Demand for Information and Stock Returns: Evidence from EDGAR

Pingle Wang*

June 14, 2019

Abstract

This paper studies the information acquisition process by investors using a novel dataset that tracks filing downloads on the SEC's EDGAR. Demand for 10-K filings predicts short-term positive return spread, and demand for 8-K filings predicts long-term negative return spread. The striking difference in the effects of 10-K and 8-K attention on stock prices can be attributed to the different viewing patterns. Demand for 10-K filings represents a general demand for assets, where 10-K visitors are typically infrequent visitors who never download any filings of the firm in past quarters. The effects are higher for attention-grabbing stocks. In the meanwhile, the demand for 8-K filings comes from frequent and local viewers with a potential information advantage. Such attention reduces information asymmetry of the firm, which decreases the cost of capital and explains the persistent underperformance of high 8-K attention stocks.

*Simon Business School, University of Rochester, Rochester, NY 14627. Phone: (505) 273-1067. E-mail: pingle.wang@simon.rochester.edu.

1 Introduction

Information plays a central role in determining asset prices. There is a large and extensive literature on how the supply side of information affects prices. Recent papers have focused on the economic announcement (Savor and Wilson (2013), Lucca and Moench (2015)), corporate disclosures (Lawrence (2013),Hwang and Kim (2017)), and media coverage (Fang and Peress (2009)). However, how investors acquire and process the information and its effect on asset prices is less well understood.

Information demand affects asset prices in two completely different channels. On the one hand, the demand for information is driven by the demand for assets. Since investors have limited attention (Barber and Odean (2007)), the demand for information reflects their asset selection preference. As a result, the information demand can predict positive future returns. On the other hand, investors who frequently search for and study the information of a firm reduce the proportion of private information and reduce information asymmetry (Grossman and Stiglitz (1980)). Therefore, it becomes less risky for an uninformed investor to hold the asset, the cost of capital reduces (Easley and O'Hara (2004)), prices increase contemporaneously, and the demand for information predicts lower future returns. Current literature points uniformly to the evidence of the asset selection channel (see Da, Engelberg, and Gao (2011) and Ben-Rephael, Carlin, Da, and Israelsen (2017)). However, the empirical evidence on how demand for information reduces information asymmetry is still lacking. It is hard to empirically disentangle the two channels and quantify their effects, as it requires measures on heterogeneous information demand, which are lacked in the current literature.

In this paper, I study how information acquisition by investors on the EDGAR filing system affects stock prices using a large sample of the U.S. public firms from 2003 to 2016. The U.S. Securities and Exchange Commission (SEC) made the EDGAR server log files publicly available recently. Unlike Google/Bloomberg search index, the log data keep track of filing download by each visitor with a unique identifier, so that I can measure not only the level of overall demand for information but also the information content and individual viewing patterns. Among a large set of filing types in EDGAR, I focus on the three most important filings regarding firms' financial and operational conditions, the annual report 10-K, the quarterly report 10-Q, and material information disclosure 8-K[1].

The Form 10-K/Qs and 8-Ks differ in the information content. The Form 10-K/Qs provide investors comprehensive financial and operation statements, which is useful for investors to make investment decisions. Moreover, the Management's Discussion and Analysis provides investors the managers' perspective in evaluating the firm's performance and risk going forward. By the time the reports are released to the public, most of the time-sensitive information is revealed to the market. For example, firms tend to have earnings announcement before releasing the annual/quarterly reports. Forms 8-K are used to notify investors with material information of the firm. The information disclosed was only privately known by insiders and event specific. It relies on investors to collect and process the disclosed information and incorporate it into the market. Therefore, the demand for 10-K/Qs is more likely to capture the asset selection channel, whereas the demand for 8-Ks captures the information asymmetry reduction channel.

Another important distinction between 10-K/Qs and 8-Ks is their reporting frequency. The Form 10-K/Qs are filed with strong seasonality. The Form 8-Ks are filed irregularly, but much more frequent than 10-K/Qs. The average firm files twelve Form 8-Ks, one Form 10-K, and three Form 10-Qs per year. As a result, the attention to 8-K filings is much more likely to respond to the newly disclosed information, which is captured by the information asymmetry reduction channel. The attention to 10-K/Q filings is more likely to respond to the stale information, which is dominated by the asset selection channel.

I show that the aggregated demand for firm filings predicts higher future returns, which is consistent with the general findings in the literature (Barber and Odean (2007); Da et al. (2011); Ben-Rephael et al. (2017)). That is, the asset selection channel is the dominant force of the demand for information. By focusing on the heterogeneous information content, I find

[1]Previous studies on information disclosure also focuses extensively on 10-K/Q and 8-K filings. See Livnat and Zhang (2012) and Gibbons, Iliev, and Kalodimos (2019), for example.

that the demand for 10-K filings predicts higher future returns, the demand for 8-K filings predicts lower future returns, and the demand for 10-Q filings does not predict future returns. The insignificant result for 10-Q filings is due to two reasons. First, the correlation between 10-K and 10-Q views is very high (0.9 in the monthly panel setting). Therefore, 10-Q views do not provide additional variations in explaining return variations beyond 10-Ks. Second, the information in 10-Ks is required to be audited, which is not the case for 10-Qs. Form 10-Ks also provide much detailed information than 10-Qs, especially in the MD&A section.

To further investigate the information acquisition through information asymmetry reduction channel, I decompose the attention to 8-K filings into two parts: the demand for unscheduled 8-K filings, and the demand for scheduled 8-K filings. Scheduled 8-K filings do not disclose any new information to the market[2]. Therefore, the demand for scheduled filings does not reduce the information asymmetry of the firm. Unscheduled 8-K filings, on the contrary, disclose material information, and it requires investors to timely collect and process the information and incorporate it into the market. As a result, demand for unscheduled filings transforms the previously private information into public information, decreases information asymmetry between insiders and investors, and reduces the future cost of capital. I show that the demand for unscheduled 8-K filings predicts lower future returns, and the demand for scheduled 8-K filings does not have any predictability.

I also test the two channels of information acquisition using the standard portfolio sort approach. The long/short portfolio sorted on 10-K views earns around 65 basis points equal-weighted alpha per month, and the portfolio sorted on 8-K views earns -50 basis points equal-weighted alpha[3]. The results are consistent with the Fama Macbeth regression result. Demand for 10-K captures investors' general demand for the asset, thus predicting positive alphas. Demand for 8-K reduces information asymmetry of the firm, thus predicting a lower risk-adjusted return. Besides the difference in return directions the two channels can predict,

[2]Around 10% of the total 8-K filings are scheduled filings.

[3]Throughout the paper, I use Fama-French five factors and UMD factor as the testing model unless specified otherwise.

they also imply drastically different return patterns in the long-term. For the asset selection channel, the predictability of information demand should be short-lived. In a perfectly efficient market, a demand shock will only impact prices contemporaneously. Under certain frictions, the demand shock will be incorporated into prices in a short period, resulting in an alpha decay pattern. The alpha of the 10-K portfolio is consistent with such pattern, starting from an average of 82 basis points (bps) in the formation month and decreasing to around 20 bps in the second holding month. On the contrary, if the demand for information reduces information asymmetry and firms' cost of capital, then the contemporaneous price should go up to reflect the permanent decrease in risk-adjusted returns in the future. Therefore, the portfolio that captures the spread in information demand will have a positive alpha in the formation period, followed by a persistent and negative alpha in the holding periods. The long/short portfolio sorted by 8-K attention exhibits such return pattern. At the formation month, the 8-K portfolio yields a positive alpha of 18 bps. The alpha of the 8-K portfolio then becomes and remains negative (around -60 bps) throughout the next 12 months.

To have better control for the information supply, I also run tests at the weekly level. As firms on average file one 8-K filing per month, it is more feasible in a weekly panel to control for the 8-K disclosures and study the interaction effect between 8-K disclosures and 8-K demand. Conditional on firms with 8-K disclosures in a given week, stocks with high 8-K views earn high contemporaneous returns and low future returns. Moreover, the result is robust in a smaller sample, in which I also control for the Bloomberg search index and Google Trends.

The results do not imply that the attention to 10-K filings does not reduce information asymmetry. It simply points out that, the asset selection channel is the dominant force here in determining asset returns, given the infrequent disclosure requirement. Besides the standard financial and operation statement, management forecast and risk evaluation are also included in the Form 10-K. Investors who pay attention to these sections could potentially learn important aspects of firms from the managers' perspective and reduce information asymmetry. However, because of the infrequent disclosure of 10-K filings, the asset selection channel dominates the information asymmetry reduction channel most of the time. The granularity of the dataset allows me to disentangle the two channels by focusing on a small subset of the firms. In a given week, I limit my sample to a set of firms just disclosed 10-Ks in the week. I then sort stocks by the size-adjusted view counts of the newly disclosed 10-K filings and form long/short portfolios. I show that the effect of 10-K attention flips to the information asymmetry reduction 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-K in a given week.

Next, I directly test the effect of information demand through information asymmetry reduction channel. In a monthly panel regression, I show that a standard deviation increase in the 8-K attention reduces the next month Amihud (2002) measure by 25 percentiles. The result is robust using Corwin and Schultz (2012) quote spread measure and Easley, Kiefer, and O'Hara (1997) probability of informed trading measure[4]. On the contrary, the demand for other filing types have no significant effects on information asymmetry proxies. Moreover, I show that the magnitude of alpha is greater when the level of ex-ante information asymmetry is high. Using the previous quarter analyst forecast dispersion as a proxy for ex-ante information asymmetry, I find that the 8-K portfolio earns -71 bps per month in alpha for firms with high forecast dispersion and only -13 bps for firms with low dispersion, with the difference being highly significant as well.

I also document the heterogeneous effect of 8-K demand on prices through the cost of information acquisition channel. Verrecchia (1982) extends Grossman and Stiglitz (1980) and provides comparative statics regarding the cost of information acquisition. In particular, ceteris paribus, the informativeness of price is nondecreasing as information acquisition costs are reduced. I use the firm-level share of local viewers and the share of frequent viewers to capture the variation in the cost of information acquisition. Local viewers and frequent

[4]I obtain quarterly PIN measure from Brown and Hillegeist (2007) and aggregate my attention variables into quarterly frequency. The result is available upon request.

viewers tend to have information advantage in terms of collecting and processing information. Firms that have a higher level of local viewers and frequent viewers have a lower cost of information acquisition. My finding is supportive of Verrecchia (1982). The effect of information demand on prices is larger when more demand comes from local and frequent viewers who have a relatively low cost of information acquisition.

Moreover, I show that the effect of 8-K attention is a function of the information content provided in the filings. The demand for 8-K filings predicts stock returns through the information asymmetry channel if the information disclosed was not previously known to the general public. I use the three-day abnormal return around event date as a proxy for the importance of information content. If the abnormal return is large and positive (negative), it is likely that the firm disclosed private and good (bad) news that is not foreseen by the public. Using abnormal return around event date, I show that the effect of 8-K attention on stock returns exhibits a "v-shape" relative to the abnormal returns. That is, demand for 8-K information predicts stock returns better when the information disclosed is private, regardless of whether it is good or bad. When the information contained in the filing is anticipated by the market, investors do not have much to learn, leading to poorer performance of the 8-K portfolio.

Lastly, I test the demand for asset channel, where investors' demand for 10-K filings reflects their preference for the assets. Barber and Odean (2007) shows that investors consider purchasing stocks that are highly visible. Therefore, the demand for 10-K filings should capture the demand for assets better among high visible stocks than among low visible ones. I use the maximum daily return during the month and abnormal trading volume to proxy for stock visibility. I find that the long/short portfolio sorted by 10-K attention earns a monthly alpha of 87 basis points (bps hereafter) among high visible stocks and only 31 bps among low visible stocks. Moreover, the spread in alpha is concentrated among firms with a large population of new visitors who have not downloaded any filings of the firms in the past three months. These new visitors are more likely to be the ones who consider buying the assets.

My paper contributes to the literature that analyzes cross-section stock returns and investor attention. Da et al. (2011) and Ben-Rephael et al. (2017) show that spikes in Google and Bloomberg search volumes can predict positive future stock returns. Different from their papers, my paper focuses on the heterogeneous information content acquired by investors and investors' patterns of information acquisition. Demand for information not only predicts positive and short-term cross-section stock returns through the asset demanding channel but also predicts negative and long-term returns through the reduction in information asymmetry channel.

The paper also fits into the literature that studies information asymmetry. Easley and O'Hara (2004) builds a theoretical model and shows information asymmetry increases the cost of capital. Brown, Hillegeist, and Lo (2004) and Brown and Hillegeist (2007) show empirically that firm disclosures reduce information asymmetry. In this paper, I show that the demand for information also reduces information asymmetry, which affects the cross-section stock returns.

A few recent papers also use the EDGAR log data. Lee, Ma, and Wang (2015) identifies peer firms through a "co-search" algorithm. Loughran and McDonald (2017) shows that investor attention to firm filings is a scarce resource. Bauguess, Cooney, and Hanley (2018) uses EDGAR log data and studies IPO pricing. They focus on Form S-1 (IPO's initial registration statement) and show that investor attention can predict IPO success and initial stock returns. Chen, Cohen, Gurun, Lou, and Malloy (2018) studies how mutual fund managers acquire information on firms and insiders. Gibbons et al. (2019) shows that analysts rely on EDGAR filings to make forecasts.

The paper proceeds as follows. Section 2 discusses the sample selection and provides an overview of EDGAR log data. Section 3 provides my main results on heterogeneous effects of information demand on stock prices. Section 4 discusses the mechanisms of 8-K demand on prices. Section 5 shows the differential effects of 8-K through the cost of information acquisition and information content channels. Section 6 shows that demand for 10-K captures investors' preference for assets. Section 7 concludes.

2 Data and Sample Selection

I combine data from several sources to execute the paper. I use CRSP, Compustat, and I/B/E/S to obtain stock related information, the Thompson Reuters to obtain institutional ownership data, the EDGAR server log to obtain daily log of page requests for SEC filings[5], 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[6]. I also control for Google Trends and Bloomberg News Heat Index. Google Trends data provide within-firm daily Google search volume index and are often used to capture retail investors' attention. Bloomberg index captures the news search volume by Bloomberg users and are used to capture institutional investors' attention. The data start from 2010/02/17.

