Beyond attention: the causal effect of media on information production *

Bruce Yuan Li †

First version: May 30, 2018; Current version: November 15, 2018

Latest version can be found here

Abstract

This paper shows that media coverage causes institutional investors to gather more information and analysts to produce more earning forecasts. I exploit random variation in the visual salience of corporate press releases to financial journalists to proxy for media coverage. Doubling the amount of media coverage increases the number of EDGAR searches by 37% and the number of analyst forecasts by 78% in a two-day period. The evidence is consistent with the theories of rational attention allocation: investors allocate resources to media-covered events as the ex-ante variances of returns are higher. Analysts cater to the increased information demand by responding to media-reported events. The results suggest that different information channels do interact, and financial media complements other channels.

Keywords: media, analyst, EDGAR search, institutional investor
JEL Codes: G12, G14

*I am indebted to my advisor, Gerard Hoberg, and my committee members, Kenneth Ahern, Oguz Ozbas, Yongxiang Wang, and TJ Wong, for their invaluable help and guidance. I also thank Heitor Almeida, William Cong, Itay Goldstein, Zhiguo He, Christopher Jones, E Han Kim, John Matsusaka, Sheridan Titman, Song Ma, Emily Nix, Christopher Parsons, Andrei Shleifer, and Ben Zhang, participants at the seminars at the University of Southern California for helpful comments and suggestions. All remaining errors are my own.

†PhD Student at Marshall School of Business, USC. Email: yli268@marshall.usc.edu

A well-functioning financial market requires information to flow efficiently between corporations and investors. Such information flows are often facilitated by a variety of information channels, including financial media, analysts, and investors' information acquisition. While the literature has studied how each of these channels individually affects the information environment and stock trading, little is known about whether and how these channels interact[1]. Uncovering the potential interactions is important to understand the process through which stock prices incorporate information. This paper's main contribution is to show a complementary relationship between financial media and other information channels: media coverage on Dow Jones Newswire leads to more information acquisition of institutional investors and earnings forecasts of analysts.

Whether and how financial media affect other information channels is ex-ante unclear. While many papers document that media coverage affects stock trading, they conclude that retail investors, who lack access to new information and are inattentive, most likely drive the results. It is thus possible that media coverage is just a sideshow for more sophisticated market participants. However, sophisticated investors, who are informationally resourceful and follow the market closely, could also respond to media coverage for strategic reasons. One possibility is that they might allocate their research to corporate releases that did not receive media coverage, thus to profit from delayed market responses (Cohen and Frazzini, 2008; Menzly and Ozbas, 2010). Alternatively, they might focus their information acquisition on media-covered events, where the ex-ante variance of returns is higher, leading to higher reward for having more precise signals (Kacperczyk, Van Nieuwerburgh, and Veldkamp, 2016). The information demand from these institutional investors would further affect the coverage choices by analysts, especially when the analysts have stronger

[1]Tetlock (2014) provides an excellent literature review for papers that study the effect of media on the financial market. Some more recent work includes Blankespoor, deHaan, and Zhu (2018); Fedyk (2018) among others. Papers that study the effect of analysts on information include Brennan, Jegadeesh, and Swaminathan (1993); Brennan and Subrahmanyam (1995); Womack (1996); Hong, Lim, and Stein (2000); Gleason and Lee (2003); Ivković and Jegadeesh (2004); Livnat and Mendenhall (2006); Kelly and Ljungqvist (2012); Bradley et al. (2014) among others. Empirical work on how investors actively acquire or search for information includes Da, Engelberg, and Gao (2011); Drake, Guest, and Twedt (2014); Lee, Ma, and Wang (2015); Loughran and McDonald (2017); Chen et al. (2017) among others.

demand-catering incentives. Therefore, whether and how media affects other information channels is ultimately an empirical question.

To empirically study the relationships, this paper tests when a firm issues a new press release, whether media coverage of the event by Dow Jones Newswire changes the information acquisition of investors and the number of analysts issuing earnings forecasts for the firm. A significant challenge for answering these questions, however, is that endogenous event- or firm-characteristics would affect the coverage decisions of both media and other information channels. For example, an earning surprise would attract media coverage, but at the same time also induce analysts to update their earning forecasts, thus creating a positive correlation between the responses from different information channels.

This paper tackles the above identification challenge by exploiting random shocks that affect the amount of media coverage. The shocks arise because of the unique way that wire journalists process information and produce news. For a newswire journalist, the typical workflow is to monitor real-time press release feeds, select newsworthy events, and quickly replay the main points to their subscribers. They need to perform the tasks quickly as different wire media compete over speed. Moreover, wire journalists generally do not specialize in industries. Instead, they cover events for the whole market[2]. A particular challenge for these wire journalists is that press releases do not arrive at a constant rate; they instead cluster at specific times within a day. Thus when large clusters of firms issue press releases at the same time, the amount of information faced by wire journalists will be challenging to process. However, the same issue is less relevant for other more focused market participants.

Key to the identification is that in these busy times, some press releases receive more media coverage than their adjacent peers because they are exogenously more visually salient to journalists. Most real-time feed systems, which the journalists are monitoring, position new content at the top of the interface, pushing old content down and to later

[2]We can see this from the job description of a Dow Jones journalist: "The desk covers all stocks from the largest FTSE100s to the smallest AIM companies across the whole range of industries and subject matters"

pages. Therefore for each press release, the duration for which it stays visually salient to journalists is determined by the speed of new releases pushing it out of the prominent position of the user interface (e.g., the first page of the screen). Importantly, such "on-screen" time can vary a lot for press releases within the same busy cluster. For a press release queued near the beginning of the cluster, it is followed by the whole cluster; thus its on-screen time is short. In comparison, a press release queued near the end of the cluster has much fewer press releases after it, and it will stay on the screen for longer. This paper uses a tight time window (first 10 seconds of an hour) to define clusters, thus the queuing orders in this short-interval queue are likely free of the concern for strategic timing, as a firm would not strategically release its press release at 4:00:01 rather than 4:00:09. On top of that, the number of press releases from other (most likely unrelated) firms will shift the size of the cluster and further affect the on-screen time, and these press releases from unrelated firms are unlikely to be correlated with the press release timing of any single firm.

This paper uses the on-screen time as an instrumental variable to media coverage. In testing the relevance condition, I find that a shorter on-screen time leads to fewer media coverage. Using all the corporate press releases issued in the first 10 seconds of an hour, which is the busiest time within an hour, I find that for a press release, if the number of following releases in the next 30 seconds doubles, thereby its on-screen time shortens, the amount of news coverage drops by 9.3%. Stated in percentage points, the probability of coverage decreases from 48% to 43% on the press release day.

The validity of the instrument requires the exclusion restriction condition to hold: the variation of on-screen time affects analysts and investors only through media. I find evidence supporting the exclusion restriction condition. First, omitted variables likely cannot explain the effect of on-screen time on media coverage. Using covariate balancing tests, I first show that the proxy of on-screen time is uncorrelated with a wide array of firm- and event-characteristics, while the same characteristics significantly predict the amount of media coverage. I further guard against omitted variables with a rigid set of fixed effects, and the effect remains highly robust. Finally, falsification tests and subsample tests further suggest that it is the journalists' limited cognitive capacity, rather than unobserved variables or economic properties of the press releases that drives the effect. Second, the effect is most likely unique to wire journalists and does not directly apply to investors or analysts. The on-screen time matters for wire journalists because their job is to monitor the event stream for the *whole market* in *real-time*. In comparison, investors and analysts often actively search or set up alarms to get information. When they do monitor a real-time event stream, they may also apply filters to focus on the specific industries they specialize in. I find the on-screen time is uncorrelated with the amount of new information from the same industry. Moreover, I find that when the shorter on-screen time is caused by press releases from private firms, which likely only impact wire journalists but not analysts and investors, the same results obtain.

With the on-screen time as an instrumental variable, I first find that media coverage significantly increases investors' information acquisition, which is measured by the number of internet requests on the SEC EDGAR system. Using the SEC EDGAR log database, I find that press releases with longer on-screen time receive more EDGAR searches in the next two days. Furthermore, significant results only exist for searches made by humans and not for requests made by web crawlers, further validating the exclusion restriction condition.

The positive effects likely go beyond the simple effects on attention because investors who are more attentive show similar or even stronger reactions. First, investors who have requested the firm's filings in the previous month show similar positive responses to media coverage. Since these investors already follow the same firm, they should suffer less from the "search problem" that is associated with limited attention (Barber and Odean, 2008). Second, I identify a subset of IP addresses which belong to financial institutions. While these institutional investors are supposed to be more attentive, their EDGAR searches react even stronger to the media coverage.

I also find that analysts issue more earnings forecasts when the on-screen time of a press release increases. Such results are surprising and puzzling given the conventional wisdom that analysts are ultra-informed. The body of evidence in this paper indicates that the results are most likely driven by resource-constrained analysts catering to the increasing demand from their institutional clients, whose demand for information on media-covered stocks is revealed to be stronger given the earlier-mentioned findings for EDGAR searches. Two key subsample results further support this explanation. First, the results are much stronger for firms with higher institutional holdings or more institutional investors. Second, the results are mostly from analysts who cover an above-median number of firms and thus are more likely to be resource-constrained. In short, resource-constrained analysts crowd-source the demand for information from their institutional clients and then cater to it, shifting their information production toward news-covered events. Indeed, as Brown et al. (2015) concludes from a survey of 182 analysts, "Demand from their clients is analysts' most important motivation for making profitable stock recommendations and their second most important motivation for issuing accurate earnings forecasts".

Altogether, I find that institutional investors and stock analysts, who are among the most sophisticated market participants, increase their information production when media coverage increases. The results are consistent with the framework of Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016), where skilled investors first decide how to allocate their attention to different stocks. One prediction from the model is that skilled investors allocate more resources to assets whose payoffs have high variances. Existing empirical evidence shows that media coverage increases the volatility of expected returns (Peress, 2014; Blankespoor, deHaan, and Zhu, 2018), changes the investor base (Barber and Odean, 2008), and possibly increases mispricing (Hillert, Jacobs, and Müller, 2014; Ahern and Sosyura, 2015). As a result, sophisticated investors may interpret media coverage as a reliable signal which suggests a higher reward to their information production. These investors also request more information production from analysts, who would cater to such demand by also shifting their information production to media-covered events. Importantly, the existence of resource constraints does not indicate irrationality, as it could suggest that analysts build their business models through a series of tradeoffs. An analyst with generalist skills can increase profits by covering more firms, but the cost of doing so is some sacrifice in their ability to be timely. Hence, such an analyst with a broader network of institutional clients can crowd-source where demand for updates is most intense. In contrast, an analyst with more specialized skills and a thinner network of institutional clientele can optimize timeliness, which can build reputation and profits, but this rationally comes at the cost of covering fewer firms.

The market outcomes I document are consistent with the rational allocation of investor attention and expenditures on information production. I first find that media coverage increases the trading volume for the next 5 trading days after the press releases, similar to the results in previous research (Engelberg and Parsons, 2011; Peress, 2014; Blankespoor, deHaan, and Zhu, 2018; Fedyk, 2018). What's new in this paper on is that media coverage also significantly widens the effective spread on the event day, consistent with a relative increase in informed traders. In the next two days, while media coverage does not affect absolute returns, it significantly increases the intraday price ranges. Peress (2014) attributes similar effects to "less price-sensitive traders who transact at less favorable prices". This lead-lag relationship in trading between informed and uninformed traders is also consistent with the findings in Ben-Rephael, Da, and Israelsen (2017), who document a similar lead-lag relationship in information searches. Overall, I find media neither improves nor deteriorates price efficiency, measured by the delayed response ratio from Dellavigna and Pollet (2009). The result echoes Blankespoor, deHaan, and Zhu (2018) who also find no effect of media on price efficiency. Collectively, the results are consistent with a tug of war between two types of investors, one attracted by media due to attention, and the other that consciously trades more media-covered firms, profiting by trading against these potentially uninformed traders.

1 Related research

1.1 Media and information efficiency

Existing evidence shows that news coverage by mass media influences the trading of both retail and institutional investors. However, in many cases such trade appears to be mispricing[3]. For example, Hillert, Jacobs, and Müller (2014) find that firms covered by newspapers show stronger momentum in the short-run but experience reversals in the longrun, concluding that media coverage "exacerbates investor bias". Fang, Peress, and Zheng (2014) find that some mutual fund managers persistently buy more media-covered stocks, and they underperform other fund managers who show lower propensities to buy media stocks. The situation becomes worse as mass media often has biased incentives that can impair its accuracy. Gurun and Butler (2012) show that local media uses less negative words when reporting local firms due to the advertisement income from these firms. Ahern and Sosyura (2015) show that media has strong incentives to report sensational news and investors tend to overreact to the merger rumors.

This line of research suggests that mispricing will likely arise when limited-attention investors over rely on mass media for information. Yet such conclusion does not necessarily extend to other professional financial media that serves sophisticated investors. Media like Dow Jones Newswire or Bloomberg builds their businesses by providing accurate information to investors in a timely fashion. The journalists in these agencies closely monitor a variety of information sources, select the relevant events, and quickly report to their users. If these agencies are more efficient at selecting important and filtering out irrelevant ones, then delegating the information selection task to these media is a natural division of labor for investors. However, we don't have much evidence on whether there are any biases or inefficiency in these media as well.

This paper fills the gap by documenting a novel setting where inefficiency could arise

[3] I refer readers to Tetlock (2014) for an excellent review of the literature of media.

due to the clustering of press releases and the limited cognitive capacity of wire journalists. During these busy times, media coverage decision is affected by how visually salient a press release is to the journalists. This paper provides extensive evidence to show that the visual salience of a press release is not related to the firm- or event-characteristics, but instead driven by exogenous factors.

Such a unique setting also contributes to the literature that studies the causal effect of media coverage. Engelberg and Parsons (2011) use extreme local weather events as exogenous shocks to news delivery. Peress (2014) uses a set of newspaper strike events as exogenous shocks to news production. Blankespoor, deHaan, and Zhu (2018) use the staggered implementation of robo-journalism to study the causal impacts of synthesizing information from analysts and other sources. Fedyk (2018) uses the random positioning of news on Bloomberg terminals to study the effects of being on the front page. This paper introduces a novel identification strategy that stems from the inefficiency in media production. Compared with previous work, the strategy also applies to a more representative and larger sample.

1.2 Interaction between media and other information channels

This paper also contribute to the literature that studies the interaction between different information channels in the financial market. Existing literature mostly focuses on the interaction between sell-side analysts and sophisticated investors. Kacperczyk and Seru (2007) find that fund managers with higher skills rely less on the public information from stock analysts. Chen et al. (2017) find that when the information production from analysts exogenously decreases due to the closures and mergers of brokerage firms, sophisticated investors scale up their information acquisition. This paper provides novel evidence that media could also impact the information production of investors and analysts. The results challenge the conventional wisdom that media does not matter for sophisticated market participants who have better information access and higher information processing skills.

