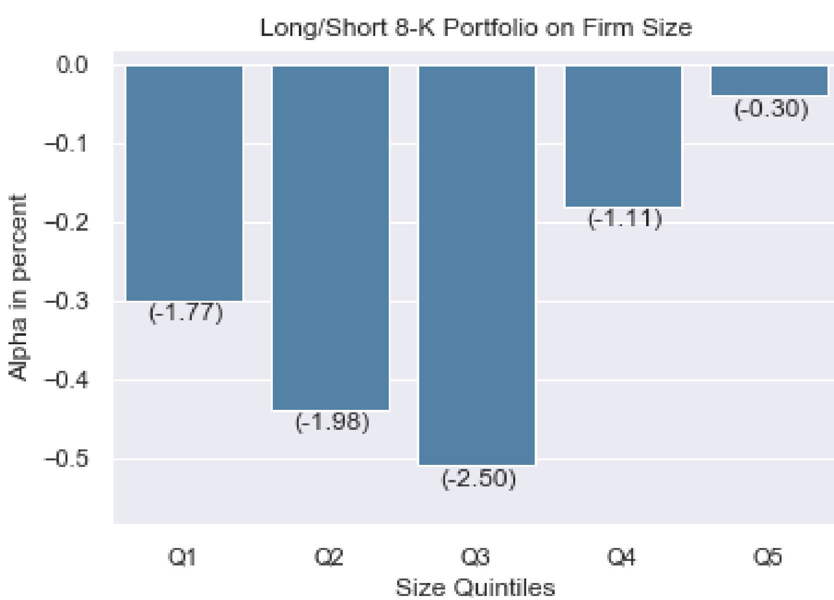


Figure A4

10-K downloads Conditional on 8-K downloads

The figure shows the time-series of 10-K viewing activity, conditional on whether the visitor also viewed any 8-K filings of the firm in the past three months. $views^{only}_{10K}$ is the total number of 10-K downloads by visitors who have not downloaded any 8-K filings of the firm. $views^{both}_{10K}$ is the total number of 10-K downloads by visitors who have downloaded one or more 8-K filings of the firm.

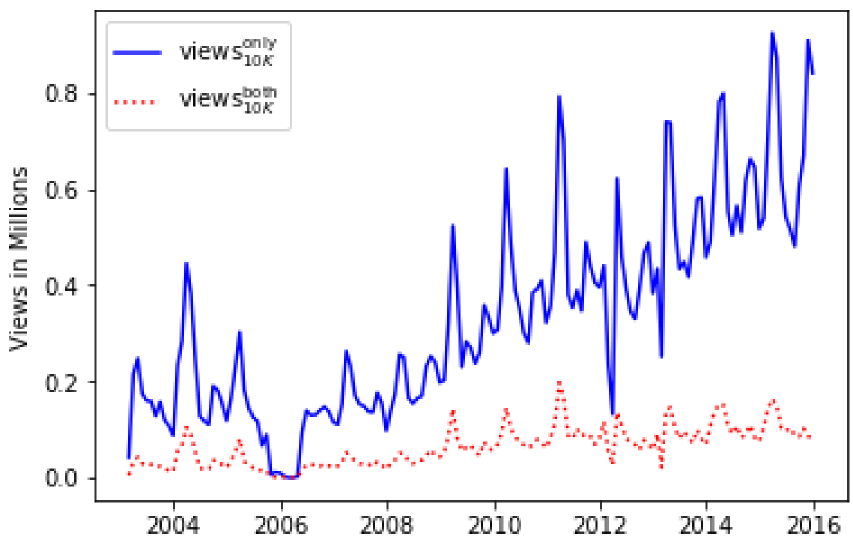
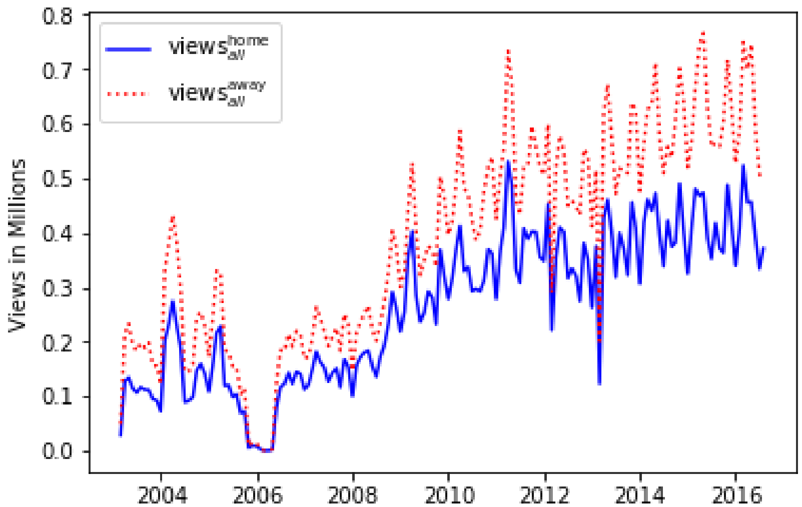