The sample starts in 2003 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 valid earning announcement data in I/B/E/S. I also require that stocks in the CRSP have matched identifiers in the SEC EDGAR. 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 microcap stocks. For the main analysis, I will use the larger sample. All my results are robust when using the smaller sample.

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

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

2.1 The EDGAR Server Log

The SEC EDGAR server 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[7], the index page flag[8], server status code[9], 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 use automated programs to crawl and download SEC filings, and the EDGAR log files flag not all crawling activities. Following Lee et al. (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. I first reduce the sample by excluding requests with the index page flag or server status code above 300, which leaves me with 9.84 billion records. 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. Forms 10-K and 10-Q contain comprehensive reports of the firm performance for the recent fiscal year and quarter, respectively. Form 8-K is a report of unscheduled material events that are important to the shareholders and the SEC. These set of forms represent the most relevant information of a firm's operation and performance and are important to investors and financial analysts. I then get the physical locations of IP addresses in the record. Finally, I have a sample of log requests with 1.36 billion records.

For each filing request, I classify it along three dimensions, the filing type (Forms 10-Q, 10-K, and 8-K), the geographical distance between the requested IP and firm headquarters,

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

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

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

and whether the IP is a frequent visitor of the firm.

Classifying filing requests by filing types has several advantages. First, different filing types represent different information flows. Forms 10-K/Q are filed periodically, containing financial and operational statements of the firm over the past year/quarter. Thus, they provide investors with a comprehensive overview of the firm. Forms 8-K are filed whenever firms are required to disclose material information, or managers judge it necessary to disclose voluntarily[10]. Therefore, 8-K filings contain the most up-to-date information.

Furthermore, if the distance between the locations of the IP address and the firm headquarter is less (greater) than 400 miles, I classify the request as a home (away) request. I then aggregate the file requests at the firm and the month level.

2.2 Overview of EDGAR Downloads

Figure 1 shows the monthly aggregated file downloads in my final sample. I separate crawling activities ("robots") from human viewing activities ("human"). Figure 1a shows the plot for all filing types. There has been an increasing trend for viewing activities on EDGAR. 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 1b to 1d show the monthly aggregated plot by file types. The strong seasonality in 10-K and 10-Q downloads are driven by the filing cycles. 10-K downloads consist around half of all downloads, with the remaining half split by 8-K and 10-Q downloads.

2.3 Difference in 10-K and 8-K Views

The EDGAR log data keep track of each IP visit so that I can distinguish viewing patterns of 10-K and 8-K visitors. Figure 2 shows that the majority of 10-K visitors never downloaded any 8-K filings of the firm in the past quarter. Moreover, 10-K visitors are different from 8-K visitors in terms of geographical distance to headquarters and the visiting frequency.

[10] Voluntary disclosures are often categorized by the SEC into Item 8.01 and Item 7.01.

Figure 3 shows the number of filing downloads by geographical distance groups. I classify a filing view into the home group if the distance between the location of the IP address and the firm headquarter is less than 400 miles. Otherwise, I classify it into the away group. The cut-off of 400 miles is about the diameter of an average state in the U.S. I denote the number of home (away) filing views as $views_k^{home}$ ($views_k^{away}$). The result suggests that a large proportion of downloads are made by users geographically close to the firm headquarter, especially for 8-K filings. 10-K filings, on the other hand, face a wider range of audience, as the gap between away $views_{10K}^{away}$ and $views_{10K}^{home}$ widens.

Figure 4 plots the time-series averages of frequent visitor ratios by 10-K and 8-K visitors. I classify a filing download as frequent if the IP address has downloaded one or more filings of the firm in the past three months. Frequent visitors constitute around 15% of 10-K visitors and around 35% of 8-K visitors.

3 Heterogeneous Effects of EDGAR Attention on Stock Returns

In this section, I use two approaches to study how demand for information affects asset prices. First, I use Fama-Macbeth (1973) cross-section regression. Second, I form long/short portfolios based on their attention levels.

3.1 Fama-Macbeth (1973) Approach

I first study the relation between future stock returns and the overall attention on EDGAR. I run Fama-Macbeth (1973) regression of monthly individual stock returns from month $t+1$ on investor attention variables from month t. *log views*$_k$ is the monthly natural logarithm of form k's total viewing counts. I consider forms 10-K, 10-Q, and 8-K.

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 log of the book-to-market ratio; *log(ME)* is the natural log of the firm market capitalization; *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 attention and future returns. Gervais, Kaniel, and Mingelgrin (2001) and Barber and Odean (2007) document that abnormal trading volume increases a firm's visibility, which could affect both attention 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 attention variables capture investor attention to firms' annual and quarterly filings, I include earning surprise and earning drift from the most recent earnings announcement. *SUE* is the unexpected quarterly earnings scaled by market cap; *Earning Drift* is the sum of daily returns in three days around earning announcement. Lastly, I control for firm disclosure. *file 8K*, *file 10K*, and *file 10Q* are the numbers of Form 8-K, 10-K, and 10-Q issued by the firm on the EDGAR, respectively.

Column 1 of Table 1 shows the baseline result. Asset growth, firm size, operating profit, unexpected earnings, abnormal trading volume, and abnormal earning announcement returns can explain cross-section stock returns. I then add the overall investor attention to EDGAR filings, *log* $views_{all}$, to the previous specification. *log* $views_{all}$ is the natural log of all filing views of the firm in the current month. Column (2) shows the regression result. The estimate of *log* $views_{all}$ is positive and significant. Firms with high filing views earn a premium of roughly 18.3 basis points per month (2.2% per year). The effect of aggregated demand for information is consistent with the literature, that the asset selection channel plays a dominant force in determining stock returns.

To disentangle the asset selection channel and information asymmetry reduction channel, I split the overall views by their filing types. Demand for 10-K and 10-Q is more likely to capture the general demand for the asset, 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. Forms 8-K, on the other hand, are filed irregularly and contains information that is privately known by insiders. Demand for 8-K transforms the disclosed information into public information and reduces the information asymmetry of the firm.

The result is shown in Column (3). The coefficient estimates of $\textit{log views}_{10K}$ are positive (39 bps per month) and highly significant with t-stat of 7.42, which is consistent with 10-K views capturing the asset demand. The coefficient of $\textit{log views}_{10Q}$ is insignificant, and its magnitude is relatively small. There are two driving forces of the result. First, the correlation between 10-K views and 10-Q views is 0.91 over the full panel. As a result, 10-Q views do not provide additional variation beyond 10-K views in explaining stock returns in the future. Second, the substance and quality of forms 10-K and 10-Q differ. Forms 10-K are required to be audited, whereas Forms 10-Q are not. In addition, the "Management's Discussion and Analysis" section in Forms 10-K are much more detailed than in Forms 10-Q[11]. As a result, Form 10-K is a more reliable source for investment reference than Form 10-Q.

The demand for 8-K captures the information asymmetry reduction channel. Demand for 8-K filings reduces the information asymmetry of the firm. As a result, investors require a lower risk premium to hold the asset. Firms with high 8-K filing attention earn less return in the future (12 bps per month), as can be seen from the negative and significant coefficient estimates of $\textit{log views}_{8K}$. To further sharpen the result, I split the 8-K views into the scheduled and unscheduled 8-K views. The scheduled 8-K filing includes pre-scheduled event, such as earnings announcements and annual shareholder meetings. These scheduled reports typically contain information that is known by the market. In Column (4), the coefficient of unscheduled 8-K views is negative and significant, which further supports the hypothesis.

In Column (5), I control for the change in Google Trends and media news coverage. The

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

result is robust, but the sample is smaller than the ones in previous columns.

3.2 Portfolio Sort Approach

In this section, I study the effect of investor attention to firm filings using a portfolio sorting approach. Since large firms naturally receive higher views than small firms, it is important to control for firm sizes when sorting on views, especially for the 10-K views[12]. At each month, I first run a cross-section regression of the natural log views on the natural log of lag firm size and extract regression residuals as the size-adjusted attention. I then sort stocks into quintiles by the size-adjusted attention. 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.

Panel A of Table 2 shows the portfolio sort result for 10-K views. The monthly one-month holding return of the equal-weighted long/short portfolio is 0.9% and highly significant. After controlling for common pricing factors, the average alpha is around 0.65% per month. Moreover, the effect of 10-K attention is short-termed, which can be seen from the insignificant alphas with three or twelve holding months.

Panel B of Table 2 shows the portfolio sort result for 8-K views. Consistent with the result in Fama-Macbeth (1973) regression, the equal-weighted long/short portfolio earns a monthly alpha of -0.56%. Moreover, the effect of 8-K attention is long-lasting. The 12-month holding alpha is around -0.6% per month and highly significant.

Panels C and D of Table 2 shows the factor loadings of one-month 10-K and 8-K portfolios. The alphas of 10-K (8-K) portfolios are monotonically increasing (decreasing) with the level of 10-K (8-K) views.

The two channels of information demand have not only opposite predictions in the short

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

term, but also suggest distinct patterns in the long term. For the asset selection channel, we should see the strongest evidence of 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 investment decision making. For the information asymmetry reduction channel, the contemporaneous price should increase to reflect the risk reduction going forward, followed by a permanent decrease in risk premium. Therefore, the contemporaneous return spread is positive, and future return spread is negative and persistent.

Figure 6 shows the persistence of long/short portfolio return and alpha. I sort stocks by size-adjusted 10-K (8-K) attention into quintiles at month t, and study the long/short portfolio returns at month $t + k$, where k ranges from 0 to 12. For the 10-K portfolio, the return and alpha are pronounced the most in the portfolio formation month and decay very quickly. The average alpha is 86 bps at month 0, 24 bps at month 2, and only 6 bps at month 4. The 8-K portfolio shows the opposite pattern. The portfolio return at formation month is positive and then becomes negative and highly persistent over the next 12 months. The distinct long term return patterns of 10-K and 8-K portfolios strongly support the hypothesis, that the demand for 10-K filings captures the general asset demand and the demand for 8-K filings reduces the information asymmetry of the firm.

Since I use size-adjusted views as the sorting variables, it is interesting to see how 10-K and 8-K portfolios perform under different size groups. At each month, I first sort stocks by their previous month market capitalization into quintiles. Conditional on each size quintile, I then sort stocks by the size-adjust 10-K (8-K) views into quintiles and form the long/short portfolio by buying the top quintile stocks and selling the bottom quintile stocks.

Figure 7 shows the result. Portfolios sorted on size-adjusted 10-K views yield positive and significant alphas across all size quintiles. The result is the strongest in the small size quintile, yielding 1.2% alpha per month. The magnitude of the alpha decreases with firm size. Both the 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 a larger price impact than large stocks. Moreover, small stocks have a tighter short-selling constrain than large stocks, which also limits the potential arbitrage opportunities and results in a large price increase.

Portfolios sorted on size-adjusted 8-K views yield negative alphas across all size quintiles but only have significant alphas for the bottom three size quintiles. 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. The result is consistent with the information asymmetry reduction hypothesis. Small firms have less media/analyst coverage and institutional holding than large firms. Investors of small firms face a higher degree of information asymmetry and rely more on themselves in processing and incorporating the disclosed information. Therefore, demand for 8-K filings has a stronger effect in small firms than in large ones.

3.3 Weekly Frequency Result

This section demonstrates the heterogeneous effect of information acquisition on stock prices at the weekly frequency, which allows me to better control for the information supply. Forms 10-K (10-Q) are filed once (three times) a year in general. Forms 8-K are filed irregularly, but once a month on average. Therefore, an analysis conducted at the weekly frequency can better disentangle the effects of information supply and demand on prices.

I aggregate the daily stock returns and daily views to weekly frequency (Friday close to Friday close). My main variable of interest is $\log(views_t^k)$, which is the natural logarithm of total views 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 $\textit{Earnings Release}_t$ is equal to one if the firm releases its earnings in week t. For a subset of the analysis[13], I also

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

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. (2017), 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 in any day of the week is above its 90 percentile in the past month.

Table 4 shows the weekly regression result. I regress the weekly stock returns on the demand for filings, controlling for the supply side of information, firm characteristics, lag returns, and time fixed effects. To better capture the interaction between supply and demand of information under information asymmetry channel, I add the 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 the next week stock returns.

As shown in Columns (1) and (2) of Table 4, the coefficient estimates of $\log(views_t^{10K})$ are all positive and significant, consistent with the asset selection hypothesis. The coefficient estimate of the interaction term between 8-K supply and demand is positive and significant in Column (1), and is negative and significant in the Column (2). The results support the information asymmetry reduction channel. When firms release new information through 8-K filings, it relies on investors to acquire the information, thus reducing the information asymmetry. Therefore, the contemporaneous price increases, and future risk premium decreases. Both the asset selection channel and information asymmetry reduction channel are robust after controlling for the Bloomberg and Google Trends search indexes, which are shown in Columns (3) and (4).

Figure 8 replicates the analysis in Figure 6 using weekly data, and the results are consistent. At the end of each week, I sort stocks by the size-adjusted weekly 10-K and 8-K views into quintiles, and form long/short portfolios. Portfolios are held throughout the next

24 weeks. The alphas of portfolios at each holding week is plotted. For 8-K portfolios, I limit the set of stocks that filed 8-K filings in the week, as the evidence suggested in Table 4 shows that the effect of 8-K demand is stronger, conditional on the supply of information.

The results do not imply that the demand for 10-K filings does not reduce information asymmetry. It merely states that the asset selection channel dominates, and it is hard to empirically disentangle the two channels because of the low disclosing 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 10-Ks in a week. I then sort these stocks on the size-adjusted view counts of the newly issued 10-K filings into quintiles. The result is plotted in Figure 9. Conditional on firms just issued 10-Ks in week 0, firms with high attention to the new document yield higher returns in the contemporaneous week, and lower returns in the upcoming weeks. The effect of 10-K demand flips to the information asymmetry reduction 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-K in a given week.