The positive relationship between media coverage and the information production is consistent with at least two types of theories. The first possibility is that media coverage could increase the precision of the signal generated from the research by investors or analysts. Goldstein and Yang (2015) show a model where there are two fundamentals that affect the security payoff. They show that as the signal of one fundamental become more precise, investors will have incentives to acquire more information about the other fundamental. In the context of media and other information channels, the two fundamentals can be investor sentiment and the true valuation of the firm. If media coverage helps investors to better gauge investor sentiment, then investors may also extend their research about the fundamental value of the firm. Empirically media coverage has large impacts on investor sentiments. Tetlock (2007) documents the sentiment expressed in a Wall Street Journal column can predict even the aggregate trading next day, and Dougal et al. (2012) further exploit the exogenous rotation of the column writers and find the writer fixed effects increase the predictive power for the aggregate returns on the next day.

Another consistent theory is the rational attention allocation theory by Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016). In their model, investors first make decision about how to allocate their attention over a portfolio of firms. One prediction of their model is that more attention is allocated to assets with high prior payoff variance. The effect of media on the return volatility is well documented in the media literature (Peress, 2014; Blankespoor, deHaan, and Zhu, 2018). As a result, media coverage might serve as a reliable signal for the higher reward to the research of investors, who would then rationally allocate their attention and research to media-covered events.

It is important to note that these two possible explanations need not to be mutually exclusive, and this paper does not have enough evidence to empirically distinguish the two. However, a majority of the news articles by Dow Jones Newswires are concise summaries of the underlying events. These articles aim to be factual and accurate. As a result, these articles might have much smaller impacts on investor sentiment than articles from mass media. Therefore the context of the paper is closer to the rational attention allocation model.

2 Inefficient media coverage

2.1 Institution background

This paper studies whether Dow Jones Newswire is efficiently covering corporate disclosures. Information disclosed by firms is one of the most important sources of new information and often spurs large market movements. More importantly, most informational production by media, investors and analysts takes place closely following corporate disclosures. To see this point, in Figure 1, I plot the percentages of all the news articles, analysts earning forecasts, and web requests on EDGAR that are produced on different days following corporate disclosures, which are measured by corporate press releases[4]. We see that over 35% of all the news articles are published on days when firms issued new press releases. Most earning forecasts by analysts are issued immediately following corporate disclosures, with almost 50% of all the forecasts being published within two days after press release issuance. The evidence is consistent with the findings in Altınkılıç, Balashov, and Hansen (2013), who show that over 50% of analysts forecasts are issued following earning or guidance reports. Even EDGAR requests peak after corporate disclosures. About 20% of all the requests are made within two days of press releases. In this paper, I adopt an event-study approach around corporate press releases, and the evidence shows that these events represents important periods where most information production happens.

[Figure 1 here]

[4]The sample covers all the news articles and analysts forecasts covering the firms in my sample during 2004-2017. The EDGAR data is from Jan. 2004 to Jun. 2017. The press releases data includes all the press releases on the top 4 press release wires.

Since Regulation Fair Disclosure (Reg FD), press releases become an increasingly popular method for corporate disclosures, given their fast delivery and broad reach[5]. Typically, firms will choose one press release wire service to publish their announcements. Four wires take most of the market share in the US, namely, PR Newswire, Business Wire, Market Wire, and GlobeNewswire (Solomon and Soltes, 2012). The press releases can cover many topics. In my sample, the top three topics are "earnings", "products-services" (e.g., new product releases), and "labor-issues" (e.g., executive appointments).

These corporate press releases serve as one the most important inputs for journalists working at financial wire media like Dow Jones Newswire, Thompson Reuters, and Bloomberg. As different financial wire media competes over publication speed, the typical workflow for these wire journalists is to monitor press release feeds, select ones that could have market impacts, quickly summarize the contents, and publish the article to their subscribers. In the process, they produce two types of news articles: "news flash", which only contains a headline summarizing the event, or "full article", which contains some additional analysis and synthesizing information from other sources (Drake, Guest, and Twedt, 2014). Different from other market participants who also monitor information in real-time, wire journalists typically monitor a much wider set of information. Analysts and investors are usually industry- or firm-focused. When having limited attention, they may choose to focus on sector-specific rather than firm-specific information (Peng and Xiong, 2006). In comparison, the job for wire journalists is to report important news in a timely manner, without providing in-depth analysis. Therefore, wire journalists often cover the events for the whole market instead of specializing in any industries. For example, the job description for a wire journalist at the Dow Jones Newswire writes[6]

[5]Reg FD implicitly encourages the use of press releases due to its fast dissemination speed and wide reach of investors. Reg FD states that "technological developments have made it much easier for issuers to disseminate information broadly. Whereas issuers once may have had to rely on analysts to serve as information intermediaries, issuers now can use a variety of methods to communicate directly with the market. In addition to press releases, these methods include, among others, Internet webcasting and teleconferencing". Similar argument can be found in Neuhierl, Scherbina, and Schlusche (2013). The full content of Reg FD can be found at https://www.sec.gov/rules/final/33-7881.htm

[6]See the journalist's bio at http://www.wsj.com/news/author/8056

> Ian currently manages the U.K. companies desk, overseeing corporate news flashes and quick fire fills for both the Dow Jones Newswire and The Wall Street Journal's website. The desk covers all stocks from the largest FTSE100s to the smallest AIM companies across the whole range of industries and subject matters.

In the next two sections, I present evidence that corporate press releases cluster at specific times. Given the unique way that wire journalists monitor and process information, some important events may be inefficiently missed during busy times.

2.2 Press release clustering

In this section, I show that firms issue most of their press releases in non-trading hours. Moreover, within each hour, press releases heavily cluster at the exact-hour and half-hour points, thus creating possible cognitive burdens on wire journalists who need to process a large amount of information at those times.

I start by compiling a comprehensive sample of press releases using RavenPack PR Edition. The data contains all the press releases published in over 10 press release wires from 2004 to 2017[7]. Importantly, all of the top 4 news wires (PR Newswire, Business Wire, Globe Newswire, and Marketwired) are included. RavenPack also links these press releases to the public firms and provides the CUSIP number(s) associated with each release.

I impose a set of filters to generate the sample for this paper. First I require the `RELEVANCE` score to be 100 to ensure the firm is the main subject of the press release, as described in the RavenPack user manual. In addition, I only include the top 4 wires to exclude foreign press release wires, and remove duplicated releases by requiring the `ENS` score to be 100. Next, I require the firm to be in the Compustat/CRSP universe, and the press release is issued on a trading day. Finally, to better associate media coverage with

[7]Table 14 in the Online Appendix provides the full list of sources and the number of press releases issued in each source

press releases, I keep only observations where a firm only issues one press release on that day. Table 15 in the Online Appendix A shows the detailed data cleaning process as well as how the number of observations changes after each step.

[Figure 2 here]

Figure 2(a) shows that the press releases cluster in non-trading hours. To plot the figure, I split a 24-hour day into 288 five-minute bins. For each firm, I calculate the percentage of press releases published in each five-minute bin. Finally for each bin, I calculate the average percentage of the 9,386 firms. The bars in Figure 2(a) represent the average percentages. The first observation from Figure 2(a) is the contrast between trading and non-trading hours. The pre-market period (7-9AM) and the post-market period (4PM) contain the majority of the press releases, while the trading hours contain a much smaller number of press releases. Moreover, we see that the number of press releases also varies a lot within each hour. For example, while 8AM is a busy hour, the majority of the press releases are issued in the first five minutes (8:00 - 8:04) and the five minutes after the half-hour (8:30 - 8:34). To better see the pattern, I use a darker shade to denote the two five-minute bins after the exact-hour or half-hour points, and we can see that the darker bars stand out in almost every hour.

The pattern that press releases cluster at the exact-hour or half-hour points is more obvious in Figure 2(b), which plots the percentage of press releases by the minute of its publication time. On average, over 25% of a firm's press releases are issued in the first minute of an hour, and almost 15% issued in the thirty-first minute. These two minutes collectively consist over 40% of all press releases. The preference to make announcements at these "integral" points is not hard to understand. When we make appointments, we naturally tend to schedule events at these "integral" points like the exact hour or the half-hour. The same social convention applies when managers determine disclosure times. Think about a management team discussing when to disclose the new earning results. The plan very likely will be “to disclose at 8 o’clock sharp” rather than “let’s do 8:03”.

In the next section, I present evidence that when many firms choose to issue press releases at these integral points, the total number of press releases that wire journalists need to process can quickly goes up, causing possible cognitive burdens that make journalists more likely to miss some releases.

2.3 The blink effect: “on-screen time” and media coverage

Wire journalists may find it harder to efficiently process all the information when press releases arrive in clusters. Most real-time systems, which the wire journalists monitor, follow the ”first-in-first-out” rule: new content always shows up at the top of the user interface, pushing current content down and then onto next pages. In these systems, the number of new press release determines how quickly the user interface updates. Therefore, during busy times like 8:00 or 16:00, the computer screen in front of the journalists would quickly update as new releases keep coming in.

This paper asks a simple question to test the possible inefficiency: during these busy times, does the length of time that a press release stays in a prominent position (first page) of the compuster screen affect the press release’s media coverage? In this section, I show that a shorter time on the journalists’ computer screen (“on-screen” time) of a press release indeed leads to less coverage by Dow Jones Newswire. For parsiI call this effect the “blink effect”, as the journalists may blink-and-miss the press release. In the next section, I discuss the exclusion condition that on-screen time is exogenous to the underlying events and not likely driven by firm- or event-characteristics. I also find that the “blink effect” is most relevant to wire journalists and unlikely impacting analysts or investors directly. Finally I show that the blink effect is robust to sample selection, different news measures, and the functional form of the dependent variable.

The on-screen time of each press release is determined by the number of new press releases that are issued after it. The more new press releases that are issued in the next thirty seconds to a minute, the more quickly the press release is pushed to less visible positions. Therefore, for a press release j, I proxy its on-screen time by the number of new press releases that are published after j in the next 30 seconds (for parsimony, I refer to this number as NPRA hereafter). As the NPRA of a press release increases, the press release's on-screen time shortens. While the data in RavenPack PR Edition starts from 2004, the timestamps do not contain the precise second until April 1, 2006. Therefore the final sample of the paper starts from April 1, 2006 to allow for precise measurement of NPRA.

To measure busy clusters, I use press releases that are published in the first 10 seconds of each hour. An additional benefit of using such a tight frame is to ensure that the press releases in these clusters are similar in nature. Furthermore, the concern for strategic disclosure within a 10-second frame is also unlikely. In practice it is almost impossible to precisely control the second of publication. My results also hold if I relax this 10-second frame to a larger one, which is discussed in details in Section 5.

The media coverage data comes from RavenPack DJ Edition, which includes all the news articles published on the Dow Jones Newswire from 2000 to 2017[8]. Besides requiring the `RELEVANCE` score to be 100, I also drop news articles covering analysts analysis and stock market reactions to avoid reverse causality[9]. I define abnormal media coverage, *AbnNews*, as the log of 1 plus the number of news articles from Dow Jones Newswire minus the log of 1 plus the firm's average number of news articles per day in the past 60 calendar days.

Table 1 shows the summary statistics of the variables. To control for the effect of outliers, I winsorize all the variables at the 1% and 99% level[10]. My main regression sample contains 131,683 press releases from 7,503 unique firms. On average, a firm issues

[8]The data in the RavenPack DJ Edition also includes press releases, so one need to drop observations whose `NEWS-TYPE` is `PRESS-RELEASE`

[9]To do that, I drop news articles whose topic variable `GROUP` is in 'analyst-ratings', 'credit-ratings', 'order-imbalances', 'technical-analysis', 'stock-prices', or 'price-targets'

[10]All of the results in this paper are robust to removing the winsorizations

about 12.50 press releases in a year, and 3.81 of them will be in the first 10 seconds of an hour. On average 48% of the press releases will be covered by Dow Jones Newswire, and each press release receives an average of 1.75 media articles.

To formally test whether the on-screen time affects the amount of media coverage, I estimate the following regression using the sample of 131,683 press releases.

$$AbnNews_{ijt} = \beta \log(NPRA_{ijt} + 1) + \boldsymbol{\alpha} + \varepsilon_{ijt} \tag{1}$$

In the regression, $AbnNews_{ijt}$ is the abnormal media coverage measure. The key independent variable is NPRA, which measures the number of new press releases issued immediately after the press release j in the next 30 seconds. $\boldsymbol{\alpha}$ represents a set of fixed effects that I am going to include in the regression.

[Table 2 here]

I find that a larger NRPA, thus a shorter on-screen time, significantly decreases the amount of media coverage even after controlling for a rigid set of firm- and event-characteristics. Table 2 shows the results. As a benchmark, in Column (1) the regression only includes the date-hour fixed effects to control for the differences across clusters. The coefficient estimate is highly significant and economically large. Doubling NPRA[11], thus the on-screen time of a press release decreases, will decrease the amount of abnormal news by 13.4%. I incrementally introduce more fixed effects to control for possible omitted variables at the firm and the press release level. Column (2) further includes firm fixed effects to control for firm-invariant characteristics. Compared with the coefficient estimate in Column (1), the new estimate only slightly decreases in magnitude (from -0.134 to -0.132). Column (3) introduces firm-year interacted fixed effects, and the analysis is therefore to compare two press releases issued by the same firm in the same year. The more rigid firm-characteristics

[11]The standard deviation of log(NPRA+1) is 1.11, thus doubling NPRA is close to an increase of one standard deviation.

controls actually increase the magnitude of the coefficient estimate. The coefficient estimate in Column (3) is -0.163. These results reveal that the endogenous factors at the firm-level likely work against me finding the blink effect.

To control for the characteristics of different press releases, I utilize the topic measures developed by RavenPack. RavenPack applies sophisticated textual analysis to categorize press releases into different topic groups. These topic measures are essentially the summaries of the underlying events. In Column (4) of Table 2, I further include the fixed effects that control for broad topic classifications (28 unique groups). With the new fixed effects, Column (4) still shows a significant estimate of -0.099. In Column (5), I further control for a more detailed topic classification and include the new topic fixed effects (143 topic groups) in the regressions. The richer fixed effects only slightly decrease the magnitude of the coefficient estimate, which now becomes -0.093 with a t-statistics of -7.74.

These results show that the on-screen time significantly affects the amount of media coverage on the event day. To show that this effect is indeed a form of inefficiency, I need to show that the on-screen time is not driven by any factors that measures the importance of the events. In Table 2 the the significant effect persists even if I include a rigid set of fixed effects which absorb many firm or event-characteristics. Instead of keep adding more controls, in the next section, I explicitly test whether NPRA, my proxy for the on-screen time, is exogenous.

2.4 Exclusion condition

The randomness of the on-screen time comes from two possible sources: the ordering within the cluster, and the number of press releases from other (most likely unrelated) firms. In this section, I first conduct covariate balancing tests to show that my proxy for the on-screen time is indeed uncorrelated with the observable firm- or event-characteristics. Second, I conduct a series of falsification tests to show that it is the journalists' limited cognitive capacity, rather than unobserved economic properties of the press releases, that drives the effect. Finally, I present evidence that shows the blink effect is unlikely affecting investors and analysts.

2.4.1 Covariate balancing test

I start by testing whether the proxy for the on-screen time, NPRA, is correlated with any firm or event characteristics. Because I argue that NPRA is random, then the null hypothesis, that NPRA is uncorrelated with any of the observable characteristics, should hold.