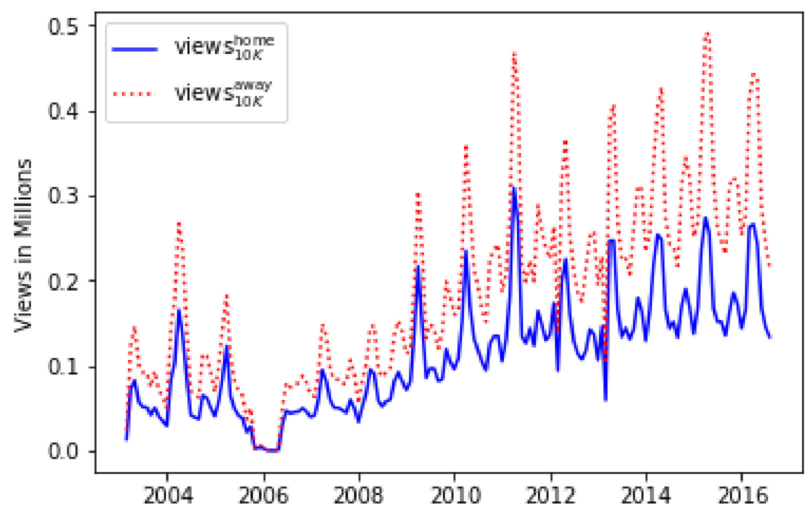
Figure A5

Viewing Activities by Geographical Distance

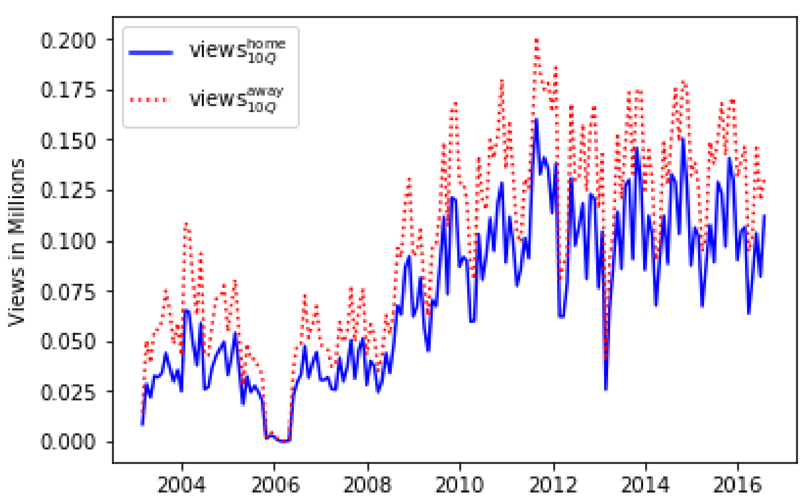
The figure shows the number of downloads by geographical distance. I classify a filing view as home if the distance between the locations of viewing IP and headquarter is less than 400 miles.