4 Mechanisms of 8-K Demand on Stock Returns

Easley and O'Hara (2004) documents that investors demand higher returns for stocks with more private information. Boot and Thakor (2001) suggests that disclosing information that is only known to informed investors decreases the information advantage informed investors have over the uninformed. However, little study has shown the effect of information demand on information asymmetry, as past literature mainly focuses on the supply side. In this section, I show that investors' 8-K attention decreases the proportion of private information, which then leads to a reduction in information asymmetry. As a result, stocks become less risky for uninformed investors to hold and expected returns decrease. Therefore, we observe that stocks with high 8-K attention underperform stocks with low 8-K attention and the underperformance is highly persistent over time. Moreover, the effect of 8-K attention should also depend on the ex-ante information asymmetry the firm is facing.

Table 5 shows the monthly panel regression results of future information asymmetry on investor attention to firm filings, controlling for firm characteristics and information disclosure. I use Amihud (2002) measure and quote spread estimated following Corwin and Schultz (2012). The coefficient estimates of $log\ views_{8K}$ is negative and significant, suggesting that higher 8-K attention is associated with lower information asymmetry in the next month. Moreover, the economic magnitude of the coefficient is large. One standard deviation of 8-K attention can move Amihud measure by 25 percentiles.

The effect of 8-K attention on stock returns should be larger when the ex-ante information asymmetry is higher. To proxy for ex-ante information asymmetry, I use Amihud illiquidity measure and previous quarter analyst forecast dispersion. I first sort stocks by information asymmetry measures into terciles. Conditional on each tercile, I sort stocks by size-adjusted 8-K attention into quintiles. Table 6 shows the portfolio double-sort results for 8-K attention and information asymmetry. When Amihud measure is low, the alpha of long/short 8-K attention portfolio is -12 bps per month. When Amihud measure is high, the magnitude of alpha increases to -72 bps per month. The result is similar using analyst forecast dispersion.

5 Heterogeneous Effect of 8-K Demand

This section shows heterogeneous effects of 8-K demand on stock returns. The heterogeneity stems from the cost of information acquisition and the information contents of the filings.

5.1 Cost of Information Acquisition

The cost of information acquisition plays an important part in reducing information asymmetry. In Verrecchia (1982) Corollary 4, the informativeness of price is nondecreasing as information acquisition costs are reduced. Although I do not directly observe the cost of information acquisition of each investor, an investor's past information acquisition history and his/her geographical location are observed in the data. I use the firm-level share of local demand and share of frequent viewers to capture the cost of information acquisition.

Local investors have information advantage to collect and process information over non-local investors. 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 of the firm in the past quarter than one who did not. Therefore, frequent visitor ratio defined in Figure 4 can be used as a proxy for the cost of information acquisition. The higher the frequent visitor ratio is, the lower the cost of information acquisition.

Panel A of Table 7 shows the portfolio double-sort results by investor attention and the average distance of viewer location to firms' headquarters. For each stock at 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 size-adjusted views into quintiles. The effect of 8-K attention is mainly concentrated in the low (-48 bps/month) and medium (-65 bps/month) distance terciles, and much weaker in the high (-18 bps/month) distance tercile. Moreover, the difference between high and low terciles is statistically significant.

Panel B of Table 7 studies the attention effects, conditional on visitors' past visiting patterns. For each firm-month, I calculate the proportion of frequent visitors. I then double sort stocks by the frequency ratio into terciles and by size-adjusted 8-K views into quintiles. Portfolios sorted by 8-K attention show significant and negative alphas when views are from frequent visitors (-42 bps/month). When the frequent ratio is low, however, the 8-K portfolio yields an insignificant alpha of -25 bps per month.

5.2 Information Content

Moreover, the effect of 8-K attention should be a function of the information content provided in the filings. The demand for information only reduces the information asymmetry 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 material agreement and change of officers, require investor attention to interpret the information. Therefore, it is important to see how the effect of information demand 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[14]. 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 standard deviation of 12%.

The second measure is constructed using textual analysis. 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 the 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 difference between the realized market response and the expected market response, and use this "un-

[14]Starting 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.

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

To see how the effect of 8-K attention varies with the importance of information content of the filing, I double sort stocks by the size-adjusted 8-K views and the above two measures. The result is shown in Table 8. 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 8-K attention and abnormal return exhibits a "V-shape". The effect of 8-K attention 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 reduce 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 abnormal return around event/post date, it is likely that the market has already taken into account the information content, which leaves investors not too much to learn in the first place. As a result, the spread in 8-K attention does not predict future returns well.

6 Mechanism of 10-K Demand on Stock Returns

Barber and Odean (2007) documents that attention is a scarce resource, and demand for assets is rooted from the stocks that grab investor attention. When investors make purchasing decisions for a stock, 10-K filings provide the most comprehensive coverage of the operational and financial details of a firm. Therefore, the effect of attention to 10-K filings on stock prices is a byproduct of demand shocks to assets. That is, the demand for asset drives up the demand for 10-K filings and stock prices. As a result, we should expect the effect of 10-K attention on stock prices to be higher for attention-grabbing stocks, where the demand shock is potentially higher.

I use stocks with high abnormal trading volume and high daily absolute returns to proxy for attention-grabbing stocks. Abnormal trading volume and daily absolute returns are constructed as the following,

$$abvol_{i,t} = \frac{vol_{i,t} - \bar{vol}_{i,t-1,t-12}}{std\_vol_{i,t-1,t-12}}, \tag{1}$$

$$max\_dret_{i,t} = \max_{d \in t} \mid ret_{i,t,d} \mid \tag{2}$$

where $\bar{vol}_{i,t-1,t-12}$ and $std\_vol_{i,t-1,t-12}$ are the mean and standard deviation of monthly trading volume during the past 12 month, respectively. $ret_{i,t,d}$ is the daily return of stock i on month t and day d. Gervais et al. (2001) first documents that stocks with abnormally high trading volume earn return premiums in the future. The argument is that shocks to the trading volume of a stock increase its visibility, which draws investor attention and drives up stock prices. Barber and Odean (2007) uses abnormal trading volume and daily maximum return to proxy for attention-grabbing. I first sort stocks by abnormal trading volume (daily absolute returns) into terciles. Conditional on each tercile, I then sort stocks by the size-adjusted 10-K attention into quintiles.

Panel A of Table 9 show the alphas of double-sorted portfolios for abnormal trading volume and 10-K views. The last column shows the alphas of long/short attention portfolios conditional on abnormal volume terciles. For low abnormal trading volume tercile, the spread in alpha is 31 bps per month. The spread in alpha increases to 87 bps per month for stocks in high abnormal trading volume tercile. Panel B shows the result using the maximum daily return as a proxy for attention-grabbing. The result is very similar to Panel A. Moreover, the spread is mostly driven by the outperformance of high attention stocks.

7 Conclusion

In this paper, I empirically test two channels where demand for information affects asset prices using EDGAR log data. On the one hand, investor attention to form 10-Ks is associated with a general demand for the asset, so that a spike in 10-K attention predicts short-term positive future returns. The effect of 10-K attention is stronger among attention-grabbing stocks. The alpha of the 10-K attention portfolio decays sharply after the first month, consistent with the demand shock. On the other hand, investor attention to form 8-Ks decreases information asymmetry. As a result, high 8-K attention stocks persistently underperform low 8-K attention stocks due to the reduced risk premium. The effect of 8-K attention is stronger when firms have high ex-ante information asymmetry.

This paper also sheds lights on the micro-level information acquisition patterns. The striking difference in 10-K and 8-K attention is attributed to the heterogeneous viewing patterns of investors. Investors who pay attention to form 10-Ks are different from the ones who view 8-Ks, as 80% of 10-K viewers did not download any 8-K filings over the previous quarter. Moreover, 10-K investors are typically one-time viewers, in the sense that they seldom visit company filings in the past. On the contrary, 8-K investors are frequent visitors who regularly review both old and new filings. Lastly, 8-K investors have a larger proportion of local investors, who presumably have the advantage to acquire information.

References

Amihud, Yakov, 2002, Illiquidity and stock returns: cross-section and time-series effects, *Journal of financial markets* 5, 31–56.

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, 2000, Information acquisition in financial markets, *The Review of Economic Studies* 67, 79–90.

Barlevy, Gadi, and Pietro Veronesi, 2007, Information acquisition in financial markets: a correction, Technical report.

Bauguess, Scott W., John Cooney, and Kathleen Weiss Hanley, 2018, Investor Demand for Information in Newly Issued Securities, SSRN Scholarly Paper ID 2379056, Social Science Research Network, Rochester, NY.

Ben-Rephael, Azi, Bruce Carlin, Zhi Da, and Ryan Israelsen, 2017, Demand for Information and Asset Pricing, Technical Report w23274, National Bureau of Economic Research, Cambridge, MA.

Ben-Rephael, Azi, Zhi Da, Peter D Easton, and Ryan D Israelsen, 2017, Who pays attention to sec form 8-k? .

Bernile, Gennaro, Alok Kumar, and Johan Sulaeman, 2015, Home away from Home: Geography of Information and Local Investors, *The Review of Financial Studies* 28, 2009–2049.

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

Brown, Stephen, and Stephen A Hillegeist, 2007, How disclosure quality affects the level of information asymmetry, *Review of Accounting Studies* 12, 443–477.

Brown, Stephen, Stephen A Hillegeist, and Kin Lo, 2004, Conference calls and information asymmetry, *Journal of Accounting and Economics* 37, 343–366.

Chen, Huaizhi, Lauren Cohen, Umit Gurun, Dong Lou, and Christopher Malloy, 2018, IQ from IP: Simplifying Search in Portfolio Choice, Working Paper 24801, National Bureau of Economic Research.

Corwin, Shane A, and Paul Schultz, 2012, A simple way to estimate bid-ask spreads from daily high and low prices, *The Journal of Finance* 67, 719–760.

Coval, Joshua D., and Tobias J. Moskowitz, 1999, Home Bias at Home: Local Equity Preference in Domestic Portfolios, *The Journal of Finance* 54, 2045–2073.

Coval, Joshua D., and Tobias J. Moskowitz, 2001, The Geography of Investment: Informed Trading and Asset Prices, *Journal of Political Economy* 109, 811–841.

Cziraki, Peter, Jordi Mondria, and Thomas Wu, 2018, Asymmetric Attention and Stock Returns, SSRN Scholarly Paper ID 1772821, Social Science Research Network, Rochester, NY.

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

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, Nicholas M Kiefer, and Maureen O'Hara, 1997, One day in the life of a very common stock, *The Review of Financial Studies* 10, 805–835.

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

Engelberg, Joseph, R. David McLean, and Jeffrey Pontiff, 2017, Anomalies and News, SSRN Scholarly Paper ID 2631228, Social Science Research Network, Rochester, NY.

Epstein, Larry G., and Martin Schneider, 2008, Ambiguity, Information Quality, and Asset Pricing, *The Journal of Finance* 63, 197–228.

Fang, Lily, and Joel Peress, 2009, Media coverage and the cross-section of stock returns, *The Journal of Finance* 64, 2023–2052.

García, Diego, and Øyvind Norli, 2012, Geographic dispersion and stock returns, *Journal of Financial Economics* 106, 547–565.

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, 2019, Analyst information acquisition via edgar, *Available at SSRN 3112761* .

Green, Jeremiah, John RM Hand, and X Frank Zhang, 2017, The characteristics that provide independent information about average us monthly stock returns, *The Review of Financial Studies* 30, 4389–4436.

Grossman, Sanford J, and Joseph E Stiglitz, 1976, Information and competitive price systems, *The American Economic Review* 246–253.

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

Hong, Harrison, and Jeremy C Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, *The Journal of finance* 54, 2143–2184.

Hwang, Byoung-Hyoun, and Hugh Hoikwang Kim, 2017, It pays to write well, *Journal of Financial Economics* 124, 373–394.

Lawrence, Alastair, 2013, Individual investors and financial disclosure, *Journal of Accounting and Economics* 56, 130–147.

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.

Lev, Baruch, and Stephen H Penman, 1990, Voluntary forecast disclosure, nondisclosure, and stock prices, *Journal of Accounting Research* 28, 49–76.

Livnat, Joshua, and Yuan Zhang, 2012, Information interpretation or information discovery: which role of analysts do investors value more?, *Review of Accounting Studies* 17, 612–641.

Loughran, Tim, and Bill McDonald, 2014, Information decay and financial disclosures, Technical report, working paper, University of Notre Dame.

Loughran, Tim, and Bill McDonald, 2017, The Use of EDGAR Filings by Investors, *Journal of Behavioral Finance* 18, 231–248.

Lucca, David O, and Emanuel Moench, 2015, The pre-fomc announcement drift, *The Journal of Finance* 70, 329–371.

Manela, Asaf, 2014, The value of diffusing information, *Journal of Financial Economics* 111, 181–199.

Nieuwerburgh, Stijn Van, and Laura Veldkamp, 2009, Information Immobility and the Home Bias Puzzle, *The Journal of Finance* 64, 1187–1215.

Patton, Andrew J., and Michela Verardo, 2012, Does Beta Move with News? Firm-Specific Information Flows and Learning about Profitability, *The Review of Financial Studies* 25, 2789–2839.

Savor, Pavel, and Mungo Wilson, 2013, How Much Do Investors Care About Macroeconomic Risk? Evidence from Scheduled Economic Announcements, *Journal of Financial and Quantitative Analysis* 48, 343–375.

Savor, Pavel, and Mungo Wilson, 2014, Asset pricing: A tale of two days, *Journal of Financial Economics* 113, 171–201.

Soroka, Stuart N, 2006, Good news and bad news: Asymmetric responses to economic information, *The journal of Politics* 68, 372–385.