To test the hypothesis, I regress the log(1 + NPRA) on different characteristics that are shown to affect media coverage choices (Solomon, 2012). These includes firm-characteristics including Tobin's Q, total asset, and firm age, and event-characteristics like event sentiment or the title length[12]. RavenPack adopts sophisticated natural language processing techniques and expert reviews to generate a set of sentiment scores. The goal of these sentiment scores is to create sufficient event summaries and to assist in trading. For example, the press release in which Sanofi announces positive results for a trial study on June 6th, 2015[13], receives an event sentiment score (ESS) of 87. In comparison, the press release in which Micron disclosed decreases in demand receives an ESS score of 17[14].

[Table 3 here]

Table 3 shows that none of these characteristics significantly predict NPRA. In Columns (1) to (5), I regress the log(1 + NPRA) on these observables individually. None of them exhibits significant estimate. The coefficient estimate remains insignificant even when all the characteristics are included as independent variables. Their joint F-statistics is 1.02 with a p-value of 0.41.

[12]While a better measure would be the length of the full press release, RavenPack does not provide any this or any similar measure. I am working on merging the RavenPack data with the press release data from LexisNexis to generate the full-article length measure.

[13]See http://mediaroom.sanofi.com/sanofi-announces-positive-results-for-toujeo-in-phase-iii-study-extension-in-japanese-people-with-uncontrolled-diabetes-2/ for the press release

[14]See http://investors.micron.com/releasedetail.cfm?ReleaseID=440412 for the press release

In sharp contrast, these covariates significantly predict the amount of media coverage. In Column (8) of Table 3, I use the log of 1 plus the number of media articles on the event day as the dependent variable. All the coefficient estimates are significant, and jointly their F-statistics is 200. The result shows that these covariates are of great relevance to media coverage. Yet all the results do not reject the null hypothesis that NPRA is uncorrelated with any of them.

2.4.2 Falsification test

In this section, I conduct a series of falsification tests to show that it is indeed the limited cognitive capacity of journalists, instead of unobserved omitted variables, that drives the blink effect.

[Table 4 here]

First, in parallel to human journalists, automated algorithms in Dow Jones Newswire will also redistribute some press releases. These algorithms are typically based on editorial judgments about which firms or event types are relevant to the market. The exclusion condition is that NPRA should not correlate with these economic factors, thus we would expect NPRA to have little impacts on such automated coverage. Indeed, Column (1) of Table 4 shows an insignificant estimate of -0.005 when the dependent variable, DJPR, is a dummy that equals to 1 if a press release is covered by the automated coverage. Another natural placebo test is to see whether the media coverage prior to the press release day, which could represent the existing interest about the coming disclosures, changes with NPRA. Column (2) of Table 4 also shows an insignificant estimate of -0.003.

If the cognitive capacity is the cause, then the blink effect should be stronger in busier times, and disappear during times that are not busy. Consistent with the hypothesis, I find similar blink effect in the press releases published in the 31st minute of each hour. As shown in Figure 2(b), the 31st minute, which is the half-hour point, also holds clusters of press releases. Column (3) of Table 4 shows a significant estimate of -0.085, indicating a similar blink effect in the 31st minute as well.

The result shows a drastic change when I estimate the same regression using press releases from all other minutes (excluding the 1st and 31st minutes). Column (4) of Table 4 in fact shows a positive coefficient estimate of 0.051. The result shows that the endogenous factors very likely will work against me finding the blink effect. When issuing important press releases, firm would spend more efforts to ensure that the press releases are issued on time, thus the press release is more likely to be at the early part of a cluster and has higher NPRA. Such endogenous forces will work against me finding a negative effect of NPRA on media coverage.

I also separate the sample into the busy hours (7-9AM and 4PM) and all non-busy hours (all other hours), based on the pattern in Figure 2(a). Similarly, we would expect the effect to be stronger in the busy hours and insignificant in the non-busy hours. Column (5) and (6) of Table 4 show consistent results. The coefficient estimate is -0.112 in the busy hours, and becomes insignificant in all other hours. Finally, I also sort the sample into quintiles by NPRA, and test how the blink effect changes with quintiles. Column (7) of Table 4 shows that the effects are insignificant in the first two quintiles (lower NPRA) and become significant and stronger in quintiles that have higher NPRA.

To sum up, the covariate balancing tests and the falsification tests shows that omitted variables unlikely drive the negative relationship between the on-screen time and media coverage. The evidence is consistent with that the limited attention capacity of journalists is the cause, as the effects become stronger in busier times and disappear in non-busy times.

2.4.3 Blink effect for investors and analyst?

So far, the evidence shows that a press release's on-screen time significantly impacts its media coverage. Moreover, the on-screen time is unlikely correlated with firm or event-level characteristics. Therefore the on-screen time is also a prime candidate to be an instrumental variable to the amount of media coverage. However, the exclusion condition further requires that the instrumental variable does not directly affect investors or analysts, which are the subject of the analysis later. In this section, I show that investors and analysts are unlikely to experience the same blink effect.

First of all, Table 4 shows that the blink effect only happens in situations where the amount of new information is so large that the cognitive capacity is stretched to the limit. Compared with wire journalists, investors and analysts follow a smaller set of firms. Almost all the real-time systems allow for information filters, thus investors and analysts can monitor an event stream that is individually customized. As a result, when the amount of information for the whole market is too much to handle, the information exposure to investors and analysts is still easily manageable at the industry level. As supporting evidence, Column (6) of Table 3 shows that the on-screen time is uncorrelated with the total number of press releases issued by firms in the same industry (2-digit SIC[15]).

The second necessary condition for the blink effect is that the user needs to monitor the event stream in real-time. Wire journalists do not directly control the information inflow because their job is to screen the information as it comes. Investors and analysts, on the other hand, may adopt other information acquisition methods. They could actively search (Da, Engelberg, and Gao, 2011; Ben-Rephael, Da, and Israelsen, 2017) or set up alarms only for the firms or events that they intend to follow. In these modes, the on-screen time of press release becomes an irrelevant factor.

In Section 5, I provide additional empirical evidence to support the hypothesis that the on-screen time is not directly impacting investors or analysts. The idea for the tests comes from the assumption that wire journalists follow the press releases from private firms while investor and analysts don't. When the on-screen time decreases due to the new press release issued by private firms, the effect should be equally strong if all the

[15]The result is robust to using 1-digit SIC industry, or the text-based industry classification by Hoberg and Phillips (2016)

effects are indeed through media. I find consistent evidence with the hypothesis, and Section 5 introduces the detailed results.

2.5 Robustness

This section shows that the blink effect is robust to news measures, sample selection, and the functional form of the dependent variable.

[Table 5 here]

In Columns (1) - (3) of Table 5, I re-estimate the blink effect using different news measures. In Column (1), I drop possibly duplicated media articles by requiring a news' novelty score (`ENS`) to be 100 and still find a significant estimate (-0.085). In Column (2) and (3), I separately test how NPRA affects flash news and full articles. The effect on the flash news is much stronger than the effect on full news, consistent with that the production of flash news has higher pressure in speed.

The results are also robust to the sample construction, alleviating concerns about possible selection biases. First of all, from the data I am not able to distinguish the press releases that are pre-scheduled from the ones that are issued at the spot. Empirically, most pre-scheduled press releases will be published in the first second. In Column (4) of Table 5, I exclude all the press releases issued in the first second of an hour. The coefficient estimate is -0.091 with a t-statistics of -5.62, suggesting that whether a press release is pre-scheduled or not is unlikely biasing my result. Second, one potential limitation of the paper is that I could not address firms' endogenous choices to issue any press release or not. While the covariate balancing tests show that the on-screen time is not correlated with the sentiment of the press releases, I further test whether the results are driven by frequent press release issuers. In Column (5), I exclude observations where the total number of press releases issued by a firm in a year is above the 75th percentile. The coefficient remains significant and only slightly drops in magnitude (-0.085 vs. -0.093), suggesting the results are not driven by those frequent issuers. Third, the main regressions control for a rigid set of fixed effects like firm-year or date-hour fixed effects. If a firm only issued one press release in a year, or the first ten seconds of an hour only contain one press releases, these observations will be completely absorbed by the fixed effects. To ensure that including these observations does not inflate the t-statistics, in Columns (6) and (7) of Table 5, I drop these observations. The coefficient estimates are identical to the base case (as expected) and the t-statistics are still above 8. Finally, I also test whether the result hold in a larger sample by expanding the sample to the press releases issued in the first 30 seconds of an hour, and Column (8) again shows significant estimates.

I further show the results are also robust to the functional form of the dependent variable. Column (9) uses the log of 1 plus the raw number of media articles on the event day, and shows that the result is robust to using the abnormal measure. Column (10) shows that the result is also robust to the using the log measure. The dependent variable in Column (10) is the raw number of of media articles on the event day minus the average daily number of media articles in the past 60 days. The result is close to the baseline estimate both in the significance and magnitude. Finally, Column (11) uses a dummy variable that equals to 1 if there is any media coverage on the event day as the dependent variable. I find that doubling NPRA decreases the probability of news coverage by 4.5%, representing a 9.4% decrease from the sample mean (48%).

3 The causal effect of media on information production

Because the "on-screen" time of a press release impacts media coverage, while at the same time is exogenous to the underlying event and most directly impacts wire journalists, it could act as an instrument for me to study the causal effects of coverage by Dow Jones Newswire. While existing research has provided ample evidence on how media coverage affects stock trading and returns, less is known about whether and how media coverage impacts other information channels. In this section, I show evidence that media coverage causes investors to acquire more information and analysts to produce more earning forecasts. Because Table 4 reveals that the blink effect is only significant in busy hours (7-9AM and 4PM), the following analyses will only use the press releases published in the first 10 seconds of these busy hours.

3.1 Investors' information acquisition through EDGAR searches

I start by testing whether media coverage increases the information acquisition of investors. To measure information acquisition, I use the EDGAR server log files from SEC[16]. When someone accesses a web page or a filing on the EDGAR system, the log data will create an observation with (1) the IP address of the visitor, (2) the CIK number, a unique firm identifier used by EDGAR, (3) the accession number, a unique filing document ID, and (4) the timestamp of the web request. The data contains all the web traffic records from Feb. 2003 to Jun. 2017. Following Loughran and McDonald (2017), I apply several filters: omit index page requests to avoid double counting and drop requests with server codes of 300 or higher. Then I test the following regression

$$\text{AbnEdgar}_{ijt} = \beta \log(NPRA_{ijt} + 1) + \boldsymbol{\alpha} + \varepsilon_{ijt} \tag{2}$$

$AbnEdgar_{ijt}$ is the abnormal number of EDGAR searches about firm i. The abnormal number of EDGAR searches is defined as the log of 1 plus the number of EDGAR searches about firm i on day t minus the log of 1 plus the average number of EDGAR searches about firm i in the 60 days prior to the press release. In the regression, I control for the same set of fixed effects as in Table 2 Column (5), including firm-year, date-hour, and detailed topic fixed effects.

[Table 6 here]

[16]Details on the data can be found in https://www.sec.gov/dera/data/edgar-log-file-data-set.html. Also see Loughran and McDonald (2017) for a general discussion about the data.

I find that media coverage significantly increases the abnormal number of EDGAR searches from humans, but not the searches from web crawlers. To identify the searches from possible human users, following Lee, Ma, and Wang (2015), I tag an IP address as a web crawler if it searches for more than 50 firms in a day[17]. Column (1) of Table 6 first shows insignificant estimate of β when the dependent variable is the total abnormal number of EDGAR searches. However, once I decompose the searches into human and web crawler searches, as shown in Columns (2) and (3), respectively, we observe a stark contrast. Media coverage significantly increases the number of EDGAR requests from human users: doubling NPRA reduces the amount of abnormal number of EDGAR searches by 4.4%. In sharp contrast, I observe no significant effect for requests made by web crawlers. Supposedly, these web crawlers operate following algorithms that track certain firms or filing types. The insignificant coefficient again validates the exclusion requirement as any economic factors that determine the algorithms do not correlate with NPRA.

Figure 3(b) plots the coefficient estimate of β in Equation 2 using other event days. I re-estimate the Equation 2 and replace the dependent variable by the abnormal number of EDGAR searches on other event days. The coefficient estimate is insignificant on the day prior to the press release, consistent with the falsification test in Table 4. The result suggests that NPRA is uncorrelated with ex-ante firm characteristics. The coefficient estimates become significant on the event day and persist for about 2 days. By day 5, the coefficient estimate almost completely converted back to 0. Overall, the impulse-response pattern is consistent with the blink effect being a random transitory shock.

Are all the effects completely driven by investor inattention? In other words, will investors not know about the press releases at all if media does not report it? I find that investor inattention is not likely the only cause.

First, I find that investors who already follow the firm on EDGAR systems also signif-

[17]This method leads to conservative measures of human EDGAR searches, as when multiple users share one outbound IP address search fro more than 50 firms in a day, this IP might be categorized as a web crawler.

icantly increases their EDGAR searches after media coverage. As shown by Chen et al. (2017), investors (mutual fund managers) exhibit very persistent searching activities on the EDGAR system. Moreover, the investors who already searched for a firm's filings in the past suffer less from the search problem associated with inattention (Barber and Odean, 2008) because they already know about the firm. If inattention is the only cause for the effect of media on EDGAR searches, we would expect a weaker effect from these existing EDGAR users. However, I actually find a slightly stronger effect for these investors. In Column (5) of Table 6, I only include human EDGAR searches where the same IP address has accessed any document from the same firm in the previous month. The coefficient estimate is -0.046, suggesting a 4.6% decrease in abnormal EDGAR searches when NPRA doubles.

Second, I also find institutional investors also strongly respond to the media coverage. Institutional investors are among the most sophisticated participants in the financial market, and relative to retail investors, they should have suffer less from inattention. To identify searches from institutional investors, I first match IP addresses to known institutions which have an autonomous system number (ASN). The IP-block and ASN link file comes from MaxMind[18]. The link file specifies different ranges of IP addresses that are assigned to different institutions. Because these institutions are not limited to the finance industry, so the second step is to identify which ones are actually financial institutions. I use two methods to do that. First, I directly search for finance-related words like "bank" or "fund" in the institution's names. Second, I compile a list of names from all 13F institutions, and use a name matching algorithm[19] to identify institutions that are in the universe of 13F institutions. Online Appendix B further describes these two methods in details. Columns (5) and (6) of Table 6 show the regression result for the EDGAR searches from financial institutions[20]. In Column (5), I identify institutional investors by searching

[18]https://dev.maxmind.com/geoip/geoip2/geolite2/

[19]The algorithm is based the code written by Jim Bessen, available at http://goo.gl/m4AdZ.

[20]In counting the searches from institutional investors, I only exclude EDGAR searches where the user-agent explicitly state the search is from a web crawler. The choice is because when the previous crawler filter

for finance-related words, while in Column (6), I identify institutional investors by matching to the names of 13F institutions. Both columns show significant estimates. Notably, Column (5) shows an even larger magnitude, -0.057, than the baseline estimate.