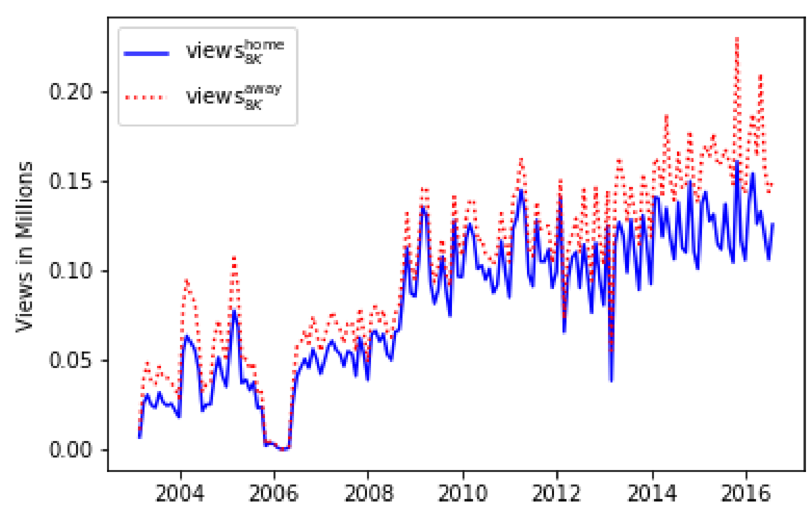
(a) All Filings



(b) 10-K Filings



(c) 10-Q Filings



(d) 8-K Filings

Figure A6

Time-series Recurring Visitor Ratios

The figure shows the time-series plot of recurring visitor ratios by 10-K and 8-K visitors. For each firm and IP address, I classify a filing view as recurring if the IP address submitted requests to view the company filings during the past three months. At each month, I then calculate the cross-section average of recurring ratios by 10-K and 8-K filings.

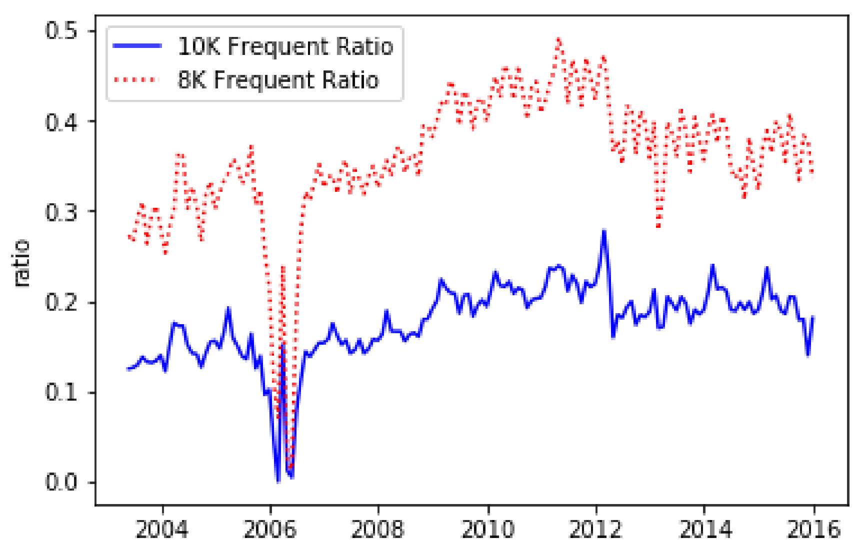
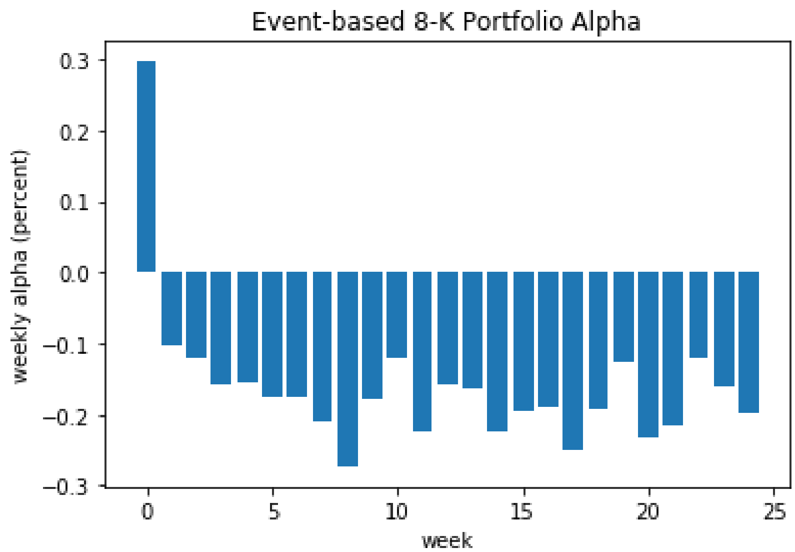