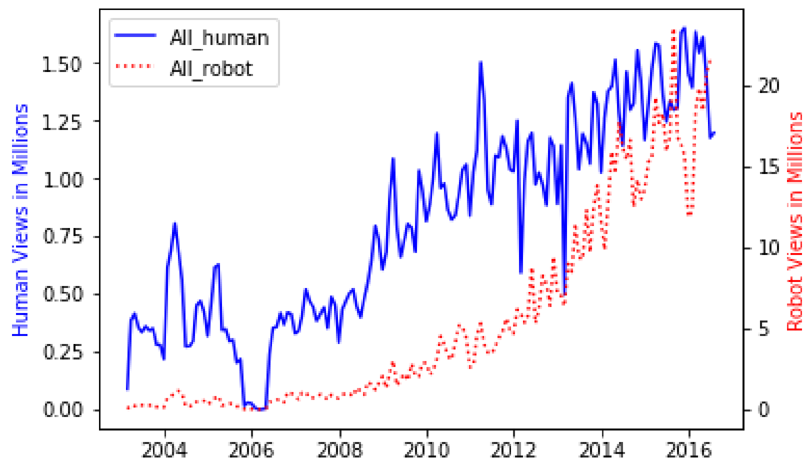
Tetlock, Paul C, 2010, Does public financial news resolve asymmetric information?, *The Review of Financial Studies* 23, 3520–3557.

Tetlock, Paul C, 2011, All the news that's fit to reprint: Do investors react to stale information?, *The Review of Financial Studies* 24, 1481–1512.

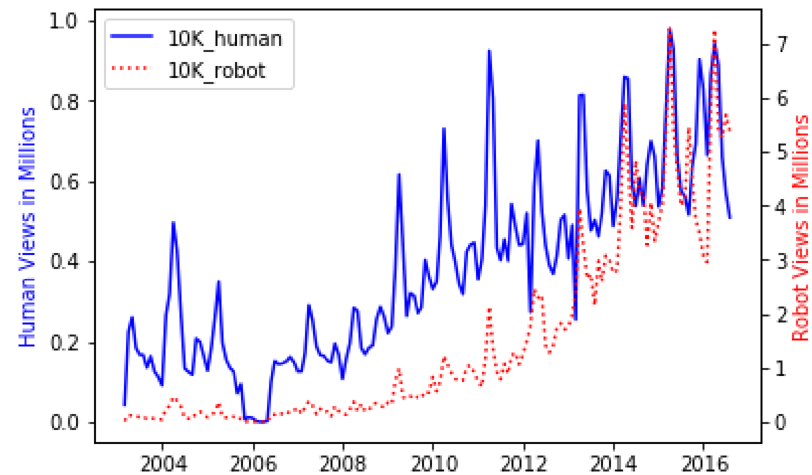
Veldkamp, Laura L., 2006, Information Markets and the Comovement of Asset Prices, *The Review of Economic Studies* 73, 823–845.

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

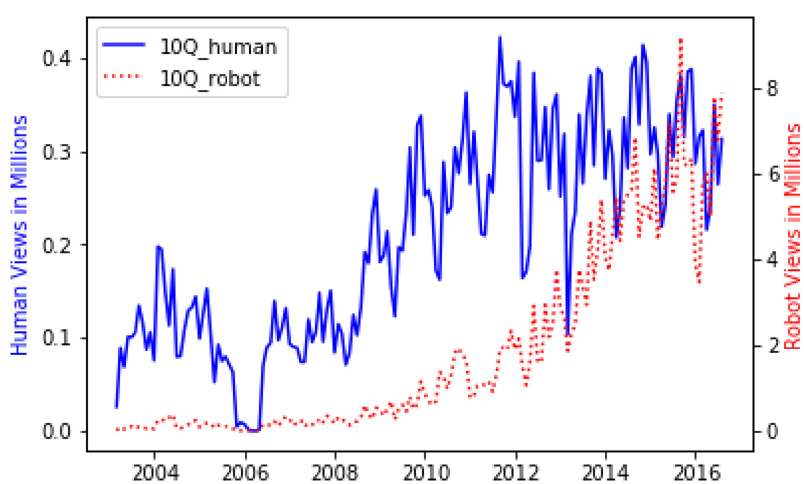
Figure 1
Time-series EDGAR Viewing Activity
The figure shows the monthly aggregated number of views on EDGAR Log system. Following Lee et al. (2015), I separate crawling activities ("robot") from human viewing activities ("human"). Figures (b) to (d) show the number of views for 10-K, 10-Q, and 8-K filings, respectively.



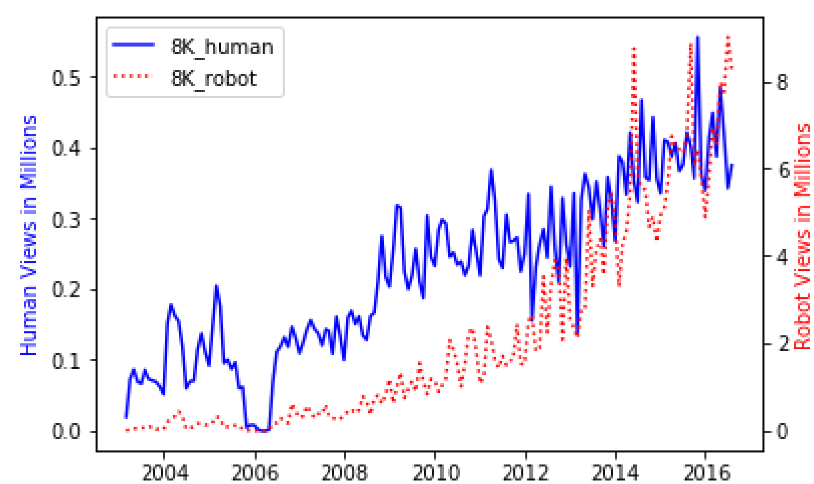
(a) All Filings



(b) 10-K Filings



(c) 10-Q Filings



(d) 8-K Filings

Figure 2
10-K Views Conditional on 8-K Views

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. $views_{10K}^{only}$ is the total number of 10-K views by visitors who have not downloaded any 8-K filings of the firm. $views_{10K}^{both}$ is the total number of 10-K views by visitors who have downloaded one or more 8-K filings of the firm.

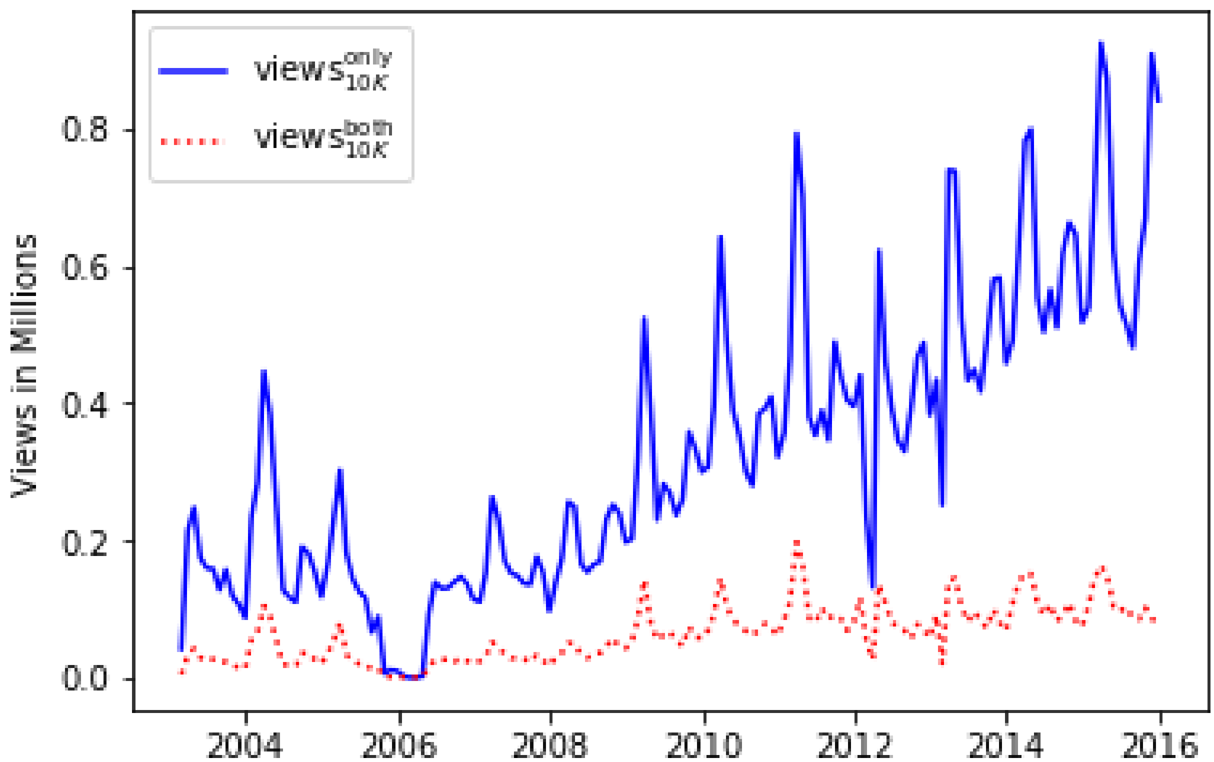
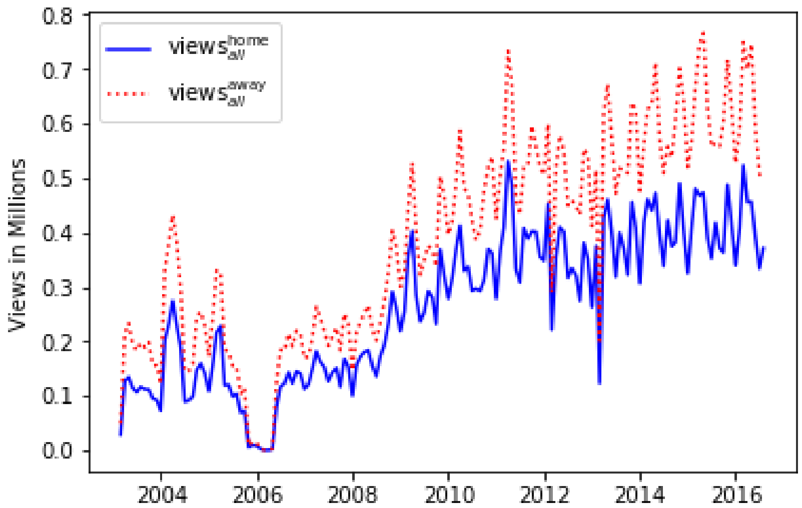


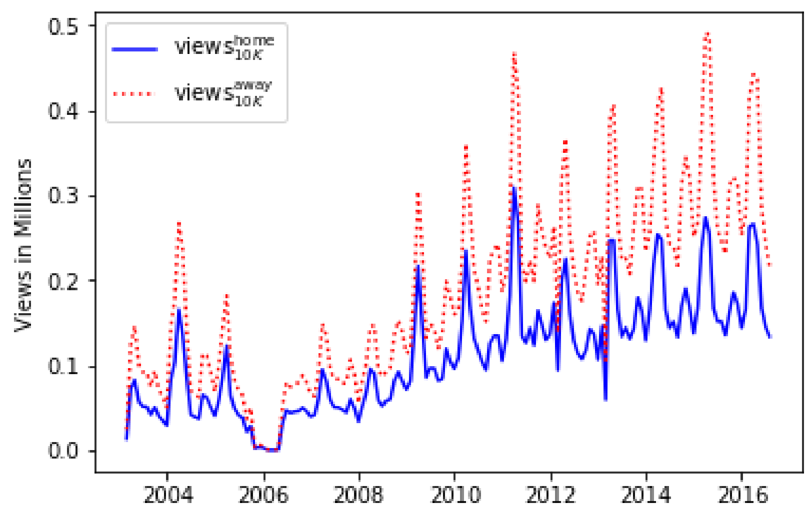
Figure 3

Viewing Activities by Geographical Distance

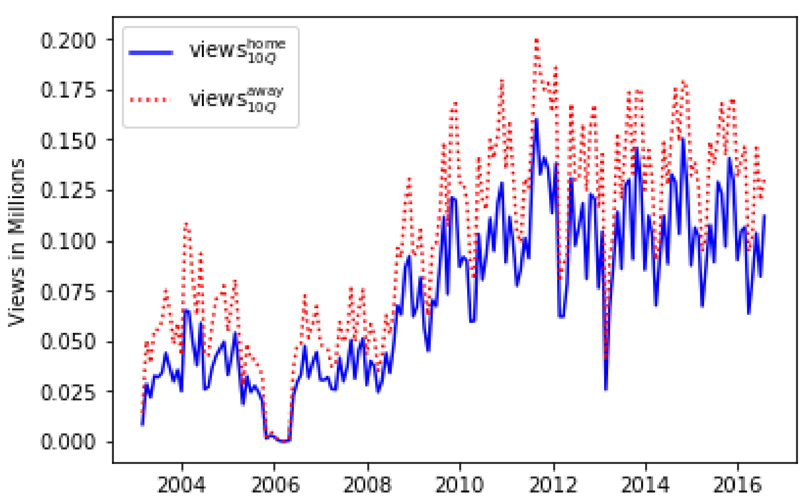
The figure shows the number of views by geographical distance. I classify a filing view as home if the distance between the locations of viewing IP and headquarter is less than 400 miles.



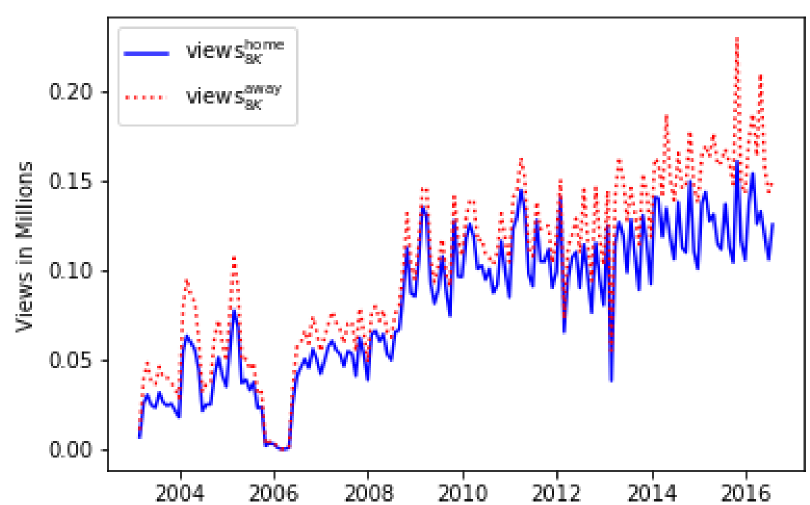
(a) All Filings



(b) 10-K Filings



(c) 10-Q Filings



(d) 8-K Filings

Figure 4
Time-series Frequent Visitor Ratios

The figure shows the time-series plot of frequent visitor ratios by 10-K and 8-K visitors. For each firm and IP address, I classify a filing view as frequent 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 frequent ratios by 10-K and 8-K filings.