Finally, I find that the results are robust to using two-stage least square estimations. In the first stage, I regress the amount of abnormal news coverage on the event day on the instrument, NPRA. The first stage result is the same as reported in Column (5) of Table 4. The conditional F-stat for the instrument is 25.5, well above the rule of thumb threshold of 10. In the second stage, I regress the abnormal number of EDGAR searches from the day after event day, on the predicted abnormal news from the first stage. Column (7) of Table 6 shows a significant effect in the second stage: doubling the amount of news coverage on the press release day, the abnormal number of EDGAR searches on the next day will increase by 30%.

These results, that existing EDGAR users and institutional investors also increase their information acquisition when media coverage increases, are difficult to be fully explained by a pure attention effect. A rational framework may explain these responses. Because media coverage is empirically associated with an increase in return volatility and a possible change in investor composition favoring noise traders, sophisticated traders might rationally allocate more resources to media-covered events in order to extract higher expected returns from trading. To further test the explanation, I examine whether analysts are also impacted by media coverage. The test is motivated by the increased number of EDGAR searches from institutional investors, who are also the major clients of analysts. Will the increased information demand from institutional investors also extend to their requests to analysts? Will analysts, who are unlikely inattentive to the firms they follow, change their information production with more media coverage? I explore these questions in the next section.

is imposed, most institutional IP addresses will be eliminated. As previously noted, multiple users within an institution might share one out-bound IP address. Including potential web-crawlers, as shown in the Column (3) of Table 6, works against me finding significant results.

3.2 Analysts' information production

This section tests whether and how media coverage affects the information production of analysts. From the supply side, media coverage is unlikely to change the production function of analysts. First, unlike investors who face the search problem, stock analysts focus on a small set of firms that in the same or related industries. It is their job to closely follow these firms and provide timely analysis. Thus it is hard to think that analysts would miss a corporate press release if no media covers it, given that the ability to collect information is relative advantage of analysts. Second, analysts have better information access (like direct communication with the management) and higher information processing skills. It is not clear whether analysts could learn any useful information from media articles produced by wire journalists. These articles are often just summaries of the original press releases. The null hypothesis that media coverage has no effect on analysts is therefore appealing.

However, the coverage decision, that whether an analyst issue any new earning forecast for a firm, is also affected by the demand side, which is institutional clients. As noted in a survey of 182 analysts from Brown et al. (2015), "Demand from their clients is analysts' most important motivation for making profitable stock recommendations and their second most important motivation for issuing accurate earnings forecasts". As institutional investors exhibit increased information demand from more EDGAR searches, it is possible that institutional investors also request more information from analysts, who might in turn prioritize these requests and shift more resources to cover media covered events.

To test these hypotheses, I collect the analyst forecasts from the unadjusted detail file in I/B/E/S. I require the observations to have non-missing EPS forecasts, and then merge it to the press release events using CUSIP numbers and announcement dates. An analyst may issue multiple forecasts that have different period ends for the same firm on a single day. These forecasts would be separate records in the I/B/E/S data. To avoid double counting, for each firm-day, I count the number of unique analysts who issue any EPS forecast as the relevant measure. Thus multiple forecasts from the same analyst would only be count once for a firm. I then estimate the following regression

$$\text{AbnAnalyst}_{ijt} = \beta \log(NPRA_{ijt} + 1) + \boldsymbol{\alpha} + \varepsilon_{ijt} \tag{3}$$

The dependent variable, AbnAnalyst, is the abnormal number of analyst forecasts. Similar to previous abnormal measures, the abnormal number of analyst forecast is defined as the log of 1 plus the number of analysts issuing any earning forecast for firm i on day t, minus the log of 1 plus the average number of analysts issuing forecasts per day in the past 60 calendar days. The regression includes the same set of fixed effects: firm-year, date-hour, and detailed press release topic fixed effects.

[Table 7 here]

Consistent with the demand-catering hypothesis, I find that media coverage significantly increases the number of analysts that issue earning forecasts for the firm. Column (1) of Table 7 shows the regression result where the dependent variable is the abnormal analyst estimate measure on the press release day. The coefficient estimate is -0.060 with a t-statistics of -5.96. Doubling NPRA reduces the number of analyst forecasts by 6% on the event day.

The effect on the next day is even larger. Column (2) of Table 7 shows the regression results where the dependent variable is the abnormal number of analyst forecasts on the day after the press release issuance. The coefficient estimate is -0.071 with a t-statistics of -7.09. The stronger reaction is consistent with the evidence in Figure 1, where the information production of analysts peaks on the day after the press release.

Figure 3(c) shows the coefficient estimate of β in Equation 3 for different event days. Similar to the pattern of EDGAR searches, the coefficient estimate is insignificantly different from 0 on the day prior to the press release. The coefficient estimates become significant on the next two days, and fall back to the pre-event level on the third day. The coefficient gradually converges back to zero in the next three days.

To better distinguish whether the effects are through the production or demand side, I first test whether analysts of different characteristics show different responses to the media coverage. On the production side, one hypothesis is that analysts may learn additional information from media coverage thus increase the accuracy of their forecasts. To test the hypothesis, I estimate the effect of media coverage separately on the analysts of different skills. Different skilled analysts likely exhibit different learning intensities from media coverage. It is possible that lower-skilled analysts might rely on media to help them judge the importance of the events, or conversely, high-skilled analysts could better extract the subtle new information from media coverage. The key prediction here is that different-skilled analysts should react differently to media coverage.

I thus split the analysts into two groups based on their skills, which are measured by their relative forecast accuracy in the previous year. I calculate the relative accuracy following the procedures in Hong and Kubik (2003) and Ljungqvist et al. (2007)[21]. In Column (3) of Table 7, the dependent variable is the abnormal number of analyst forecasts from analysts whose relative accuracy is above the median accuracy in the previous year. Because the effect of media coverage on analysts is even stronger on the day after the press release, I thus use the cumulative abnormal number of analyst forecasts on the two days after the press release day (including) as the dependent variable. The coefficient estimate is -0.068 with a t-statistics of -6.57. In Column (4), I test for the analysts who have below-median accuracy in the previous year, and the estimated effect, -0.062, is close to the estimate in Column (3). In Column (5), the dependent variable is the difference of the two dependent variables in Columns (3) and (4), and the coefficient estimate is insignificant. These results indicate that different-skilled analysts do not react differently

[21]The procedure exactly follows the footnote 6 in Ljungqvist et al. (2007). For analyst i covering firm k in year t, I first calculate the absolute forecast error using the following steps. (1) get the analysts most recent forecast of year-end EPS issued between Jan. 1 and Jun. 30, (2) calculate the difference with the subsequent realized earnings, (3) scale the difference by previous year-end price. Then for all the analysts covering firm k in year t, I re-scale the absolute forecast errors so that the most and least accurate analysts scores one and zero, respectively. Finally, analyst i's relative forecast accuracy in year t is his/her average score across the the stocks he/she covers over years t-2 to t.

to media coverage.

[Table 8 here]

I next test whether a demand-side reason could explain the positive effects. As the evidence in the EDGAR searches shows, institutional investors exhibit increased information demand for media-covered events. It is possible that these institutional investors may also request more information from analysts. If analysts have unlimited resources and could easily scale up their information production, then the increased demand would not necessarily impact these analysts because they can always cover all the events that are important, even for those that are not covered by media. However, if the resource constraints for analysts bind and analysts can only cover a subset of events, then the demand-side shock will shift analysts' research to these media-covered events. I find consistent results with the hypothesis. I define resource constraints by the number of firms that an analyst covers in the previous year. Similar to the previous analysts, I split the analysts into two groups based whether they covered an above-median (more-constrained) or below-median number (less-constrained) of firms. Column (6) and (7) of Table 7 separately show the regression results for the more-constrained and less-constrained analysts. The coefficient estimate is much larger for the constrained analysts in Column (6) compared with the less-constrained analysts in Column (7) (-0.090 vs. -0.013). Column (8) further shows that the difference between the two groups are significant.

As additional evidence for the demand-catering hypothesis, I also test whether the effects are stronger for firms with higher institutional holdings or more institutional investors. As the number of institutional investors increases, an analyst is more likely to experience a positive demand shock, thus the effect of media on analysts should be stronger. To test the hypothesis, I first sort the sample into two groups by its (1) institutional ownership or (2) the number of institutional investors in the previous year. The institutional holding data comes from Thomson Reuters. Because the blink effect itself (first-stage regression) could be different in these two subsamples, I thus estimate the 2SLS regressions separately using the two subsamples and compare the treatment effect from the second-stage. Column (1) and (2) of Table 8 show the second-stage estimates. The dependent variables are the cumulative abnormal number of analyst forecasts on the two days after the press release, and the key independent variable is the predicted abnormal news coverage on the event day from the first stage regressions. The estimated treatment effect of media coverage is about twice the size in the high-institution holding group of the estimated effect in the low-institution holding group. In Columns (5) and (6) of Table 8, we see similar differences when I split the sample using the number of institutional investors. Consistent with the hypothesis, I find that the effect on analysts is stronger for firms with more institutional investors.

Overall, this paper documents a large effect of media coverage on the information production of analysts. As the number of media coverage increases, more analysts also issue earning forecasts for the media-covered firms. I find that such a effect is unlikely driven by a learning mechanism where analysts extract additional information from media coverage. Rather, a demand-catering mechanism, where institutional investors increasingly request more information from analysts for media-covered events, is most consistent with the evidence. Throughout the analysis, I have precluded the possibility that analysts are inattentive to these press releases. While this paper could not fully reject such hypothesis, the existing literature and the institution background provides little empirical support to this view[22].

4 Market outcomes

The rational attention allocation hypothesis in this paper suggests that sophisticated investors focus on media-covered events to obtain higher returns. This section shows that

[22]For example, Hirshleifer, Hsu, and Li (2013) use analyst coverage as a proxy for investor attention. The implicit assumption is that analysts are well attentive to the firm information.

media coverage attracts trade from both possible informed and noise traders. Collectively, the results are consistent with a tug of war between these two type of investors: while the trading volume significantly increases, the overall price efficiency is not affected.

The market outcome variables come from a variety of sources. The trading volume, intraday price range, and daily stock return data comes from CRSP. I calculate the abnormal daily turnover following Tetlock (2010) as the log of 1 plus the turnover minus the log of 1 plus the average daily turnover in the past 60 trading days. The effected spread variables, both equally-weighted and value-weighted, is generated from TAQ data following Goyenko, Holden, and Trzcinka (2009). More specifically, it is calculated as $2|\log(P_k) - \log(M_k)|$, where P_k is the price of the trade, and M_k is the mid-point of the consolidated BBO at the time of the trade. I define abnormal spread following Blankespoor, deHaan, and Zhu (2018) as the effective spread over the average daily effective spread in the past 60 trading days[23]. Following Peress (2014), I define daily range as the log of the daily high price minus the log of the daily low. Finally, I calculate abnormal returns by subtracting the CRSP value-weighted index return from the daily raw returns. Because the sample contains press releases from both pre- and post-market hours, I now define the event day, or day 0, as the first trading day after the press release is issued.

[Table 9 here]

I start by showing that media coverage significantly increases trading volumes. I estimate a similar regression as Equation 7 but with different dependent variables.Column (1) of Table 9 shows the regression results where the dependent variable is the abnormal turnover on the event day. Doubling NPRA, the abnormal turnover will decrease by 7.7%. Figure 3(d) further plots the coefficient estimate for different event days. The significant effect of media coverage on the trading volume persists in the next three trading days, and

[23]This definition slightly differs from the definition in Blankespoor, deHaan, and Zhu (2018), who use the window [-40, -11] to calculate the average effective spread. I use the window of past 60 trading days for consistency with other tests. My results are robust to using the window of [-40, -11] to calculate the average daily effective spread.

finally converges to 0 on the 5th trading day. The large effect on the trading volume is consistent with a growing literature that find similar effects. Barber and Odean (2008) find that news coverage significantly increases the buying activities of retail investors. Engelberg and Parsons (2011) find that investors trade a stock more when local newspapers cover the firm's earnings. Peress (2014) documents a 12% decrease in trading volume during strike days when newspapers could not be produced or delivered. Blankespoor, deHaan, and Zhu (2018) find that trading volume increases by approximately 11% when the firm is covered by machine-generated news articles. Fedyk (2018) shows that news articles at the front-page induces 280% higher trading volumes than other similar news at less prominent positions in the next 10 minutes. The large effect documented in this paper confirms that my unique setting is highly relevant to the financial market.

As evidence for the participation of informed traders, I find that the effective spread widens on the event day with more media coverage. In Columns (2) and (3), I use the equal-weighted and value-weighted effective spread on the event day as the dependent variable. The coefficient estimates in both columns are significant. Doubling NPRA would decrease the equal-weighted (value-weighted) abnormal effective spread by 3.1% (3.5%), suggesting an increase in the relative proportion of informed trades. The result is consistent with the increased information production of institutional traders documented before. The effect on the effective spread disappears in the next two trading days, as shown in Columns (7) and (8) of Table 9. The coefficient estimate for the average daily abnormal effective spread is insignificant and quantitatively tiny (-0.001 for equal-weighted and 0.004 for value-weighted). Indeed, as analysts earning forecasts increase the amount of public information and more retail traders participate in the trading, the information asymmetry for market makers should decrease.

The evidence also shows that media coverage attracts more price-insensitive traders on days after the press release day. More specifically, I find on days [1, 2], while media coverage does not change the overall level of return, it increases the intra-day trading price range. Column (9) of Table 9 shows regression result where the absolute cumulative abnormal return on days [1, 2] is the dependent variable. The coefficient estimate is only -0.001 and insignificant. In contrast, Column (10), where the dependent variable is the average daily price range on days [1, 2], shows significant coefficient estimate (-0.077). Such results are similar to the findings in Peress (2014), who finds that the absence of newspaper decreases the intraday price range while the aggregate level of return is unchanged. Peress (2014) attributes similar effects to "less price-sensitive traders who transact at less favorable prices".

Collectively, the results suggest a tug of war between two types of investors, one attracted by media due to attention, and the other that consciously trades more media-covered firms, profiting by trading against these potentially uninformed traders. The lead-lag responses, where the informed trade happens on the event day while less price-sensitive trade happens on the next two days, are possibly caused by the slow information diffusion to retail traders, who might rely on mass media for information (Blankespoor, deHaan, and Zhu, 2018). The pattern also echoes the findings in Ben-Rephael, Da, and Israelsen (2017), who document a similar lead-lag relationship in information searches. It is likely that these two types of traders would have the opposite effects on price efficiency. Consistently, I find the overall price efficiency is unchanged by media coverage. To measure price efficiency, I use the delayed response ratio measure from Dellavigna and Pollet (2009). While the original work of Dellavigna and Pollet (2009) uses a period of 75 days to measure the total cumulative returns, I construct the delayed response ratio measure using a range of periods, as my sample is not restricted to earning announcements, which is what Dellavigna and Pollet (2009) focus on. Across all the measures, as shown in Table 10, none of them show significant results. The results are in line with the findings in Blankespoor, deHaan, and Zhu (2018) who find that news does not improve nor impede the price efficiency.

[Table 10 here]

5 Robustness

This section addresses several concerns that might bias the results of the paper.