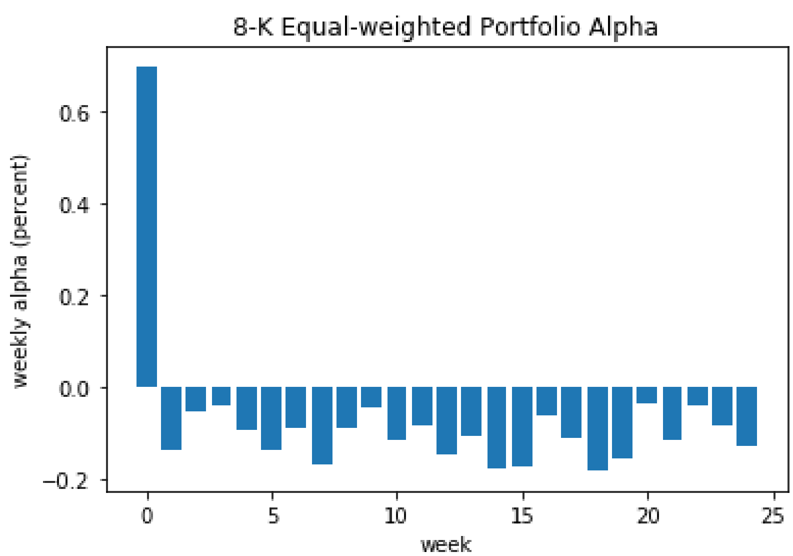
Figure A7

Long-term Performance of 8-K portfolio: Alternative Estimations

The figure shows the long-term alphas of the long/short portfolio sorted by 8-K downloads. In Panel a, at each given date and within each NYSE size quintile, I sort stocks with unscheduled 8-K disclosure by the newly issued filing downloads into quintiles. I then estimate the stock-level factor loadings using one-year daily return before the disclosure and calculate the estimated alpha for the next 25 weeks. I then plot the average difference in alpha between the top quintile stocks and the bottom quintile stocks. In Panel b, stocks with unscheduled 8-K disclosure are sorted by residual views. Residual views are extracted by running a cross-sectional regression of the natural log of unscheduled 8-K views one week after filing on a set of control variables and their square terms. Control variables include the natural log of firm size, the natural log of the number of analysts, abnormal trading volume, and idiosyncratic volatility in the past 12 months. All control variables are lagged for one month.



(a) Size-adjusted 8-K Portfolio



(b) Residual 8-K Portfolio

Table A1

Summary statistics

The table shows the summary statistics of main variables at the firm-month level. $views_{10K}$ is the number of 10-K filing downloads. $views_{10Q}$ is the number of 10-Q filing downloads. $views_{8K}$ is the number of 8-K filing downloads. Asset Growth is the annual percentage change in total assets. log(BM) is the natural logarithm of book-to-market ratio. log(ME) is the natural logarithm of market capitalization. Operating Profit is the revenue minus cost of goods sold, SG&A expenses, and interest expense, divided by lagged common shareholders' equity. Abnormal Trading Volume is the difference between trading volume and previous 12-month average trading volume, scaled by the standard deviation of previous 12-month trading volume. SUE is the unexpected quarterly earnings (adjusted by median forecast earnings) divided by fiscal-quarter-end market capitalization. Earning Drift is the sum of daily returns in three days around earnings announcement. Media Coverage is the total number of news in covered by Ravenpack. file 10K/10Q/8K is the number of 10-K/10-Q/8-K filings in the month.