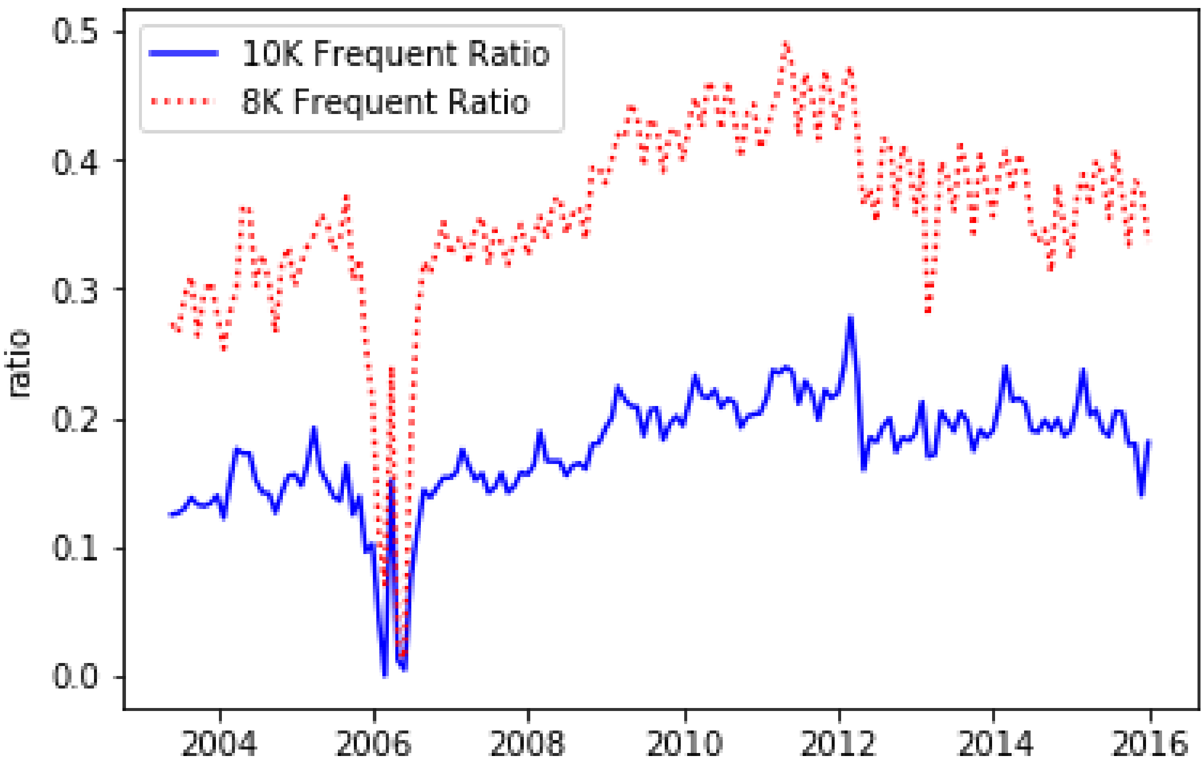


Figure 5

Investor Attention Histogram by Firm Sizes

The figure shows the histogram of investor attention on EDGAR, grouped by firm sizes. The horizontal axis is the natural log of monthly filing views 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-size firm is defined with a firm market cap between 20% and 80% NYSE percentile.

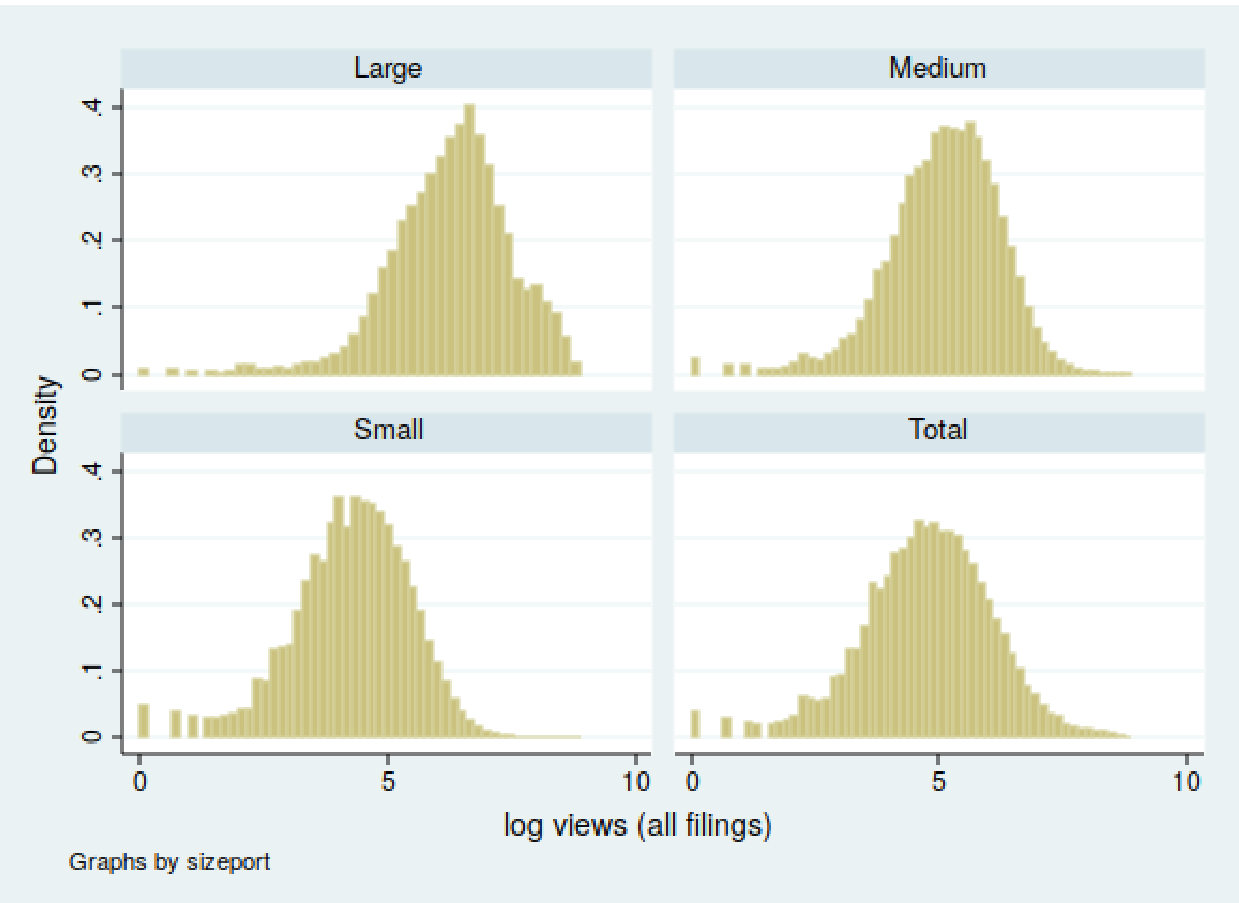


Figure 6

Long/Short Investor Attention Portfolio Return

The figure shows the monthly long/short portfolio returns and alphas for up to 12 months. Stocks are sorted by the size-adjusted 10-K (8-K) views into quintiles. I then form long/short portfolios and plot the next 12-month average portfolio return and Fama French five-factor alphas. For 8-K views, I only focus on the views on unscheduled filings.

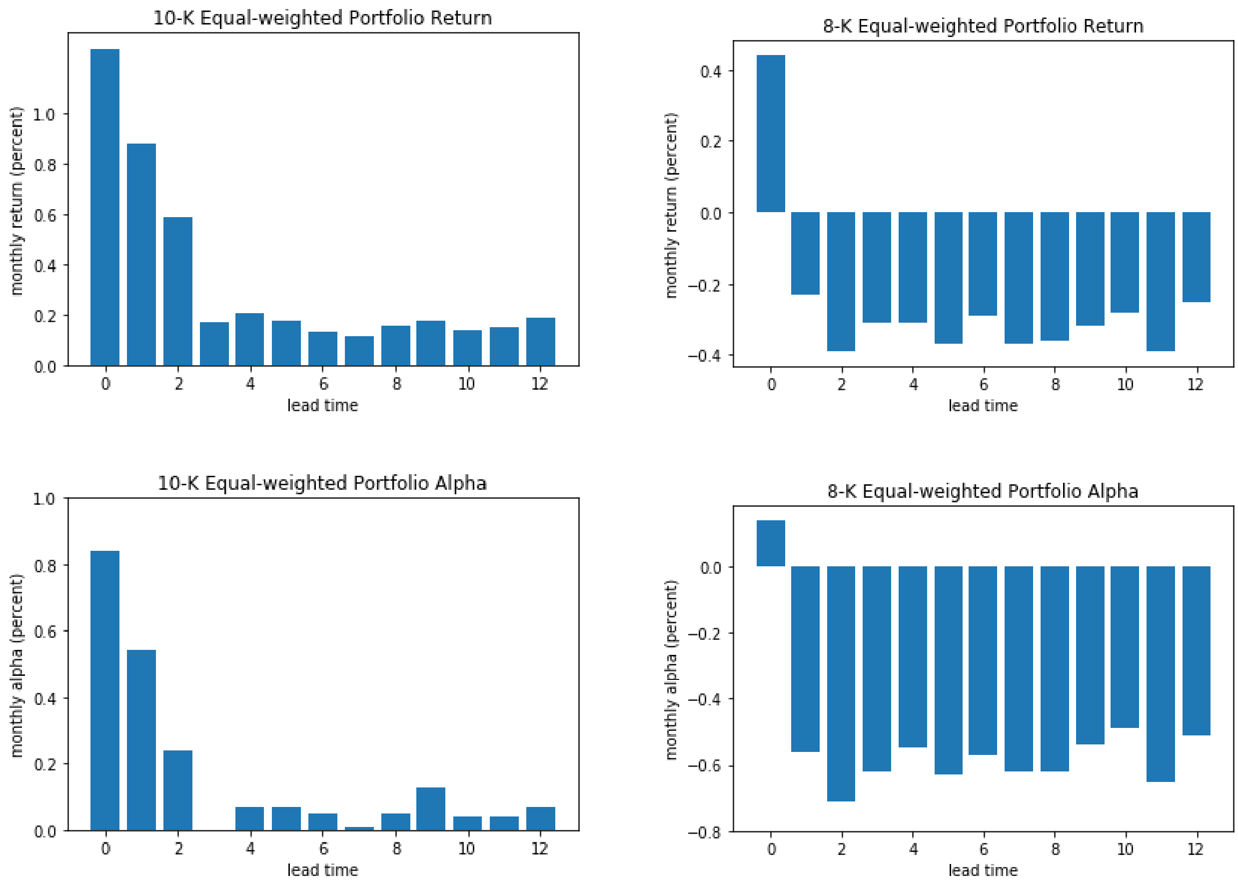
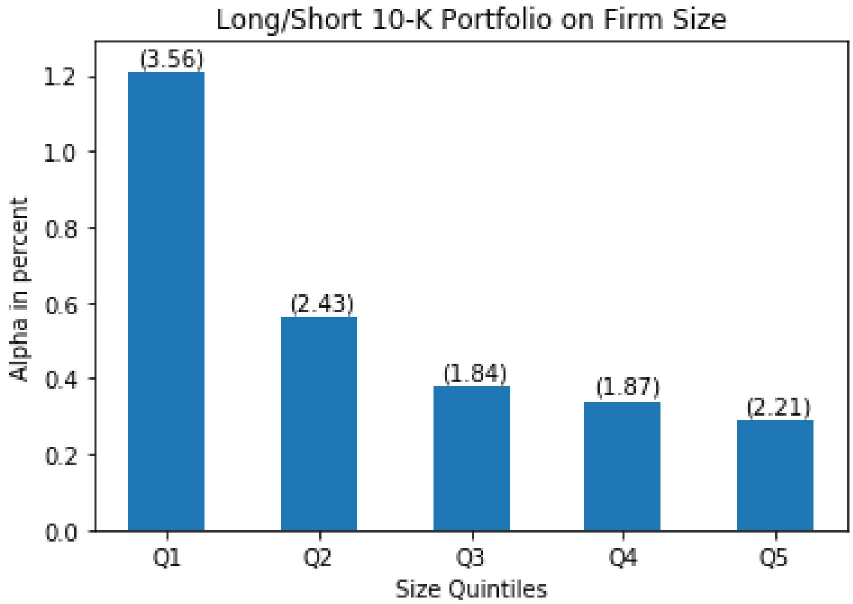


Figure 7

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

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



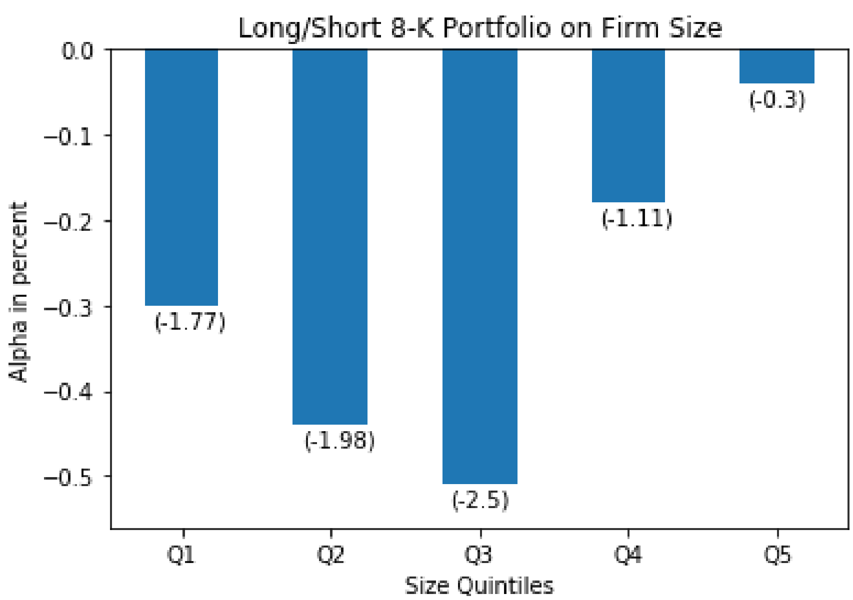


Figure 8

Long/Short Investor Attention Portfolio - Weekly Returns

The figure shows the weekly Fama French 5-factor alphas of 10-K and 8-K portfolios. Stocks are sorted by the size-adjusted weekly views at the end of Friday. Long/short portfolios are held throughout the next 24 weeks.