5.1 Are the effects through media coverage?

This paper argues that all the changes in EDGAR searches and analyst forecasts are causally impacted by media coverage. Such interpretation might not be valid if variation in the on-screen time can directly impact investors and analysts. As a result, the observed effects are not necessarily through journalists but from a common shock to all. However, this paper argues that the measured on-screen time is unique to wire journalists due to the way they monitor and process information. The measure is uncorrelated with the amount of information exposure at the industry level. In this section, I provide additional empirical support to show that the effects are mostly likely through journalists.

An ideal experiment to test whether the effects are through journalists is to find another shock that only impacts journalists. If the effects are not through journalists, such exogenous shock should cause no changes in analysts and investors. One possible candidate for the shock is the number of press release from private firms. While some investors might also care about news from private firms, it is unlikely that they will monitor the press releases from these private firms in real-time, because such information is generally not tradable. In comparison, wire journalists also cover news events from private firms, and the press release from private firms affect wire journalists in the same way as the releases from public firms do. Therefore, the variation in the number of press releases from private firms is likely a shock that only affects wire journalists.

Given the above assumption, I first test whether the composition of NPRA, that whether the following press releases are from private or public firms, changes the effect on analysts and investors. Suppose the effects are not through journalists, then for a press release, as the proportion of its following press releases that are from private firms increases, the effects on investors and analysts should decrease (as the absolute number of press release from public firms decreases), while the effect on journalists should stay unchanged (as the total number of press releases does not change). I test this hypothesis by including an interaction term, log(NPRA + 1) x % Priv, where % Priv is the percentage of press releases in NPRA that are issued by private firms. The dependent variables come from the main tests before: the abnormal news coverage, the two-day cumulative abnormal number of EDGAR searches, the two-day cumulative abnormal number of analyst forecasts, the abnormal turnover on the event day, the absolute abnormal return on the event day, and the price range on the event day. Panel (A) of Table 11 shows that the coefficient estimate for the interaction term is not significant in all the tests, while the coefficient estimates for log(NPRA+1) are almost identical as in my baseline results.

[Table 11 here]

The insignificant effect could also be due to the power of the test. Therefore I next test whether the variation in the number of press releases from private firms can generate similar effects as documented in my main tests. In Panel (B) of Table 11, I directly use log(NPRPriv + 1), the number of press releases from private firms in the next 30 seconds following the press release, as the key independent variable. To control for the information exposure at the industry level, I also include log(NPRInd + 1), which is the total number of press releases from firms in the same industry on that day in the regression. The results show that all coefficient estimates for log(NPRPriv+1) are significant, and the magnitude are also similar to the effects in previous tables.

Altogether, the results show that the investors and analysts do not react differently to press releases issued by private firms. The evidence is most consistent with the hypothesis that the effects in this paper are through journalists, meaning interpreting the results as the causal effects of media is valid.

5.2 Pre-scheduled press releases

One limitation of the paper is that I could not distinguish press release that are pre-scheduled from the ones that are manually submitted on the spot. To better understand whether this limitation severely bias my results, I exclude all the press releases that are published in the first second of an hour, as the first second would contain a much larger proportion of pre-scheduled press releases. Panel (A) of Table 12 shows the regression results. All the coefficient estimates are almost identical to my results in the previous tables, suggesting that the concern about pre-scheduled press releases is unlikely impacting the results.

[Table 12 here]

6 Conclusion

This paper documents that wire journalists may inefficiently select and report corporate events when they reach their cognitive limits. The problem arises because corporate press releases overcrowd at specific times, forcing journalists to process a lot of information quickly. During these busy times, the exogenous variation of how visually salient a press release stays on the journalists' computer screen affects the amount of media coverage.

This paper finds strong support that the salience of press releases is affected by exogenous shocks that most directly impact wire journalists. This unique setting thus allows me to study the causal effect of media coverage on other market participants. I find that media coverage significantly increases the information acquisition of investors, and the effect also exists in investors who are less likely inattentive. The increased information demand from institutional investors also impacts stock analysts, who also issue more earning forecasts for media-covered events. The effect is most prominent for resource-constrained analysts and becomes stronger in firms with higher institutional ownership.

The results of the paper are consistent with the rational attention framework in Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016). Sophisticated investors may rationally allocate their learning to media-covered events, which are empirically associated with higher price volatility and possibly more mispricing. Their increased information demand also affects analysts, especially the resource-constrained analysts who would crowdsource their coverage decision based on the clients' needs. The market outcomes are consistent with the tug of war between these two types of investors. The effective spread increases with media coverage only on the event day. On the next two days, the intra-day price range increases with more media coverage while the overall absolute return is unchanged.

As new data and technology brings more information to the financial market, a division of labor between different information intermediaries, including media, analysts, and investors themselves, could greatly improve the efficiency to process new information. However, as different information intermediaries become more intertwined, the inefficiency from one member could now affect others as well. This paper provides novel evidence that inefficiency in the wire media could also impact other information channels and cause large market reaction. Moreover, I find that the equilibrium forces do not necessarily reverse such exogenous shift in attention. Even those more sophisticated market participants, who suffer less from inattention, might find it more rewarding to chase the media-covered events.

References

Ahern, K. R., and D. Sosyura. 2015. Rumor has it: Sensationalism in financial media. *Rev. Financ. Stud.* 28:2050–93.

Altınkılıç, O., V. S. Balashov, and R. S. Hansen. 2013. Are analysts' forecasts informative to the general public? *Management Science* 59:2550–65.

Barber, B. M., and T. Odean. 2008. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies* 21:785–818.

Ben-Rephael, A., Z. Da, and R. D. Israelsen. 2017. It depends on where you search: Institutional investor attention and underreaction to news. *Review of Financial Studies* 30:3009–47.

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

Bradley, D., J. Clarke, S. Lee, and C. Ornthanalai. 2014. Are analysts recommendations informative? intraday evidence on the impact of time stamp delays. *The Journal of Finance* 69:645–73.

Brennan, M. J., N. Jegadeesh, and B. Swaminathan. 1993. Investment analysis and the adjustment of stock prices to common information. *Review of Financial Studies* 6:799–824.

Brennan, M. J., and A. Subrahmanyam. 1995. Investment analysis and price formation in securities markets. *Journal of Financial Economics* 38:361–81.

Brown, L. D., A. C. Call, M. B. Clement, and N. Y. Sharp. 2015. Inside the "black box" of Sell-Side financial analysts. *Journal of Accounting Research* 53:1–47.

Chen, H., L. Cohen, U. Gurun, D. Lou, and C. Malloy. 2017. IQ from IP: Simplifying search in portfolio choice. *Working paper* .

Cohen, L., and A. Frazzini. 2008. Economic links and predictable returns. *Journal of Finance* 63:1977–2011.

Da, Z., J. Engelberg, and P. Gao. 2011. In search of attention. *Journal of Finance* 66:1461–99.

Dellavigna, S., and J. M. Pollet. 2009. Investor inattention and friday earnings announcements. *Journal of Finance* 64:709–49.

Dougal, C., J. Engelberg, D. García, and C. A. Parsons. 2012. Journalists and the stock market. *Review of Financial Studies* 25:639–79.

Drake, M. S., N. M. Guest, and B. J. Twedt. 2014. The media and mispricing: The role of the business press in the pricing of accounting information. *The Accounting Review* 89:1673–701.

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

Fang, L. H., J. Peress, and L. Zheng. 2014. Does media coverage of stocks affect mutual funds' trading and performance? *Review of Financial Studies* 27:3441–66.

Fedyk, A. 2018. Front page news: The effect of news positioning on financial markets. *Working paper* .

Gleason, C. A., and C. M. Lee. 2003. Analyst forecast revisions and market price discovery. *The Accounting Review* 78:193–225.

Goldstein, I., and L. Yang. 2015. Information diversity and complementarities in trading and information acquisition. *Journal of Finance* 70:1723–65.

Goyenko, R. Y., C. W. Holden, and C. A. Trzcinka. 2009. Do liquidity measures measure liquidity? *J. financ. econ.* 92:153–81.

Gurun, U. G., and A. W. Butler. 2012. Don't believe the hype: Local media slant, local advertising, and firm value. *Journal of Finance* 67:561–98.

Hillert, A., H. Jacobs, and S. Müller. 2014. Media makes momentum. *Review of Financial Studies* 27:3467–501.

Hirshleifer, D., P.-H. Hsu, and D. Li. 2013. Innovative efficiency and stock returns. *Journal of Financial Economics* 107:632–54.

Hoberg, G., and G. Phillips. 2016. Text-Based network industries and endogenous product differentiation. *Journal of Political Economics* 124:1423–65.

Hong, H., and J. D. Kubik. 2003. Analyzing the analysts: Career concerns and biased earnings forecasts. *Journal of Finance* 58:313–51.

Hong, H., T. Lim, and J. C. Stein. 2000. Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *Journal of Finance* 55:265–95.

Ivković, Z., and N. Jegadeesh. 2004. The timing and value of forecast and recommendation revisions. *Journal of Financial Economics* 73:433–63.

Kacperczyk, M., and A. Seru. 2007. Fund manager use of public information: New evidence on managerial skills. *Journal of Finance* 62:485–528.

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

Kelly, B., and A. Ljungqvist. 2012. Testing asymmetric-information asset pricing models. *The Review of Financial Studies* 25:1366–413.

Lee, C. M. C., P. Ma, and C. C. Y. Wang. 2015. Search-based peer firms: Aggregating investor perceptions through internet co-searches. *Journal of Financial Economics* 116:410–31.

Livnat, J., and R. R. Mendenhall. 2006. Comparing the post–earnings announcement drift for surprises calculated from analyst and time series forecasts. *Journal of Accounting Research* 44:177–205.

Ljungqvist, A., F. Marston, L. T. Starks, K. D. Wei, and H. Yan. 2007. Conflicts of interest in sell-side research and the moderating role of institutional investors. *Journal of Financial Economics* 85:420–56.

Loughran, T., and B. McDonald. 2017. The use of EDGAR filings by investors. *Journal of Behavioral Finance* 18:231–48.

Menzly, L., and O. Ozbas. 2010. Market segmentation and cross-predictability of returns.

Journal of Finance 65:1555–80.

Neuhierl, A., A. Scherbina, and B. Schlusche. 2013. Market reaction to corporate press releases. *Journal of Financial and Quantitative Analysis* 48:1207–40.

Peng, L., and W. Xiong. 2006. Investor attention, overconfidence and category learning. *Journal of Financial Economics* 80:563–602.

Peress, J. 2014. The media and the diffusion of information in financial markets: Evidence from newspaper strikes. *Journal of Finance* 69:2007–43.

Solomon, D. H. 2012. Selective publicity and stock prices. *Journal of Finance* 67:599–638.

Solomon, D. H., and E. F. Soltes. 2012. Managerial control of business press coverage. *Working paper* .

Tetlock, P. C. 2007. Giving content to investor sentiment: The role ogf media in the stock market. *Journal of Finance* 62:1139–68.

———. 2010. Does public financial news resolve asymmetric information? *Review of Financial Studies* 23:3520–57.

———. 2014. Information transmission in finance. *Annual Review of Financial Economics* 6:365–84.

Womack, K. L. 1996. Do brokerage analysts' recommendations have investment value? *Journal of Finance* 51:137–67.

Figure 1: Information production around press releases

The figure plots the percentage of media articles, analyst estimates, and EDGAR searches that are produced on different days relative to the most recent press release from the related firm. The sample includes all the news articles from Dow Jones Newswire, analyst estimates from I/B/E/S, and EDGAR searches from 2004 to 2017 (EDGAR data ends at June 2017). The sample only includes the firms that are in the RavenPack database and CRSP/Compustat Universe. The x-axis in the figure represents the number of days after the most recent press release, and the y-axis represents the percentage of observations.

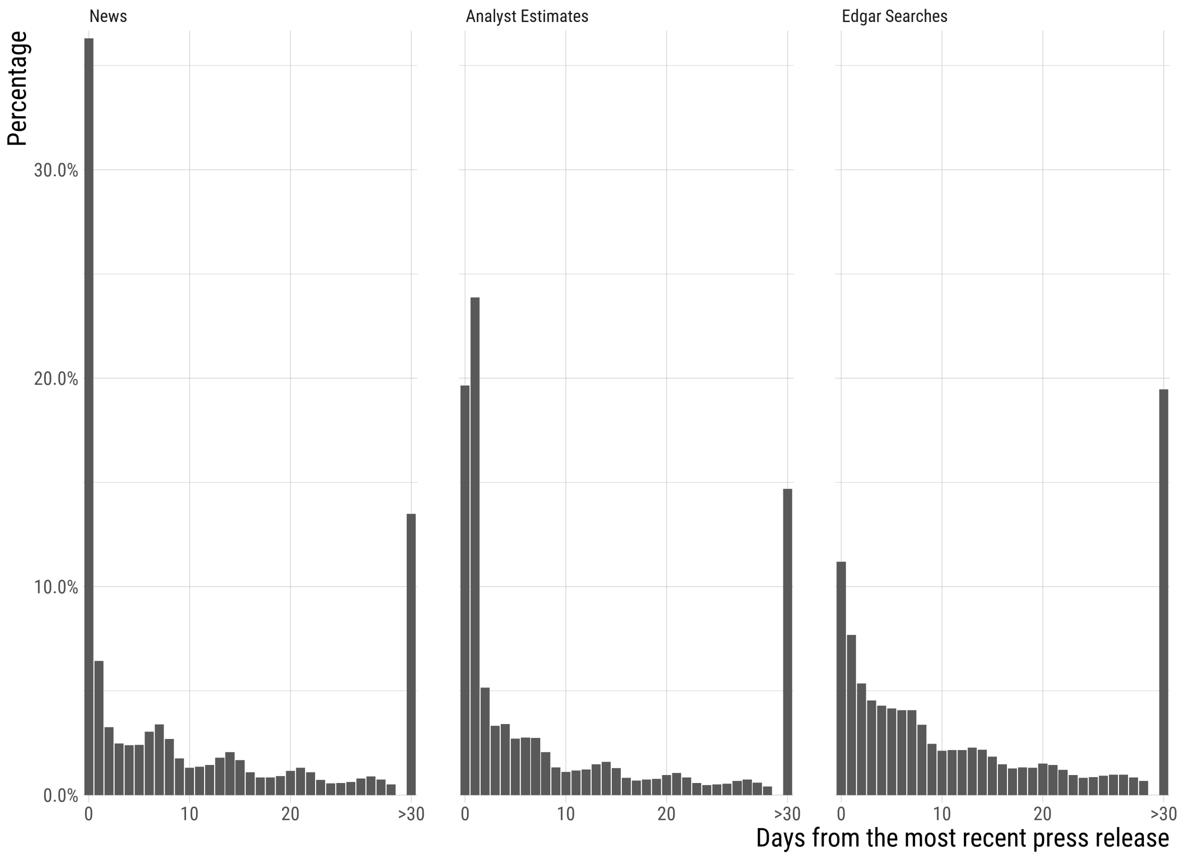
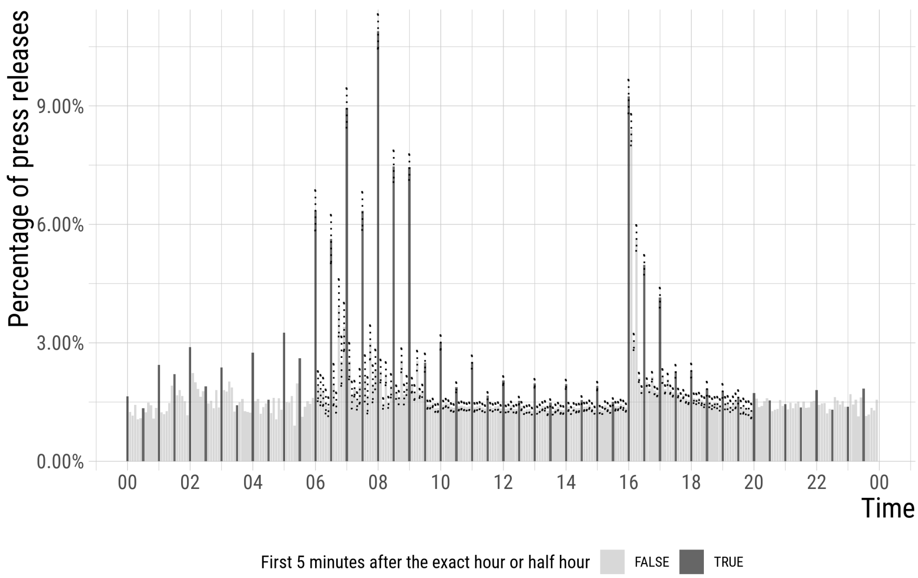


Figure 2: Press release publication time

The figure below plots the distribution of press releases publication time within a day or an hour. Panel (a) plots the average percentage of press releases published in different 5-minute intervals within a day. I split a 24-hour day into 288 5-minute intervals, and for each firm, I calculate the percentage of press releases published in each 5-minute bin. Then for each 5-minute interval I calculate the average percentage across 9,386 firms. The x-axis denotes the publication time, and the y-axis denotes the percentage. Each bar represents the average percentage of press releases published in that 5-minute interval. I also plot the 95% confidence intervals (dashed line) of the group means for hours between 6AM and 8PM. The darker bars indicate the first five minutes of each half hour, or minutes [0, 5) and [30, 34) of each hour, and the lighter bars represent the rest of 5-minute intervals. Panel (b) plots the average percentage of press releases published in each minute. The calculation method is similar as in Panel (a). The error bars represent the 95% confidence intervals.