Variable	Obs	Mean	Std. Dev.	Min	Max	P1	P25	P50	P75	P99
$views_{10K}$	502662	133.98	1032.961	0	370231	0	17	44	110	1478
$views_{10Q}$	502662	79.868	2547.154	0	1053239	0	13	32	75	587
$views_{8K}$	502662	75.786	336.343	0	133132	0	11	31	80	655
Asset Growth	502662	.103	.348	-.679	3.197	-.471	-.038	.047	.154	1.748
log(BM)	502662	.642	.622	-1.611	7.644	-.385	.29	.518	.829	3.055
log(ME)	502662	12.979	2.092	5.535	18.626	8.603	11.439	12.908	14.404	17.85
Operating Profit	502662	.694	1.182	-6.469	9.753	-3.027	.285	.537	.925	6.16
Abnormal Trading Volume	502662	.185	1.612	-2.826	19.255	-1.926	-.76	-.22	.647	6.916
SUE	502662	-.006	.16	-6.275	1.528	-.358	-.003	0	.003	.286
Earning Drift	502662	.002	.088	-.464	.524	-.24	-.04	.001	.042	.255
Media Coverage	397780	8.306	9.512	0	407	0	2	6	11	43
file 10K	502662	.089	.319	0	1	0	0	0	0	1
file 8K	502662	1.008	1.147	0	26	0	0	1	2	5
file 10Q	502662	.253	.478	0	1	0	0	0	0	1

Table A2

Univariate Sort on Residual 8-K Views

The table shows the univariate sort results for residual 8-K views. In each month and conditional on firms with unscheduled 8-K filings, residual 8-K views are extracted by running a cross-sectional regression of the natural log of unscheduled 8-K views one month after filing on a set of control variables and their square terms. Control variables include the natural log of firm size, the natural log of the number of analysts, abnormal trading volume, and idiosyncratic volatility in the past 12 months. All control variables are lagged for one month.

Panel A: Residual 8-K Views Equal-Weighted L/S Alpha

Holding Months	$Alpha^{CAPM}$	$Alpha^{FF3}$	$Alpha^{FFC}$	$Alpha^{FF5+UMD}$	$Alpha^{8-factor}$
1	-0.64***	-0.64***	-0.58***	-0.47***	-0.51***
	(-2.92)	(-3.09)	(-3.37)	(-2.75)	(-2.88)
3	-0.68***	-0.68***	-0.6***	-0.48***	-0.49***
	(-3.06)	(-3.2)	(-3.54)	(-2.83)	(-2.82)
12	-0.64***	-0.66***	-0.58***	-0.48***	-0.51***
	(-3.17)	(-3.4)	(-3.66)	(-3.0)	(-3.1)

Panel B: Factor Loadings of Residual 8-K Portfolio

Residual Views	Alpha	Market	SMB	HML	UMD	RMW	CMA
Low	0.35***	0.809***	0.591***	0.042	-0.09***	-0.173***	-0.015
	(4.22)	(33.88)	(15.0)	(1.08)	(-4.73)	(-3.27)	(-0.24)
2	0.32***	0.904***	0.729***	0.025	-0.14***	-0.155**	-0.126*
	(3.45)	(33.26)	(16.26)	(0.57)	(-6.44)	(-2.59)	(-1.72)
3	0.29***	0.98***	0.766***	0.008	-0.147***	-0.142**	0.054
	(3.3)	(37.89)	(17.97)	(0.18)	(-7.13)	(-2.49)	(0.78)
4	0.27**	1.034***	0.792***	0.06	-0.213***	-0.287***	0.012
	(2.24)	(28.88)	(13.43)	(1.03)	(-7.45)	(-3.63)	(0.12)
High	-0.13	1.077***	0.878***	0.049	-0.44***	-0.486***	0.043
	(-0.71)	(20.52)	(10.14)	(0.57)	(-10.5)	(-4.19)	(0.3)
H-L	-0.47***	0.268***	0.287***	0.007	-0.35***	-0.313***	0.058
	(-2.75)	(5.31)	(3.45)	(0.08)	(-8.69)	(-2.82)	(0.43)