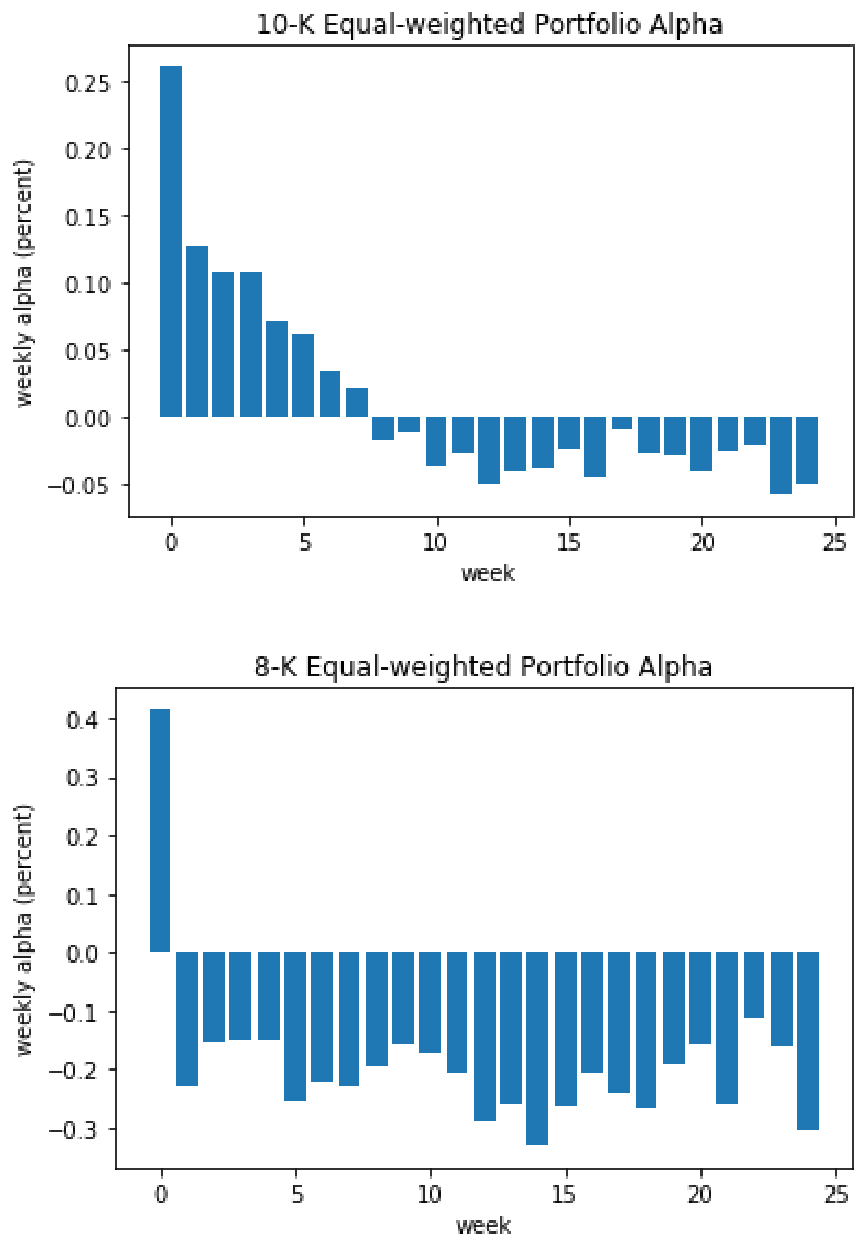


Figure 9

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 by the size-adjusted view counts of the newly disclosed 10-K filing into quintiles. Long/short portfolios are held throughout the next 24 weeks.

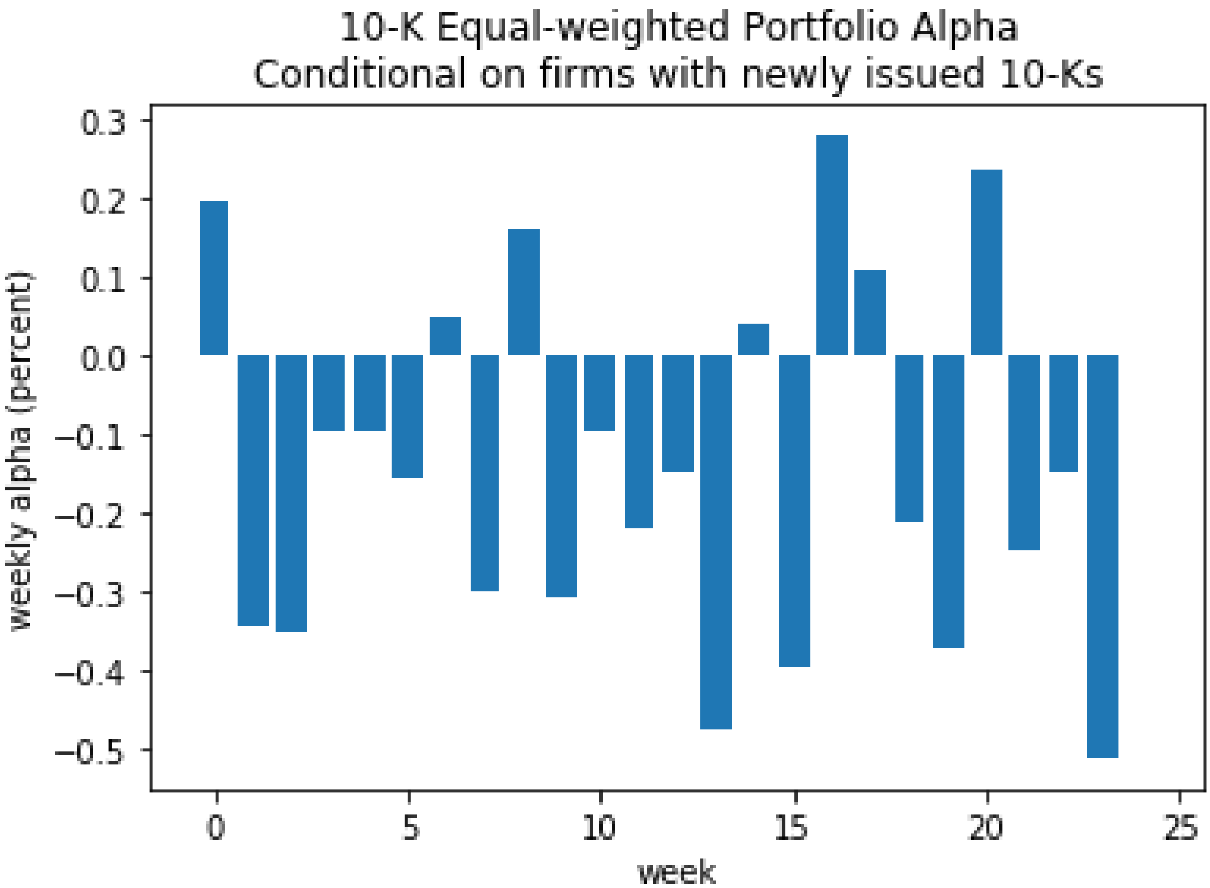


Figure 10

Demand for 8-K and Abnormal Return around Events

The figure studies the long/short portfolio of size-adjusted 8-K attention 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 size-adjusted 8-K attention and the cumulative abnormal return into 5-by-5 blocks. Conditional on each abnormal return quintile, I regress the long/short size-adjusted 8-K portfolio return on Fama French five factors and momentum 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.

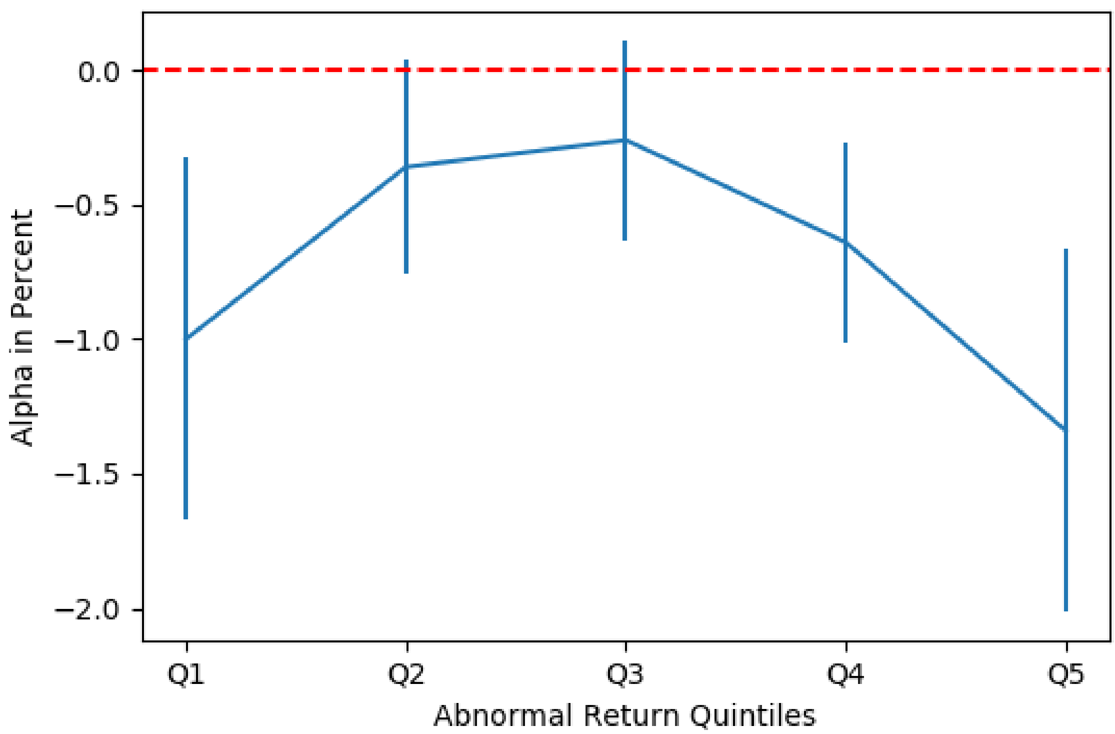


Table 1
Fama-Macbeth Regression on EDGAR Attention

The table shows results from Fama-Macbeth regressions of monthly individual stock returns on EDGAR views. The variable *log views*$_k$ is the natural log of human views of the firm for filing type k. Regressions include controls for other variables that are known to predict cross-section variation in returns. Independent variables are winsorized at one 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 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 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.

	(1) Ret	(2) Ret	(3) Ret	(4) Ret	(5) Ret
$log\ views^{full}_{all}$		0.183^{*}			
		(1.75)			
$log\ views^{full}_{10K}$			0.390^{***}	0.388^{***}	0.347^{***}
			(7.42)	(7.35)	(5.89)
$log\ views^{full}_{10Q}$			-0.0691	-0.0697	-0.0531
			(-1.15)	(-1.15)	(-1.09)
$log\ views^{full}_{8K}$			-0.120^{**}		
			(-2.23)		
$log\ views^{unscheduled}_{8K}$				-0.117^{**}	-0.174^{***}
				(-2.32)	(-3.10)
$log\ views^{scheduled}_{8K}$				0.0237	0.0128
				(0.70)	(0.30)
file 10K	0.222^{*}	0.149	-0.0693	-0.0729	-0.0797
	(1.87)	(1.22)	(-0.55)	(-0.58)	(-0.52)
file 8K	-0.0760^{***}	-0.120^{***}	-0.0584	-0.0585	-0.0417
	(-2.99)	(-4.79)	(-1.13)	(-1.20)	(-1.37)
file 10Q	-0.0725^{*}	-0.0873^{*}	-0.0531	-0.0441	-0.0471
	(-1.75)	(-1.90)	(-1.44)	(-1.49)	(-1.55)
Asset Growth	-0.723^{***}	-0.680^{***}	-0.622^{***}	-0.625^{***}	-0.527^{***}
	(-4.76)	(-4.70)	(-4.30)	(-4.32)	(-3.56)
log(BM)	0.134	0.108	0.0946	0.0937	0.0503
	(0.87)	(0.70)	(0.62)	(0.61)	(0.32)
log(ME)	-0.0696	-0.133^{*}	-0.178^{**}	-0.178^{**}	-0.0658
	(-1.39)	(-1.81)	(-2.48)	(-2.49)	(-0.90)
Operating Profit	0.0834^{**}	0.0654^{*}	0.0482	0.0477	0.0475
	(2.31)	(1.74)	(1.30)	(1.29)	(1.39)
$r_{1,0}$	-2.319^{***}	-2.397^{***}	-2.398^{***}	-2.394^{***}	-2.224^{***}
	(-3.49)	(-3.71)	(-3.72)	(-3.71)	(-3.07)
$r_{12,2}$	-0.608	-0.498	-0.496	-0.495	-0.415
	(-1.41)	(-1.30)	(-1.30)	(-1.29)	(-1.04)
Abnormal Trading Volume	0.141^{***}	0.132^{***}	0.135^{***}	0.136^{***}	0.128^{***}
	(4.08)	(3.90)	(4.02)	(4.02)	(3.35)
SUE	3.930^{***}	3.900^{***}	3.861^{***}	3.860^{***}	3.945^{***}
	(4.88)	(4.90)	(4.86)	(4.86)	(4.24)
Earning Drift	1.250^{***}	1.261^{***}	1.244^{***}	1.240^{***}	1.116^{***}
	(3.30)	(3.38)	(3.34)	(3.33)	(2.73)
Change in Google Trend					-0.108
					(-0.82)
Media Coverage					0.00314
					(0.69)
Constant	1.719^{**}	1.801^{**}	2.399^{***}	2.395^{***}	1.052
	(2.06)	(2.12)	(2.63)	(2.62)	(1.14)
N	502662	502662	502662	502662	347381
r2	0.0351	0.0385	0.0402	0.0403	0.0431
F	11.87	11.55	14.10	13.11	8.328

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

Table 2
Long/Short Portfolio by 10-K and 8-K Attention

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 run a cross-section regression of *log views*$_{10K}$ (*log views*$_{8K}$), the natural log of 10-K (8-K) views, on the natural log of lag firm size. The residuals of the regression can be interpreted as the level of 10-K (8-K) attention, after controlling for firm size. I then sort stocks by the size-adjusted log views 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 views, I only focus on the views on unscheduled filings.