(a) by 5-minute bins



(b) by minute

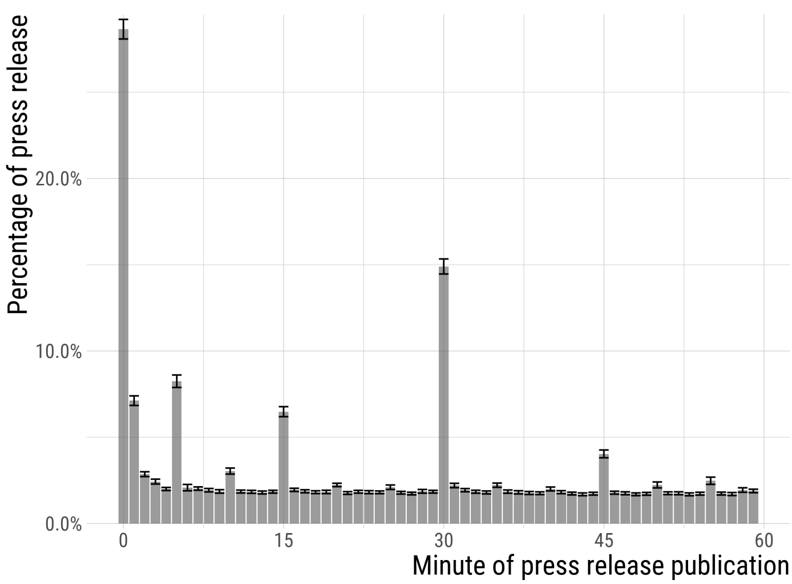


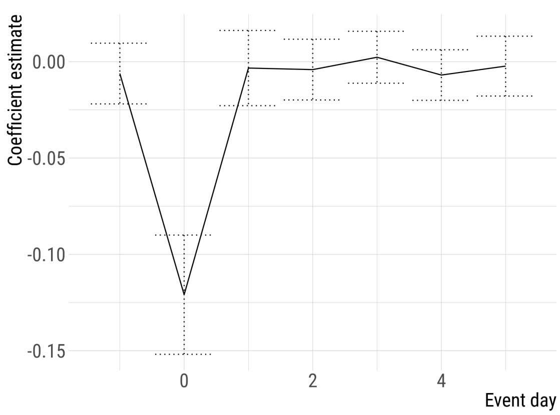
Figure 3: Coefficient estimates for different event days

This figure plots the coefficient estimates for β in the following regression

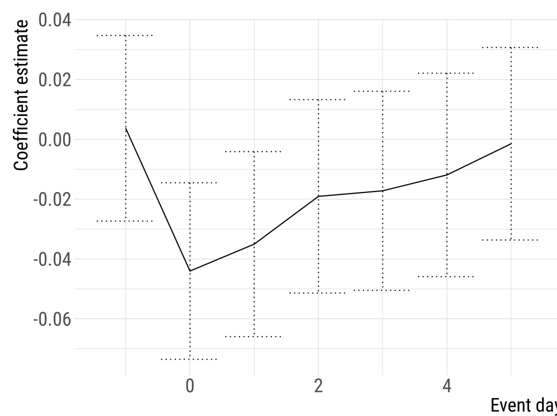
$$Y_{ijt} = \beta \log(NPRA_j + 1) + \boldsymbol{\alpha} + \varepsilon_{ijt}$$

The sample contains all the press releases that are published in the first 10 seconds in 7-9AM and 4PM. For the press release j, NPRA measures the number of following press releases that are published in the next 30 seconds. $\boldsymbol{\alpha}$ includes the firm-year fixed effects, hour-date fixed effects, and detailed press release topic fixed effects. In Figure (a), the dependent variable is the abnormal number of media coverage from Dow Jones Newswire on day t. In Figure (b), the dependent variable is the number of abnormal EDGAR searches on firm i on day t. In Figure (c), the dependent variable is the abnormal number of analysts that issue earning forecasts for firm i on day t. In Figure (d), the dependent variable is the abnormal turnover on day t. In each figure, I separately estimate the regression for event days from -1 to 5, where day 0 is the press release day. The standard errors are clustered by firm and date, and the error bars represent the 95% confidence intervals.

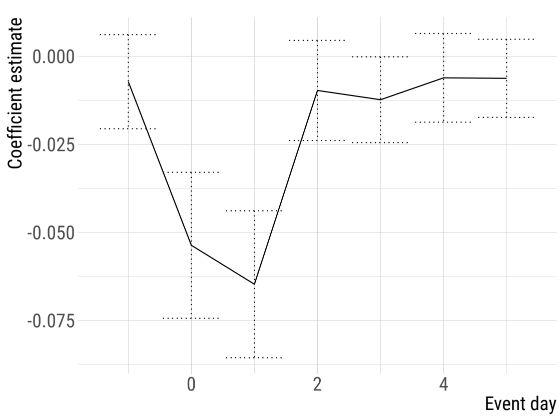
(a) AbnNews



(b) AbnEdgar



(c) AbnAnalyst



(d) AbnTurnover

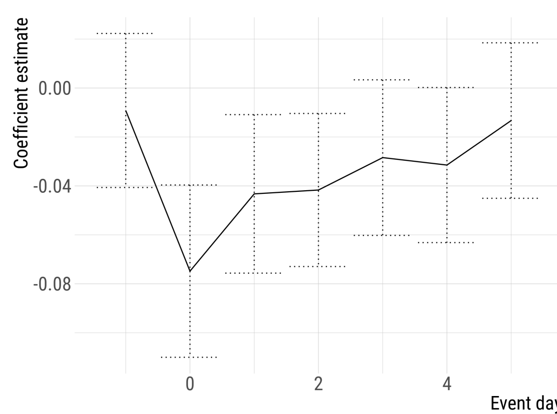


Table 1: Summary statistics

	N	Mean	Std Dev	10th	50th	90th
Panel A: Press release issuance						
Full sample						
# of PRs	738,196					
# of unique firms	8,756					
# of PRs per firm-year	59,035	12.50	9.08	3.00	11.00	23.00
# of PRs per day	2,958	242.00	113.16	113.70	226.00	400.00
Issued in the first 10 seconds of each hour						
# of PRs	131,683					
# of unique firms	7,503					
# of PRs per firm-year	34,532	3.81	3.97	1.00	2.00	8.00
# of PRs per day	2,958	43.04	21.97	19.00	39.00	72.00
Issued in the first 10 seconds of 7-9AM and 4PM						
# of PRs	80,246					
# of unique firms	6,560					
# of PRs per firm-year	24,766	3.24	3.43	1.00	2.00	7.00
# of PRs per day	2,949	26.35	15.93	10.00	23.00	48.00
Panel B: Sample summary statistics						
Issued in the first 10 seconds of each hour						
Age	131376	20.54	16.06	5.00	15.00	48.00
Q	125557	2.05	1.52	0.97	1.52	3.81
AT	130417	21006.78	86222.38	56.95	1093.07	32476.00
ESS	80246	55.35	7.90	50.00	50.00	69.00
NWord	131683	11.56	4.30	7.00	11.00	17.00
NPRA	131683	36.49	31.95	4.00	27.00	84.00
log(NPRA + 1)	131683	3.17	1.11	1.61	3.33	4.44
News	131683	1.75	2.96	0.00	0.00	5.00
AbnNews	131683	0.34	0.82	-0.46	-0.03	1.59
AbnNews_Novel	131683	0.25	0.72	-0.39	-0.06	1.43
AbnNews_Flash	131683	0.30	0.75	-0.41	-0.05	1.51
AbnNews_Full	131683	-0.05	0.44	-0.43	-0.13	0.58
News_Dummy	131683	0.48	0.50	0.00	0.00	1.00
DJPR	131683	0.85	0.36	0.00	1.00	1.00
Issued in the first 10 seconds of 7-9AM and 4PM						
NPRA	80246	51.07	32.24	13.00	46.00	96.00
log(NPRA + 1)	80246	3.71	0.79	2.64	3.85	4.57
AbnNews	80246	0.41	0.83	-0.43	0.00	1.67
AbnEdgar_Total	75673	1.82	1.54	0.00	1.94	3.86
AbnEdgar_Human	75673	0.35	0.75	-0.34	0.03	1.41

Continued on next page

Table 1 – *Continued from previous page*

	N	Mean	Std Dev	10th	50th	90th
AbnEdagr_Crawler	75667	1.65	1.56	0.00	1.63	3.79
AbnEdgar_Exist	75673	-0.48	0.99	-1.83	-0.16	0.51
AbnEdgar_Ins1	75673	-0.85	1.22	-2.53	-0.64	0.20
AbnEdgar_Ins2	75673	-0.69	1.20	-2.35	-0.28	0.48
AbnAnalyst	80246	0.18	0.62	-0.30	-0.03	1.20
AbnAnalyst_MoreAccu	80246	0.04	0.59	-0.53	-0.06	0.94
AbnAnalyst_LessAccu	80246	0.01	0.56	-0.57	-0.06	0.89
AbnAnalyst_MoreCons	80246	0.16	0.73	-0.49	-0.03	1.35
AbnAnalyst_LessCons	80246	-0.18	0.31	-0.61	-0.13	0.01
AbnTurnover	80117	0.02	0.80	-0.79	0.00	0.96
AbnSpread_EW	41832	1.11	11.48	0.64	0.97	1.54
AbnSpread_VW	41832	1.04	0.51	0.58	0.92	1.63
$\|CAR\|$	80142	2.84	5.35	0.21	1.36	6.65
Range	80142	5.09	5.62	1.33	3.53	10.35
% Priv	80246	0.04	0.04	0.00	0.03	0.09
log(NPRPriv + 1)	80246	0.83	0.71	0.00	0.69	1.79

Table 2: Blink effect: press releases' on-screen time and media coverage

This table shows the blink effect: as a press release is followed by more new releases, it stays on the screen for shorter time and ultimately receives less media coverage. I estimate the following regression

$$AbnNews_{ijt} = \beta \log(NPRA_j + 1) + \boldsymbol{\alpha} + \varepsilon_{ijt}$$

The dependent variable is the abnormal media coverage, defined as log of 1 plus the number of news articles that cover the press release issuing firm on the issuing day, minus log of 1 plus the average number of news articles per day for the firm in the past 60 days. The independent variable, NPRA, measures the number of new press releases published in the top 4 press release wires after the press release j in the next 30 seconds. This is the proxy for the on-screen time of each press release. The regression controls for the fixed effects including: date-hour, firm, firm-year, broad topic, and/or detailed topic. The regression sample contains all the press releases published in the first 10 seconds of each hour. Standard errors are clustered by firm and day, and t-statistics are reported in the parentheses. Coefficients marked with ***, **, and * are significant at the 1%, 5%, and 10% level, respectively. Detailed variable definitions can be found in Table 13 in the Appendix.

	AbnNews				
	(1)	(2)	(3)	(4)	(5)
log(NPRA + 1)	-0.134^{***}	-0.132^{***}	-0.163^{***}	-0.099^{***}	-0.093^{***}
	(−12.003)	(−12.487)	(−10.972)	(−8.048)	(−7.737)
Date-Hour FE	Y	Y	Y	Y	Y
Firm FE		Y			
Firm-Year FE			Y	Y	Y
Broad Topic FE				Y	
Detailed Topic FE					Y
Observations	131,683	131,683	131,683	131,683	131,683
Adjusted R^2	0.182	0.323	0.350	0.489	0.509

Table 3: Covariate balancing test

This table shows that the observable firm and event characteristics do not correlate with NPRA, which measures the number of new press releases published in the top 4 press release wires after the press release j in the next 30 seconds. The dependent variable for Columns (1)-(7) is log of 1 plus NPRA, and for Column (8) is the log of 1 plus the number of media articles on the Dow Jones Newswire covering the firm on the event day. The regression sample includes all the press releases published in the first 10 seconds of each hour. Standard errors are clustered by firm and day, and t-statistics are reported in the parentheses. Coefficients marked with ***, **, and * are significant at the 1%, 5%, and 10% level, respectively. Detailed variable definitions can be found in Table 13.

	log(NPRA + 1)							News
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Q	0.002						0.002	0.010***
	(0.773)						(0.538)	(3.803)
log(AT)		0.000					−0.002	0.015**
		(0.055)					(−0.223)	(1.974)
log(Age + 1)			0.006				−0.026	0.048**
			(0.326)				(−1.123)	(2.123)
log(ESS)				−0.023			−0.016	0.634***
				(−1.196)			(−0.817)	(20.311)
log(NWords)					−0.011		−0.010	−0.318***
					(−1.567)		(−1.399)	(−25.999)
log(NPRInd + 1)						0.006	0.006	0.012**
						(1.302)	(1.186)	(2.122)
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y
Date FE	Y	Y	Y	Y	Y	Y	Y	Y
Hour FE	Y	Y	Y	Y	Y	Y	Y	Y
Broad Topic FE	Y	Y	Y	Y	Y	Y	Y	Y
F-stat							1.027	200
p-value							0.405	<0.001
Observations	125,557	130,417	131,376	131,683	131,683	131,683	125,557	125,557

Table 4: Falsification tests of the blink effect

This table shows the falsification tests of the blink effect. The dependent variable in Column (1) is a dummy variable that equals to 1 if the automated algorithm republishes the press release on Dow Jones Newswire. The dependent variable in Columns (2) is the abnormal news measure on the day *before* the event (press release) day. The dependent variables in Column (3) - (7) are the same abnormal news measures used in Table 2. Column (3) and (4) show regression results using samples from different minutes. Column (3) includes all the press releases published in the first 10 seconds in the 31st minute of each hour. Column (4) includes all the press published in the first 10 seconds in all other minutes except for the 1st and the 31st minutes. Column (5) and (6) show regression results using samples from different hours. Column (5) includes all the press releases published in the first 10 seconds in 7-9AM and 4PM, and Column (6) includes all the press releases published in the first 10 seconds of all other hours. In Column (7), I sort the press releases by NPRA into quintiles, and Q2 - Q5 in the independent variables are quintile dummies, where Q5 represents the highest-NPRA quintile. Standard errors are clustered by firm and day, and t-statistics are reported in the parentheses. Coefficients marked with ***, **, and * are significant at the 1%, 5%, and 10% level, respectively. Detailed variable definitions can be found in Table 13.