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A3

Univariate Sort on 8-K Crawling Activity

The table shows the univariate sort results for 8-K crawling activity, separated by institutional and retail crawling. To control for firm sizes, I first sort stocks by size into five groups using NYSE breakpoints. Conditional on each NYSE size-group, I then sort stocks by institutional (retail) 8-K crawling into quintiles and form equal-weighted portfolios. Institutions are identified by the name of organizations that own the IP block.

Panel A: Long/Short Institution 8-K Crawling Activity

Institution Crawling	Alpha	Market	SMB	HML	UMD	RMW	CMA
Low	0.40***	0.848***	0.608***	-0.079	-0.002	-0.393***	-0.134
	(3.19)	(23.4)	(10.19)	(-1.34)	(-0.07)	(-4.92)	(-1.37)
2	0.29***	0.874***	0.643***	0.018	-0.112***	-0.232***	-0.091
	(2.74)	(28.22)	(12.6)	(0.35)	(-4.51)	(-3.39)	(-1.09)
3	0.20**	0.948***	0.671***	-0.008	-0.132***	-0.263***	-0.037
	(2.48)	(39.64)	(17.02)	(-0.21)	(-6.92)	(-4.99)	(-0.57)
4	0.13	0.959***	0.75***	0.037	-0.158***	-0.273***	-0.027
	(1.39)	(35.23)	(16.71)	(0.84)	(-7.29)	(-4.53)	(-0.37)
High	-0.10	1.097***	0.808***	0.158**	-0.348***	-0.118	-0.048
	(-0.74)	(28.0)	(12.51)	(2.44)	(-11.1)	(-1.32)	(-0.45)
H-L	-0.48***	0.246***	0.198**	0.247***	-0.344***	0.272**	0.087
	(-2.67)	(4.69)	(2.28)	(2.83)	(-8.19)	(2.27)	(0.6)

Panel B: Long/Short Retail 8-K Crawling Activity

Retail Crawling	Alpha	Market	SMB	HML	UMD	RMW	CMA
Low	0.19	0.845***	0.694***	-0.224***	-0.082***	-0.337***	-0.161*
	(1.55)	(23.74)	(11.82)	(-3.87)	(-2.88)	(-4.29)	(-1.69)
2	0.34***	0.888***	0.632***	-0.053	-0.116***	-0.246***	-0.07
	(3.45)	(30.98)	(13.39)	(-1.13)	(-5.07)	(-3.89)	(-0.91)
3	0.18*	0.95***	0.724***	0.019	-0.14***	-0.17***	-0.047
	(1.91)	(33.84)	(15.62)	(0.42)	(-6.27)	(-2.75)	(-0.62)
4	0.15*	0.979***	0.734***	0.091**	-0.171***	-0.211***	0.058
	(1.68)	(36.42)	(16.53)	(2.08)	(-7.98)	(-3.56)	(0.81)
High	-0.02	1.073***	0.805***	0.148**	-0.339***	-0.261***	-0.097
	(-0.14)	(28.02)	(12.74)	(2.38)	(-11.11)	(-3.09)	(-0.95)
H-L	-0.20	0.226***	0.108	0.374***	-0.258***	0.071	0.066
	(-1.25)	(4.74)	(1.37)	(4.86)	(-6.8)	(0.68)	(0.52)

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$