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

Holding Months	Raw Return	$Alpha^{CAPM}$	$Alpha^{FF3}$	$Alpha^{FFC}$	$Alpha^{FF5+UMD}$	$Alpha^{8-factor}$
1	0.90***	0.65***	0.65***	0.73***	0.67***	0.58***
	(3.72)	(2.93)	(2.94)	(4.17)	(3.81)	(3.27)
3	0.62***	0.38*	0.38*	0.45***	0.38**	0.29*
	(2.81)	(1.88)	(1.9)	(2.9)	(2.43)	(1.82)
12	0.50***	0.28	0.27	0.33**	0.26*	0.19
	(2.67)	(1.64)	(1.63)	(2.33)	(1.83)	(1.29)

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

Holding Months	Raw Return	$Alpha^{CAPM}$	$Alpha^{FF3}$	$Alpha^{FFC}$	$Alpha^{FF5+UMD}$	$Alpha^{8-factor}$
1	-0.23	-0.56***	-0.56***	-0.51***	-0.56***	-0.47***
	(-1.18)	(-4.11)	(-4.09)	(-4.47)	(-4.78)	(-4.01)
3	-0.28*	-0.63***	-0.63***	-0.58***	-0.62***	-0.52***
	(-1.75)	(-4.91)	(-4.88)	(-5.48)	(-5.63)	(-4.83)
12	-0.27*	-0.63***	-0.63***	-0.58***	-0.61***	-0.52***
	(-1.83)	(-5.57)	(-5.56)	(-6.32)	(-6.4)	(-5.52)

Panel C: Factor Loadings of 10-K Portfolio

level	alpha	mktrf	smb	hml	umd	rmw	cma
Low	-0.01	0.835***	0.635***	0.185***	-0.015	-0.306***	-0.209***
	(-0.14)	(33.23)	(15.33)	(4.53)	(-0.73)	(-5.53)	(-3.1)
2	0.02	0.944***	0.734***	0.098***	-0.049***	-0.311***	-0.155***
	(0.34)	(54.89)	(25.88)	(3.52)	(-3.59)	(-8.19)	(-3.35)
3	0.12	0.977***	0.753***	0.085**	-0.138***	-0.218***	-0.096*
	(1.64)	(45.98)	(21.5)	(2.48)	(-8.11)	(-4.64)	(-1.67)
4	0.39***	1.017***	0.755***	0.057	-0.236***	-0.153**	0.025
	(4.12)	(36.47)	(16.43)	(1.25)	(-10.58)	(-2.48)	(0.33)
High	0.66***	1.001***	0.751***	-0.038	-0.437***	-0.262**	0.246*
	(3.79)	(19.79)	(9.01)	(-0.47)	(-10.81)	(-2.35)	(1.81)
H-L	0.67***	0.167***	0.117	-0.223***	-0.422***	0.044	0.456***
	(3.81)	(3.25)	(1.38)	(-2.68)	(-10.32)	(0.39)	(3.3)

Panel D: Factor Loadings of 8-K Portfolio

level	alpha	mktrf	smb	hml	umd	rmw	cma
Low	0.47***	0.746***	0.62***	0.116**	-0.103***	-0.365***	-0.071
	(4.44)	(24.04)	(12.12)	(2.31)	(-4.16)	(-5.33)	(-0.85)
2	0.34***	0.87***	0.737***	0.099**	-0.152***	-0.29***	-0.092
	(3.24)	(28.49)	(14.62)	(2.01)	(-6.22)	(-4.31)	(-1.12)
3	0.37***	0.979***	0.803***	0.01	-0.145***	-0.282***	-0.025
	(3.98)	(36.0)	(17.89)	(0.23)	(-6.67)	(-4.71)	(-0.34)
4	0.19*	1.02***	0.763***	0.049	-0.206***	-0.266***	0.003
	(1.83)	(33.23)	(15.07)	(0.98)	(-8.41)	(-3.95)	(0.03)
High	-0.08	1.097***	0.69***	0.025	-0.338***	-0.24***	-0.007
	(-0.65)	(31.17)	(11.88)	(0.43)	(-12.02)	(-3.1)	(-0.07)
H-L	-0.56***	0.353***	0.075	-0.095*	-0.233***	0.127*	0.06
	(-4.78)	(10.29)	(1.33)	(-1.7)	(-8.49)	(1.68)	(0.65)

t statistics in parentheses

Table 3
Double Sort on 10-K and 8-K Attention
The table shows monthly alphas of portfolios double sorted by size-adjusted 10-K and 8-K attention. For each portfolio, I regress portfolio return on Fama French five factors and momentum factor. For 8-K views, I only focus on the views on unscheduled filings.

8-K / 10-K	Low	2	3	4	High	H-L
Low	0.04	0.05	0.20	0.40**	0.67***	0.63**
	(0.43)	(0.55)	(1.52)	(2.37)	(2.67)	(2.48)
2	-0.18	-0.03	0.12	0.23	0.68***	0.86***
	(-1.52)	(-0.29)	(1.16)	(1.6)	(3.12)	(3.49)
3	-0.20	-0.06	0.06	0.34**	1.00***	1.21***
	(-1.6)	(-0.62)	(0.57)	(2.59)	(4.57)	(4.61)
4	-0.16	-0.17	-0.11	0.22*	0.44**	0.60**
	(-0.91)	(-1.59)	(-1.07)	(1.7)	(2.31)	(2.37)
High	-1.16***	-0.45***	-0.18	-0.08	-0.02	1.14***
	(-4.8)	(-2.85)	(-1.52)	(-0.64)	(-0.12)	(4.05)
H-L	-1.21***	-0.51***	-0.39**	-0.49***	-0.72***	
	(-5.03)	(-2.89)	(-2.35)	(-2.84)	(-2.94)	

t statistics in parentheses

Table 4

Weekly Regression of Stock Returns on EDGAR Attention

The table shows results from regressions of weekly individual stock returns on EDGAR views. 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 stock returns in basis points. $views_t^k$ is the cumulative views 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 Ravenpacks 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 Trend daily index in any day of the week is above its 90 percentile in the past month. Time fixed effects are included, and standard errors are clustered by week.

	(1)	(2)	(3)	(4)
	ret_t	ret_{t+1}	ret_t	ret_{t+1}
$\log(views_t^{10K})$	14.11***	11.68***	6.809***	4.523***
	(8.58)	(7.69)	(4.27)	(2.95)
$\log(views_t^{8K})$	-0.405	-2.046	-1.955	-1.691
	(-0.26)	(-1.53)	(-0.88)	(-1.42)
$Filing\ 10K_t \times \log(views_t^{10K})$	-1.504	5.867	-2.450	-0.0753
	(-0.23)	(1.01)	(-0.39)	(-0.01)
$Filing\ 8K_t \times \log(views_t^{8K})$	11.54***	-3.262*	16.09***	-4.804*
	(4.79)	(-1.82)	(3.81)	(-1.79)
$Filing\ 10K_t$	-28.96	-25.01	0.292	0.375
	(-1.18)	(-1.27)	(0.01)	(0.01)
$Filing\ 8K_t$	-5.198	8.463*	-32.67**	11.28
	(-0.86)	(1.86)	(-2.51)	(1.16)
$Media\ Coverage_t$	34.38***	5.645***	24.83***	2.975
	(16.84)	(3.29)	(7.33)	(0.92)
$Earning\ Release_t$	36.70***	14.36***	1.390	17.51***
	(6.20)	(3.39)	(0.17)	(3.18)
AIA_t			51.67***	0.999
			(11.19)	(0.34)
$DADSVI_t$			21.89***	2.301
			(11.05)	(1.32)
lag returns	Yes	Yes	Yes	Yes
firm controls	Yes	Yes	Yes	Yes
week fe	Yes	Yes	Yes	Yes
N	2308554	2305463	529874	528806
r2	0.115	0.113	0.154	0.172
F	63.31	21.04	27.70	3.784

t statistics in parentheses

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

Table 5
Panel Regression of Private Information

The table shows the monthly panel regression of next-month information asymmetry proxy on current month investor attention to filings. The dependent variables are the next-month quote spread (Corwin and Schultz (2012)) and Amihud (2002) measure. Independent variables include the log views of filings, current information asymmetry measure, and firm characteristics shown in Table 1. Time and firm fixed effects are included. Standard errors are two-way clustered by time and firm.

	(lead asy proxy)	
	(1)	(2)
	spread	amihud
log views$_{10K}$	-0.00742	-0.113
	(-1.48)	(-1.06)
log views$_{8K}$	-0.00626*	-0.172***
	(-1.82)	(-2.69)
log views$_{10Q}$	0.00698	0.0641
	(1.54)	(0.71)
asy proxy	0.283***	0.441***
	(14.71)	(6.16)
Asset Growth	-0.0197***	0.0501
	(-2.87)	(0.55)
log(BM)	0.0145	-0.198
	(1.52)	(-0.76)
log(ME)	-0.213***	-1.843***
	(-15.76)	(-5.68)
Operating Profit	-0.00167	-0.0675
	(-0.85)	(-1.44)
$r_{1,0}$	0.0386	-0.203
	(1.42)	(-0.19)
$r_{12,2}$	-0.0384***	0.0766
	(-5.73)	(0.84)
Abnormal Trading Volume	-0.00438***	-0.0614***
	(-3.36)	(-3.84)
SUE	-0.0466	-0.740**
	(-1.54)	(-2.13)
Earning Drift	0.0214	0.556**
	(1.09)	(2.31)
file 10K	0.0116	0.00526
	(0.90)	(0.04)
file 8K	0.0268***	0.142***
	(2.93)	(2.69)
file 10Q	0.0172***	0.169**
	(2.94)	(2.15)
Time and Firm FE	Yes	Yes
N	449278	448647
adjusted R^2	0.389	0.352
F	56.97	21.78

t statistics in parentheses

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

Table 6
8-K Attention and Information Asymmetry

The table shows monthly alphas of portfolios sorted by size-adjusted 8-K views and information asymmetry. I use Amihud illiquidity measure and previous quarter earning forecast dispersion to measure ex-ante information asymmetry. For each portfolio, I regress portfolio return on Fama French five factors and momentum factor. For 8-K views, I only focus on the views 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	0.09	0.01	0.10	0.10	-0.03	-0.12
	(1.17)	(0.12)	(1.33)	(1.32)	(-0.32)	(-1.08)
2	0.21**	0.12	0.21**	0.26**	-0.12	-0.34*
	(2.22)	(1.11)	(2.01)	(2.22)	(-0.69)	(-1.71)
High	0.81***	0.72***	0.67**	0.45***	0.04	-0.72**
	(4.2)	(3.3)	(2.58)	(3.46)	(0.1)	(-2.31)
H-L	0.73***	0.71***	0.57**	0.35**	0.07	-0.6**
	(3.36)	(2.94)	(2.11)	(2.13)	(0.18)	(-2.12)

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	0.32***	0.08	0.21**	0.22**	0.18**	-0.13
	(3.68)	(0.97)	(2.6)	(2.57)	(2.07)	(-1.16)
2	0.15*	0.19**	0.20**	0.06	-0.09	-0.24*
	(1.66)	(2.02)	(2.05)	(0.58)	(-0.74)	(-1.69)
High	0.17	-0.09	-0.14	-0.15	-0.54***	-0.71***
	(1.29)	(-0.61)	(-0.88)	(-0.93)	(-2.65)	(-3.27)
H-L	-0.15	-0.18	-0.34*	-0.37*	-0.72***	-0.58***
	(-0.87)	(-0.95)	(-1.85)	(-1.91)	(-3.2)	(-2.85)

t statistics in parentheses

Table 7
8-K Attention, Distance to Headquarters, and Frequent Viewers
The table shows monthly alphas of equal-weighted portfolios sorted by size-adjusted 8-K views, conditional on geographical distance distribution to headquarters and frequent 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 frequent view if the IP address visited any firm filings in the past three months. Frequent visitor ratio is the ratio between the numbers of frequent and infrequent views. For each stock at each month, I first sort stocks by geographical distance (frequent visitor ratio) into terciles. Conditional on each tercile, I then sort stocks by size-adjusted 8-K views into quintiles. For each portfolio, I regress portfolio return on Fama French five factors and momentum factor, and report the alphas.