	DJPR	AbnNews					
		t-1	min 30	other minutes	7-9AM & 4PM	Other Hours	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log(NPRA+1)	−0.005	−0.003	−0.085***	0.051***	−0.112***	−0.055	−0.030
	(−0.714)	(−0.403)	(−3.781)	(17.157)	(−8.489)	(−0.922)	(−1.195)
log(NPRA+1) x Q2							−0.024
							(−0.761)
log(NPRA+1) x Q3							−0.062**
							(−2.043)
log(NPRA+1) x Q4							−0.097***
							(−3.106)
log(NPRA+1) x Q5							−0.101***
							(−2.884)
Date-Hour FE	Y	Y	Y	Y	Y	Y	Y
Firm-Year FE	Y	Y	Y	Y	Y	Y	Y
Detailed Topic FE	Y	Y	Y	Y	Y	Y	Y
Observations	131,683	131,683	69,185	252,532	80,246	51,437	131,683
Adjusted R^2	0.303	0.160	0.465	0.506	0.547	0.141	0.509

Table 5: Robustness of the blink effect

This table shows that the results in Table 2 are robust to news measures, sample selection, and functional form of the dependent variable. Columns (1) - (3) re-estimate Equation 1 using different news measures. Column (1) uses non-duplicated news only, whose `ENS` score is 100 in RavenPack; Column (2) uses flash news, which only contains a headline; Column (3) uses full news, which contains a headline and at least one paragraph. Columns (4) - (8) re-estimate Equation 1 using different samples. Dependent variables in Columns (4) - (8) are the abnormal number of media articles covering the firm on the press release day. From all the press releases published in the first 10 seconds in each hour: Column (4) excludes the releases published in the first second, Column (5) excludes firm-years where the total number of press releases is above the 75th-percentile, Column (6) excludes firm-years where there is only one release, Column (7) excludes date-hours where there is only one release. Column (8) includes all the press releases published in the first 30 seconds of each hour. Columns (9)-(11) change the functional form of the dependent variable. Column (9) uses the log(News + 1) as the dependent variable, rather than the abnormal news measure. Column (10) uses the raw level of the abnormal news, rather than the log change measure, as the dependent variable. Column (11) uses a dummy variable that equals to 1 if there are any news coverage on the event day, as the dependent variable. The independent variable, NPRA, measures the number of new press releases published in the top 4 press release wires after the press release j in the next 30 seconds. The standard errors in all regressions are clustered by firm and date, and t-statistics are reported in the parentheses. Coefficients marked with ***, **, and * are significant at the 1%, 5%, and 10% level, respectively. Detailed variable definitions can be found in Table 13.

Robustness to	news measures			sample					functional form		
	novel	flash	full	excl. 1st second	excl. top 25% of PR issuers	excl. if only 1 PR per firm-year	excl. if only 1 PR per date-hour	first 30 seconds	no abnormal	no log	dummy
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
log(NPRA + 1)	−0.085***	−0.085***	−0.041***	−0.091***	−0.085***	−0.093***	−0.093***	−0.045***	−0.094***	−0.371***	−0.045***
	(−8.407)	(−8.154)	(−4.931)	(−5.625)	(−5.649)	(−8.206)	(−8.330)	(−6.749)	(−7.991)	(−7.301)	(−6.161)
Date-Hour FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Firm-Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Detailed Topic FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	131,683	131,683	131,683	101,115	103,673	120,424	116,450	188,981	131,683	131,683	131,683
Adjusted R^2	0.518	0.564	0.207	0.506	0.507	0.526	0.530	0.516	0.483	0.395	0.416

Table 6: Media coverage and EDGAR requests

This table shows that media coverage affect the number of EDGAR requests. I estimate the following regression

$$AbnEdgar_{ijt} = \beta \log(NPRA_j + 1) + \boldsymbol{\alpha} + \varepsilon_{ijt}$$

The dependent variables in Columns (1) - (6) are the abnormal number of EDGAR requests, defined as the log of 1 plus the number of requests made on the press release day minus the log of 1 plus the average number of EDGAR requests per day in the past 60 days. I exclude EDGAR searches which are on the index page and have a server code above 300. Column (1) counts all the EDGAR searches in the dependent variable. Column (2) and Column (3) only include requests from humans and web crawlers, respectively. The web crawlers are defined as IP addresses that search for more than 50 companies per day. The dependent variable in Column (4) count the number of human requests from IP addresses which have searched the same firm in teh previous month. The dependent variables in Columns (5) and (6) count the number of web requests from financial institutions. Online Appendix B shows the details steps for identifying IP addresses from financial institutions. Column (6) shows the second stage results of 2SLS estimation, where the first stage results are the same as Column (5) of Table 4. The dependent variables in Columns (7) and (8) are the number of abnormal EDGAR searches on the day after the press release. Column (8) shows the OLS regression of abnormal EDGAR searches on abnormal media coverage. The regression sample includes all the press releases published in the first 10 seconds of 7-9AM and 4PM. Standard errors are clustered by firm and day, and t-statistics are reported in the parentheses. Coefficients marked with ***, **, and * are significant at the 1%, 5%, and 10% level, respectively. Detailed variable definitions can be found in Table 13.

	AbnEdgar							
	day 0						day 1	
	total	human	crawler	existing searcher	Institution		2SLS	OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(NPRA+1)	-0.024	-0.044^{***}	-0.022	-0.046^{***}	-0.057^{***}	-0.043^{**}		
	(-1.638)	(-3.117)	(-1.396)	(-2.584)	(-3.357)	(-2.454)		
$\widehat{\text{AbnNews}}$							0.301^{**}	
							(2.475)	
AbnNews								0.144^{***}
								(15.594)
Date-hour FE	Y	Y	Y	Y	Y	Y	Y	Y
Firm-Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Detailed Topic FE	Y	Y	Y	Y	Y	Y	Y	Y
Observations	75,667	75,667	75,667	75,667	75,667	75,667	75,667	75,667
Adjusted R^2	0.834	0.287	0.804	0.398	0.567	0.517	0.286	0.297

Table 7: News coverage and analyst forecasts

This table shows that media coverage impacts the number of analysts who issue earning forecasts for the same firm. I estimate the following regression

$$AbnAnalyst_{ijt} = \beta \log(NPRA_j + 1) + \boldsymbol{\alpha} + \varepsilon_{ijt}$$

The dependent variable in Column (1) is abnormal number of analyst forecasts on day 0, defined as log of 1 plus the number of analysts who issued any forecasts for the firm on the event day minus the log of 1 plus the average of daily number of analysts issuing forecasts for the firm in the past 60 days. The dependent variable in Column (2) is the abnormal number of analyst forecasts issued on day 1. The dependent variables for Columns (3) - (8) are cumulative abnormal number of analyst forecasts issued on day 0 and 1. In Columns (3) - (5), I group analysts by their relative accuracy in the previous three years. Column (3) counts forecasts from more accurate (relative accuracy above median) analysts, Column (4) counts forecasts from less accurate analysts, and Column (5) uses the difference of the two as the dependent variable. In Columns (6) - (9), I instead group analysts into two groups based on the number of companies they cover in the previous year. Column (6) only counts forecasts from analysts who covered more than median number of firms in the previous year, Column (7) only counts forecasts from analysts who cover less than the median number, and Column (8) uses the difference of the dependent variables in Columns (6) and (7) as the dependent variable. Column (9) and (10) shows the second stage results of the 2SLS estimation, where the first stage estimation is shown in Column (5) of Table 4. The dependent variables in Column (9) and (10) are the abnormal number of analyst forecasts issued on day 1. Column (10) shows OLS regression result. The regression sample includes all the press releases published in the first 10 seconds of 7-9AM and 4PM. Standard errors are clustered by firm and day, and t-statistics are reported in the parentheses. Coefficients marked with ***, **, and * are significant at the 1%, 5%, and 10% level, respectively. Detailed variable definitions can be found in Table 13.

	AbnAnalyst									
Analyst estimates on	day 0	day 1	day 0-1						day 1	
			Accuracy			# of followed firms			2SLS	OLS
			>50th	<50th	dif	>50th	<50th	dif		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
log(NPRA+1)	−0.060***	−0.071***	−0.068***	−0.062***	−0.006	−0.090***	−0.013**	−0.077***		
	(−5.960)	(−7.087)	(−6.571)	(−6.906)	(−0.758)	(−7.345)	(−2.501)	(−7.308)		
$\widehat{\text{AbnNews}}$									0.589***	
									(7.220)	
AbnNews										0.277***
										(27.346)
Date-hour FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Firm-Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Detailed Topic FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	80,246	80,246	80,246	80,246	80,246	80,246	80,246	80,246	80,246	80,246
Adjusted R^2	0.404	0.446	0.430	0.477	0.284	0.468	0.545	0.479	0.425	0.518

Table 8: Media coverage, analyst forecasts, and institutional holdings

This table shows the effects of media coverage on analyst forecasts are stronger for firms with higher institutional ownership or more institutional investors. The dependent variables in Columns (1), (2), (5), and (6) are the cumulative abnormal number of analyst forecasts issued on day 0 and 1. The dependent variables in Columns (3), (4), (7), and (8) are the abnormal media coverage on day 0. Columns (1), (2), (5), and (6) show the second stage results of 2SLS estimations, while the other four columns show the first stage regression results. Columns (1) and (3) include observations where the institutional ownership is above median in the previous year, while Columns (2) and (4) use the below-median sample. Columns (5) - (8) separate the sample by the number of institutional investors in the previous year. The regression sample includes all the press releases published in the first 10 seconds of 7-9AM and 4PM. Standard errors are clustered by firm and day, and t-statistics are reported in the parentheses. Coefficients marked with ***, **, and * are significant at the 1%, 5%, and 10% level, respectively. Detailed variable definitions can be found in Table 13.

	AbnAnalyst		AbnNews		AbnAnalyst		AbnNews	
	by % of institutional holdings				by # of institutional investors			
	<50th	>50th	<50th	>50th	<50th	>50th	<50th	>50th
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\widehat{\text{AbnNews}}$	0.519***	0.962***			0.558***	0.914***		
	(3.309)	(5.164)			(3.818)	(4.866)		
log(NPRA+1)			−0.107***	−0.125***			−0.110***	−0.121***
			(−5.087)	(−5.694)			(−5.604)	(−5.428)
Date-hour FE	Y	Y	Y	Y	Y	Y	Y	Y
Firm-Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Detailed Topic FE	Y	Y	Y	Y	Y	Y	Y	Y
Second FE	Y	Y	Y	Y	Y	Y	Y	Y
Observations	40,164	40,082	40,164	40,082	40,191	40,055	40,191	40,055
Adjusted R^2	0.460	0.484	0.529	0.561	0.468	0.491	0.535	0.559

Table 9: Media coverage and market reaction

This table shows that media coverage significantly affects trading volume, announcement returns, effective spread, and price ranges. The dependent variables in Columns (1) - (5) are from the first trading day after the press releases, and the dependent variables in Columns (6) - (10) are average measures (except for CAR) or cumulative measures (CAR) in days 1-2, the second and third trading days after the press release. The dependent variables for Columns (1) and (6) are abnormal turnover, defined as log of 1 plus the turnover minus the log of 1 plus the average daily turnover in the previous 60 trading days. The dependent variables in Columns (2), (3), (7) and (8) are abnormal effective spread measures, with equal-weighted measures in Columns (2) and (7) and value-weighted measures by dollar amount in Columns (3) and (8). The dependent variables in Columns (4) and (9) are the absolute value of cumulative abnormal returns, where abnormal returns are calculated by subtracting the CRSP value-weighted index return on the same day. The dependent variables in Columns (5) and (10) are the price ranges, defined as the log of the daily highest price minus the log of the daily lowest prices. The regression sample includes all the press releases published in the first 10 seconds of 7-9AM and 4PM. Standard errors are clustered by firm and day, and t-statistics are reported in the parentheses. Coefficients marked with ***, **, and * are significant at the 1%, 5%, and 10% level, respectively. Detailed variable definitions can be found in Table 13.

	Day 0					Days 1-2				
	AbnTurnover	AbnSpread		|CAR|	Range	AbnTurnover	AbnSpread		|CAR|	Range
		EW	VW				EW	VW		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
log(NPRA+1)	−0.077***	−0.031***	−0.035***	−0.255***	−0.272***	−0.043***	−0.001	0.004	−0.001	−0.077**
	(−4.928)	(−3.071)	(−2.600)	(−4.317)	(−4.277)	(−3.333)	(−0.090)	(0.409)	(−0.660)	(−2.047)
Firm-Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Date-hour FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Detailed Topic FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	80,145	41,835	41,835	80,145	80,145	80,162	41,464	41,463	80,162	80,162
Adjusted R^2	0.839	0.219	0.173	0.304	0.512	0.868	0.217	0.173	0.068	0.621

Table 10: Media coverage and delayed response ratio

This table shows media coverage has insignificant effects on price efficiency. The dependent variables are delayed response ratios, as used in Dellavigna and Pollet (2009). The delayed response ratio over days [0, X] is calculated as $R^{(2,X)}/R^{(0,X)}$, where $R^{(2,X)}$ is the cumulative abnormal returns over days [2, X], and $R^{(0,X)}$ is the cumulative abnormal returns over the period [0, X]. Columns (1) - (6) show results with different period length. The regression sample includes all the press releases published in the first 10 seconds of 7-9AM and 4PM. Standard errors are clustered by firm and day, and t-statistics are reported in the parentheses. Coefficients marked with ***, **, and * are significant at the 1%, 5%, and 10% level, respectively. Detailed variable definitions can be found in Table 13.

	Delayed response ratio					
	[2,5]	[2,15]	[2,30]	[2,45]	[2,60]	[2,75]
	(1)	(2)	(3)	(4)	(5)	(6)
log(NPRA+1)	−0.025	0.015	0.054	0.035	0.004	0.020
	(−0.388)	(0.298)	(1.267)	(0.962)	(0.138)	(0.695)
Firm-Year FE	Y	Y	Y	Y	Y	Y
Date-hour FE	Y	Y	Y	Y	Y	Y
Detailed Topic FE	Y	Y	Y	Y	Y	Y
Observations	80,130	80,130	80,130	80,130	80,130	80,130
Adjusted R^2	0.008	0.001	0.012	0.010	0.017	0.003

Table 11: Blink effect for analysts and investors?