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

distance/8-K views	Low	2	3	4	High	H-L
Low	0.55***	0.37***	0.31**	0.12	-0.03	-0.48***
	(4.67)	(3.48)	(2.44)	(1.28)	(-0.24)	(-3.61)
2	0.54***	0.48***	0.20	0.06	-0.11	-0.65***
	(4.27)	(3.7)	(1.55)	(0.43)	(-0.75)	(-4.21)
High	0.44***	0.17	0.47***	0.41**	0.26	-0.18*
	(3.11)	(1.2)	(3.71)	(2.55)	(0.96)	(-1.73)
H-L	-0.12	-0.21	0.16	0.28*	0.29	0.31**
	(-0.95)	(-1.56)	(1.11)	(1.69)	(1.38)	(-2.32)

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

$freq^{8K}$/8-K views	Low	2	3	4	High	H-L
Low	0.44***	0.38***	0.23*	0.45***	0.21	-0.25
	(3.17)	(3.15)	(1.77)	(3.95)	(1.64)	(-1.62)
2	0.47***	0.45***	0.30**	0.13	-0.02	-0.49***
	(3.57)	(3.63)	(2.4)	(0.94)	(-1.43)	(-2.95)
High	0.35***	0.28**	0.19	-0.11	-0.07	-0.42***
	(3.19)	(2.03)	(1.64)	(-0.82)	(-0.41)	(-2.8)
H-L	-0.06	-0.08	-0.01	-0.53***	-0.27*	-0.17*
	(-0.51)	(-0.55)	(-0.11)	(-4.02)	(-1.81)	(-1.77)

t statistics in parentheses

Table 8
8-K Attention and Information Content
The table shows monthly alphas of portfolios sorted by size-adjusted 8-K views and cumulative abnormal returns around filing and event date of unscheduled 8-K filings. 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 size-adjusted 8-K attention and the cumulative abnormal return into 5-by-3 blocks. Conditional on each abnormal return tercile, I regress the long/short size-adjusted 8-K portfolio return on Fama French five factors and momentum 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

abret/views	Low	2	3	4	High	H-L
Low	0.36**	0.22	-0.05	-0.01	-0.38*	-0.76***
	(2.05)	(1.07)	(-0.23)	(-0.04)	(-1.77)	(-3.19)
2	0.32**	0.41***	0.18	0.36***	0.11	-0.23*
	(2.5)	(3.18)	(1.57)	(3.06)	(-0.11)	(-1.78)
High	0.65***	0.53***	0.36*	0.13	-0.33	-0.99***
	(3.66)	(3.07)	(1.74)	(0.61)	(-1.64)	(-4.39)
H-L	0.28	0.31	0.41**	0.14	0.05	-0.23
	(1.43)	(1.28)	(2.03)	(0.5)	(0.21)	(-0.72)

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

unexpected abret/views	Low	2	3	4	High	H-L
Low	0.38**	0.20	-0.04	-0.05	-0.36*	-0.76***
	(2.2)	(0.96)	(-0.21)	(-0.21)	(-1.68)	(-3.15)
2	0.28**	0.40***	0.15	0.43***	0.02	-0.27
	(2.22)	(3.06)	(1.37)	(3.73)	(0.17)	(-1.65)
High	0.69***	0.48***	0.36*	0.11	-0.36*	-1.07***
	(3.94)	(2.8)	(1.82)	(0.53)	(-1.77)	(-4.69)
H-L	0.31	0.29	0.40*	0.16	0.00	-0.31
	(1.62)	(1.15)	(1.97)	(0.58)	(0.0)	(-1.45)

t statistics in parentheses

Table 9
10-K Attention and Attention-Grabbing

The table shows monthly alphas of portfolios sorted by size-adjusted 10-K views and attention-grabbing measure. I use abnormal trading volume and maximum daily absolute return to measure attention-grabbing. Abnormal trading volume is the difference between monthly trading volume and past 12-month average trading volume, scaled by the standard deviation of past 12-month trading volume. For each portfolio, I regress portfolio return on Fama French five factors and momentum factor. For 8-K views, I only focus on the views on unscheduled filings.

Panel A: Double Sort by Size-adjusted 10-K Views and Maximum Return

Max Return/Views	Low	2	3	4	High	H-L
Low	0.29***	0.31***	0.38***	0.35***	0.60***	0.31**
	(2.9)	(3.84)	(4.58)	(4.09)	(5.86)	(2.31)
2	-0.03	-0.07	0.16*	0.33***	0.52***	0.55***
	(-0.26)	(-0.83)	(1.91)	(3.55)	(3.12)	(2.84)
High	-0.36**	-0.32**	-0.02	0.45**	0.51	0.87***
	(-2.17)	(-2.22)	(-0.11)	(2.16)	(1.63)	(3.14)
H-L	-0.65***	-0.63***	-0.40**	0.10	-0.09	0.56**
	(-3.3)	(-3.41)	(-2.18)	(0.49)	(-0.31)	(1.99)

Panel B: Double Sort by Size-adjusted 10-K Views and Abnormal Trading Volume

Abnormal Trading Volume/Views	Low	2	3	4	High	H-L
Low	-0.31***	-0.23**	-0.12	-0.19	0.06	0.38*
	(-2.7)	(-2.29)	(-1.05)	(-1.33)	(0.31)	(1.87)
2	0.01	0.03	0.27***	0.46***	0.63***	0.62***
	(0.07)	(0.35)	(2.76)	(4.04)	(3.34)	(3.31)
High	0.21*	0.09	0.31***	0.72***	1.16***	0.95***
	(1.85)	(0.9)	(2.91)	(5.5)	(5.49)	(4.16)
H-L	0.52***	0.32**	0.43***	0.91***	1.10***	0.57**
	(3.41)	(2.22)	(2.91)	(5.57)	(5.3)	(2.32)

t statistics in parentheses

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 views. $views_{10Q}$ is the number of 10-Q filing views. $views_{8K}$ is the number of 8-K filing views. 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

10-K Attention, Distance to Headquarters, and Frequent Viewers

The table shows monthly alphas of equal-weighted portfolios sorted by size-adjusted 10-K views, conditional on geographical distance distribution to headquarters and frequent 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 frequent view if the IP address visited any firm filings in the past three months. Frequent visitor ratio is the ratio between the numbers of frequent and infrequent views. For each stock at each month, I first sort stocks by geographical distance (frequent visitor ratio) into terciles. Conditional on each tercile, I then sort stocks by size-adjusted 10-K views into quintiles. For each portfolio, I regress portfolio return on Fama French five factors and momentum factor, and report the alphas.

Panel A: Double Sort by Size-adjusted 10-K Views and Distance

distance/10-K views	Low	2	3	4	High	H-L
Low	-0.10	-0.04	0.09	0.09	0.23*	0.33*
	(-0.93)	(-0.43)	(0.92)	(0.78)	(1.69)	(1.77)
2	-0.22**	-0.17*	-0.09	0.23**	0.28*	0.49**
	(-2.29)	(-1.94)	(-0.97)	(2.12)	(1.7)	(2.59)
High	-0.05	0.02	0.11	0.38***	0.63***	0.68***
	(-0.37)	(0.19)	(1.08)	(2.82)	(2.91)	(3.0)
H-L	0.05	0.04	0.02	0.28	0.39*	0.35*
	(0.34)	(0.34)	(0.13)	(1.18)	(1.77)	(1.78)

Panel B: Double Sort by Size-adjusted 10-K Views and 10-K Freq Ratio

$freq^{10K}$/10-K views	Low	2	3	4	High	H-L
Low	0.00	-0.08	0.11	0.26**	0.63**	0.63**
	(0.02)	(-0.81)	(1.11)	(2.21)	(3.38)	(3.14)
2	-0.23***	-0.03	0.08	0.14	0.27	0.50**
	(-2.65)	(-0.41)	(0.91)	(1.22)	(1.52)	(2.5)
High	-0.09	-0.09	-0.00	0.13	0.40*	0.49*
	(-0.85)	(-1.08)	(-0.04)	(1.02)	(1.67)	(1.83)
H-L	-0.10	-0.02	-0.11	-0.16	-0.26	-0.14
	(-0.95)	(-0.19)	(-0.99)	(-1.15)	(-1.35)	(-1.25)

t statistics in parentheses

Table A3
EDGAR 8-K Filing Counts

The table shows the number of 8-K filings by section for all firms in the sample from 1994 to 2016. A filing can be categorized into multiple sections/items.

Item Code	Description	Section Count	Item Count
Section 1	Registrant's Business and Operations	210636	
Item 1.01	Entry into a Material Definitive Agreement		198726
Item 1.02	Termination of a Material Definitive Agreement		18355
Item 1.03	Bankruptcy or Receivership		4320
Item 1.04	Mine Safety		182
Section 2	Financial Information	369770	
Item 2.01	Completion of Acquisition or Disposition of Assets		43560
Item 2.02	Results of Operations and Financial Condition		262011
Item 2.03	Creation of a Direct Financial Obligation		61905
Item 2.04	Triggering Events That Accelerate or Increase a Direct Financial Obligation		3701
Item 2.05	Costs Associated with Exit or Disposal Activities		5518
Item 2.06	Material Impairments		3119
Section 3	Securities and Trading Markets	66905	
Item 3.01	Notice of Delisting or Failure to Satisfy a Continued Listing Rule or Standard		14068
Item 3.02	Unregistered Sales of Equity Securities		45151
Item 3.03	Material Modification to Rights of Security Holders		12249
Section 4	Matters Related to Accountants and Financial Statements	31476	
Item 4.01	Changes in Registrant's Certifying Accountant		25642
Item 4.02	Non-Reliance on Previously Issued Financial Statements or a Related Audit Report or Completed Interim Review		5968
Section 5	Corporate Governance and Management	262237	
Item 5.01	Changes in Control of Registrant		16682
Item 5.02	Departure/Election/Appointment of Directors or Officers; Compensatory Arrangements of Certain Officers		194313
Item 5.03	Amendments to Articles of Incorporation or Bylaws; Change in Fiscal Year		37421
Item 5.04	Temporary Suspension of Trading Under Registrant's Employee Benefit Plans		1116
Item 5.05	Amendment to Registrant's Code of Ethics, or Waiver of a Provision of the Code of Ethics		1932
Item 5.06	Change in Shell Company Status		1719
Item 5.07	Submission of Matters to a Vote of Security Holders		38456
Item 5.08	Shareholder Director Nominations		340
Section 6	Asset-Backed Securities	1052	
Item 6.01	ABS Informational and Computational Material		199
Item 6.02	Change of Servicer or Trustee		660
Item 6.04	Failure to Make a Required Distribution		43
Item 6.05	Securities Act Updating Disclosure		105
Section 7	Regulation FD Disclosure	207540	
Section 8	Other Events	421676	
Section 9	Financial Statements and Exhibits	968550	

Table A4

EDGAR 8-K Views by Sections

The table shows the total number of 8-K views by section for all firms in the sample from 2003 to 2016. If a filing is categorized into multiple sections/items, a single view of the filing is counted into multiple sections/items.

Item Code	Description	Section Count	Item Count
Section 1	Registrant's Business and Operations	234569	
Item 1.01	Entry into a Material Definitive Agreement		228676
Item 1.02	Termination of a Material Definitive Agreement		21985
Item 1.03	Bankruptcy or Receivership		1568
Item 1.04	Mine Safety		199
Section 2	Financial Information	410335	
Item 2.01	Completion of Acquisition or Disposition of Assets		35479
Item 2.02	Results of Operations and Financial Condition		308484
Item 2.03	Creation of a Direct Financial Obligation		69664
Item 2.04	Triggering Events That Accelerate or Increase a Direct Financial Obligation		2552
Item 2.05	Costs Associated with Exit or Disposal Activities		11464
Item 2.06	Material Impairments		5017
Section 3	Securities and Trading Markets	56773	
Item 3.01	Notice of Delisting or Failure to Satisfy a Continued Listing Rule or Standard		8991
Item 3.02	Unregistered Sales of Equity Securities		35529
Item 3.03	Material Modification to Rights of Security Holders		18957
Section 4	Matters Related to Accountants and Financial Statements	12015	
Item 4.01	Changes in Registrant's Certifying Accountant		7407
Item 4.02	Non-Reliance on Previously Issued Financial Statements or a Related Audit Report or Completed Interim Review		4714
Section 5	Corporate Governance and Management	264116	
Item 5.01	Changes in Control of Registrant		3399
Item 5.02	Departure/Election/Appointment of Directors or Officers; Compensatory Arrangements of Certain Officers		207761
Item 5.03	Amendments to Articles of Incorporation or Bylaws; Change in Fiscal Year		50110
Item 5.04	Temporary Suspension of Trading Under Registrant's Employee Benefit Plans		1412
Item 5.05	Amendment to Registrant's Code of Ethics, or Waiver of a Provision of the Code of Ethics		3293
Item 5.06	Change in Shell Company Status		613
Item 5.07	Submission of Matters to a Vote of Security Holders		39600
Item 5.08	Shareholder Director Nominations		167
Section 6	Asset-Backed Securities	25	
Item 6.01	ABS Informational and Computational Material		0
Item 6.02	Change of Servicer or Trustee		7
Item 6.04	Failure to Make a Required Distribution		0
Item 6.05	Securities Act Updating Disclosure		16
Section 7	Regulation FD Disclosure	246470	
Section 8	Other Events	320659	
Section 9	Financial Statements and Exhibits	938448	

Table A5
Investor Attention Portfolios Conditional on Media Coverage
The table shows monthly equal-weighted alphas of portfolios sorted by size-adjusted views and high media coverage dummy, which is equal to one if the number of news is higher than the past 12-month median.

Panel A: Double Sort by Size-adjusted 10-K Views and Media Coverage

Media/10-K views	Low	2	3	4	High	H-L
Low	-0.11	-0.14**	0.00	0.16	0.53***	0.64***
	(-1.27)	(-2.04)	(0.01)	(1.54)	(2.89)	(3.33)
High	-0.09	-0.19**	-0.00	0.20*	0.17	0.26
	(-0.92)	(-2.12)	(-0.05)	(1.79)	(0.93)	(1.27)
H-L	0.02	-0.04	-0.01	0.04	-0.36**	-0.38
	(0.19)	(-0.39)	(-0.05)	(0.31)	(-2.58)	(-1.52)

Panel B: Double Sort by Size-adjusted 8-K Views and Media Coverage

Media/8-K views	Low	2	3	4	High	H-L
Low	0.55***	0.35***	0.42***	0.26**	-0.01	-0.56***
	(4.42)	(2.9)	(3.47)	(2.25)	(-0.09)	(-4.02)
High	0.37***	0.27**	0.24**	0.03	-0.23	-0.61***
	(3.19)	(2.36)	(2.34)	(0.25)	(-1.6)	(-3.8)
H-L	-0.18	-0.09	-0.18	-0.23**	-0.23*	-0.05
	(-1.39)	(-0.75)	(-1.39)	(-2.08)	(-1.94)	(-0.33)

t statistics in parentheses