This table shows that investors and analysts unlikely experience similar blink effect. I estimate similar regressions as in previous tables. Both Panel (A) and (B) use all the press releases published in the first 10 seconds of 7-9AM and 4PM. Column (1) uses the abnormal media coverage on day 0 as the dependent variable. Column (2) uses the cumulative abnormal number of EDGAR searches on days 0 and 1 as the dependent variable. Column (3) uses the cumulative abnormal number of analyst forecasts on days 0 and 1 as the dependent variable. Column (4) uses the abnormal turnover on the event day as the dependent variable. Column (5) uses the abnormal spread on the event day as the dependent variable. Column (6) uses the absolute abnormal return on the event day as the dependent variable. Column (7) uses the intraday price range on the event day as the dependent variable. In Panel (A), I include an interaction term, log(NPRA + 1) x % Priv, where % Priv is the percentage of press releases in NPRA that are issued by private firms. In Panel (B), I directly regress the dependent variables on log(NPRPriv + 1), where NPRPriv is the number of press releases issued by private firms in the next 30 seconds. I also control for log(NPRInd + 1), where NPRInd is the number of press releases issued by firms in the same 2-digit SIC industry on the same day. Standard errors are clustered by firm and day, and t-statistics are reported in the parentheses. Coefficients marked with ***, **, and * are significant at the 1%, 5%, and 10% level, respectively. Detailed variable definitions can be found in Table 13.

Panel A: Interaction with the ratio of press releases from private firms							
	AbnNews	AbnEdgar	AbnAnalyst	AbnTurnover	AbnSpread	|CAR|	Range
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log(NPRA+1)	−0.125***	−0.045***	−0.094***	−0.066***	−0.032**	−0.233***	−0.236***
	(−7.669)	(−2.899)	(−6.212)	(−4.274)	(−2.465)	(−3.512)	(−3.176)
log(NPRA+1) x % Priv	0.111	−0.006	−0.007	0.150	0.021	−0.133	−0.868
	(0.479)	(−0.025)	(−0.033)	(0.678)	(0.082)	(−0.124)	(−0.788)
% Priv	−0.432	−0.020	−0.302	−0.593	−0.247	−0.456	2.182
	(−0.583)	(−0.028)	(−0.433)	(−0.825)	(−0.292)	(−0.130)	(0.617)
Firm-Year FE	Y	Y	Y	Y	Y	Y	Y
Date-hour FE	Y	Y	Y	Y	Y	Y	Y
Detailed Topic FE	Y	Y	Y	Y	Y	Y	Y
Observations	80,246	75,667	80,246	80,117	41,832	80,142	80,142
Adjusted R^2	0.547	0.340	0.473	0.277	0.220	0.304	0.517
Panel B: direct test using press releases from private firms							
	AbnNews	AbnEdgar	AbnAnalyst	AbnTurnover	AbnSpread	|CAR|	Range
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log(NPRPriv + 1)	−0.104***	−0.030**	−0.088***	−0.053***	−0.023**	−0.247***	−0.271***
	(−7.625)	(−2.148)	(−6.179)	(−3.774)	(−1.983)	(−3.898)	(−3.889)
log(NPRInd + 1)	0.030***	−0.002	0.035***	−0.004	0.002	−0.026	0.055
	(2.599)	(−0.179)	(2.831)	(−0.342)	(0.231)	(−0.478)	(1.003)
Firm-Year FE	Y	Y	Y	Y	Y	Y	Y
Date-hour FE	Y	Y	Y	Y	Y	Y	Y
Detailed Topic FE	Y	Y	Y	Y	Y	Y	Y
Observations	80,246	75,667	80,246	80,117	41,832	80,142	80,142
Adjusted R^2	0.547	0.339	0.472	0.276	0.219	0.304	0.517

Table 12: Robustness tests

This table shows the robustness of the main results in the paper. The dependent variables are: Column (1), the abnormal media coverage on day 0; Column (2), the cumulative abnormal number of EDGAR searches on days 0 and 1; Column (3), the cumulative abnormal number of analyst forecasts on days 0 and 1. Column (4), the abnormal turnover on the event day. Column (5), the abnormal spread on the event day. Column (6), the absolute abnormal return on the event day. Column (7), the intraday price range on the event day. In Panel (A), the regression sample includes all the press releases issued in the first 10 seconds, excluding the first second, of 7-9AM and 4PM. In Panel (B), the regressions further include a set of fixed effects for the second of the press release publication time. In Panel (C), the regression sample includes all the press releases in the first 10 seconds of 7-9AM and 4PM, excluding observations where the total number of press releases issued by a firm is above the 75th percentile. Standard errors are clustered by firm and day, and t-statistics are reported in the parentheses. Coefficients marked with ***, **, and * are significant at the 1%, 5%, and 10% level, respectively. Detailed variable definitions can be found in Table 13.

Panel A: Drop the 1st seconds							
	AbnNews	AbnEdgar	AbnAnalyst	AbnTurnover	AbnSpread	|CAR|	Range
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log(NPRA+1)	-0.118^{***}	-0.046^{***}	-0.093^{***}	-0.056^{***}	-0.036^{**}	-0.236^{***}	-0.246^{***}
	(-6.654)	(-2.652)	(-5.444)	(-3.122)	(-2.407)	(-2.952)	(-2.834)
Fixed Effects	Y	Y	Y	Y	Y	Y	Y
Observations	62,809	62,809	62,809	62,706	33,907	62,725	62,725
Panel B: Control for second fixed effects							
	AbnNews	AbnEdgar	AbnAnalyst	AbnTurnover	AbnSpread	|CAR|	Range
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log(NPRA+1)	-0.081^{***}	-0.044^{***}	-0.063^{***}	-0.040^{**}	-0.023^{*}	-0.156^{**}	-0.128^{*}
	(-5.064)	(-2.804)	(-4.034)	(-2.403)	(-1.960)	(-2.251)	(-1.692)
Fixed Effects	Y	Y	Y	Y	Y	Y	Y
Observations	80,246	80,246	80,246	80,117	41,832	80,142	80,142
Panel C: Drop the top 25% press release issuing firms							
	AbnNews	AbnEdgar	AbnAnalyst	AbnTurnover	AbnSpread	abs(CAR)	Range
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log(NPRA+1)	-0.112^{***}	-0.045^{***}	-0.080^{***}	-0.065^{***}	-0.026	-0.248^{***}	-0.247^{***}
	(-6.602)	(-2.613)	(-5.172)	(-2.972)	(-1.626)	(-3.031)	(-2.738)
Fixed Effects	Y	Y	Y	Y	Y	Y	Y
Observations	60,586	60,586	60,586	52,883	30,348	60,485	60,485

Appendix A. Variable Definition

Table 13: Variable definitions

Variable	Definition	Source
$NPRA_j$	The number of new press releases published immediately after the press release j in the next 30 seconds. I only include press releases published on the top 4 press release wires, namely, PRNewswire, BusinessWire, GlobeNewswire, Marketwired.	RavenPack PR Edition
$News_{it}$	The total number of news articles covering firm i on day t. The relevance score needs to be 100, meaning that firm i is the main subject of the news article.	RavenPack DJ Edition
$AbnNews_{it}$	Abnormal news coverage. The number is calculated by subtracting the average daily number of media articles covering firm i in the previous 60 days from $News_{it}$	RavenPack DJ Edition
$AbnNews\_Novel_{it}$	From AbnNews, I only include news articles whose novelty score from RavenPack (`ESS`) is 100	RavenPack DJ Edition
$AbnNews\_Flash_{it}$	From AbnNews, I only include news articles whose news type in RavenPack is `NEWS-FLASH`	RavenPack DJ Edition
$AbnNews\_Full_{it}$	From AbnNews, I only include news articles whose news type in RavenPack is `FULL-ARTICLE`	RavenPack DJ Edition
$News\_Dummy_{it}$	A dummy variable that equals to 1 if there is any media coverage on that day	RavenPack DJ Edition
$AbnEdgar\_Total_{it}$	Abnormal number of EDGAR requests about firm i on day t. I exclude requests where `idx` equals 1 (search on the index page) or the server code `code` is above 300. The abnormal measure is calculated as the log of 1 plus the number of searches on day t minus the average daily number of searches in the past 60 days.	SEC EDGAR log
$AbnEdgar\_Human_{it}$	Abnormal number of EDGAR searches from possible human users. To qualify as human, the IP address need to search less than 50 firms in a day.	SEC EDGAR log
$AbnEdgar\_Crawler_{it}$	Abnormal number of EDGAR searches from possible web crawlers. To qualify as a web crawler, the IP address need to search 50 or more firms in a day.	SEC EDGAR log
$AbnEdgar\_Exist_{it}$	Abnormal number of human EDGAR searches from IP addresses which have accessed the filings from the same firm in the previous month.	SEC EDGAR log

Continued on next page

Table 13 – *Continued from previous page*

Variable	Definition	Source
AbnEdgar_Ins1$_{it}$	Abnormal number of EDGAR searches from institutional investors. I identify institutional investors first by matching IP addresses to known institutions which have autonomous system numbers. The IP-ASN organization link file comes from MaxMind. I then search for finance-related words in the names of the institutions. Details can be found in the Online Appendix B	SEC EDGAR log
AbnEdgar_Ins2$_{it}$	Abnormal number of EDGAR searches from institutional investors. I identify institutional investors first by matching IP addresses to known institutions which have autonomous system numbers. The IP-ASN organization link file comes from MaxMind. I then match the names of these institutions to the names of all 13F institutions. Details can be found in the Online Appendix B	SEC EDGAR log
AbnAnalyst$_{it}$	Abnormal number of analyst forecasts issued on day t. For each firm-day, I count the unique number of analysts who issue any earning forecasts for firm i. Then the abnormal measure is calculated as the log of 1 plus the number of analysts issuing any earning forecast for firm i on day t, minus the log of 1 plus the average number of analysts issuing forecasts per day in the past 60 calendar days.	IBES
AbnAnalyst$_{it}$	Abnormal number of analyst forecasts issued on day t. For each firm-day, I count the unique number of analysts who issue any earning forecasts for firm i. Then the abnormal measure is calculated as the log of 1 plus the number of analysts issuing any earning forecast for firm i on day t, minus the log of 1 plus the average number of analysts issuing forecasts per day in the past 60 calendar days.	IBES

Continued on next page

Table 13 – *Continued from previous page*

Variable	Definition	Source
AbnAnalyst_MoreAccu$_{it}$	The abnormal number of analyst forecasts from analysts whose relative forecast accuracy is above median in the previous year. The relative forecast accuracy is constructed following Ljungqvist et al. (2007). For analyst i covering firm k in year t, I first calculate the absolute forecast error using the following steps. (1) get the analysts most recent forecast of year-end EPS issued between Jan. 1 and Jun. 30, (2) calculate the difference with the subsequent realized earnings, (3) scale the difference by previous year-end price. The for all the analysts covering firm k in year t, I re-scale the absolute forecast errors so that the most and least accurate analysts scores one and zero, respectively. Finally, analyst i's relative forecast accuracy in year t is his/her average score across the the stocks he/she covers over years t-2 to t.	IBES
AbnAnalyst_LessAccu$_{it}$	The abnormal number of analyst forecasts from analysts whose relative forecast accuracy is below median in the previous year.	IBES
AbnAnalyst_MoreCons$_{it}$	The abnormal number of analyst forecasts from analysts who cover an above-median number of firms per analyst in the previous year.	IBES
AbnAnalyst_LessCons$_{it}$	The abnormal number of analyst forecasts from analysts who cover a below-median number of firms per analyst in the previous year.	IBES
AbnTurnover	Abnormal turnover, calculated as the log of 1 plus the turnover minus the log of 1 plus the average daily turnover in the past 60 trading days.	CRSP
AbnSpread_EW	Equal-weighted abnormal effective spread. For each trade, I calculate the effect spread by $2\|\log(P_k)-\log(M_k)\|$, where P_k is the price of the trade, and M_k is the mid-point of the consolidated BBO at the time of the trade. I exclude corrected orders, trades with zero price or zero size, trades with condition codes B, G, J, K, L, O, T, W, or Z, and quotes in which the bid-ask spread is negative or greater than 50% of the quote midpoint. Trades on all exchanges are included. Then for each day, I calculate the equal-weighted average effective spread over all the trades. I define abnormal spread as the effective spread on day t over the average daily effective spread in the past 60 trading days.	TAQ
AbnSpread_VW	value-weighted abnormal effective spread, calculated as the dollar-value weighted average of all the effective spread within the day.	TAQ

Continued on next page

Table 13 – *Continued from previous page*

Variable	**Definition**	**Source**
|CAR|	Absolute value of the abnormal return. I calculate abnormal returns by subtracting the CRSP value-weighted index return from the daily raw returns.	CRSP
Range	Daily price range, defined as the log of the daily high price minus the log of the daily low	CRSP
% Priv	The percentage of press releases in NPRA that are issued by private firms.	RavenPack PR Edition
NPRPriv$_{it}$	The number of press releases that are from private firms and published in the next 30 seconds after the press relese j.	RavenPack PR Edition
NPRInd$_{it}$	The total number of press releases that are issued by firms from the same 2-digit SIC industry as firm i, minus 1 (the press release from firm i itself).	RavenPack PR Edition
ESS	Event sentiment score. The score is generated by RavenPack. A group of experts first read and score a sample of stories to determine the direction of impacts (positive or negative) and the degree of different event types. In total they have over 2000 types of events. New articles are then compared to these tagged events to calculate the score. Stories with scores higher than 50 are positive, and lower than 50 are negative.	RavenPack PR Edition
AT	Total asset	Compustat
Q	Tobin's Q, defined as `(at + csho x prcc_f - ceq) / at`	Compustat
Age	The number of years since publication	Compustat, CRSP
NWord	The number of words in the title of a press release	RavenPack PR Edition
DJPR	A dummy variable that equals to one if the automated algorithm from Dow Jones Newswire republishes the press release	RavenPack PR/DJ Edition

Online Appendix A. Description of the press release data

In this section, I briefly introduce the press release data from the RavenPack PR Edition. The data is from 2004 to 2017 (in my paper), and covers a variety of press release wires. Table 14 below shows the total number of press releases by sources.

Table 14: Press release counts by sources

The table below shows the number of press releases from each sources in the RavenPack PR Edition. The data covers 2004 to 2017. The sample in this table is constructed with several data filters: (1) the press release has a relevance score of 100 for the mentioned firm, (2) the mentioned firm has a valid CUSIP number, and (3) the CUSIP number is in the CRSP-Compustat merged database. My final sample only includes press releases in the top 4 press release wires: PR Newswire, Business Wire, Globe Newswire, and Marketwired.

Source	Number of press release
PR Newswire	**734,068**
Business Wire	**507,727**
Globe Newswire	**160,447**
Marketwired	**99,779**
Canadian Newswire	44,245
LSE Regulatory News Service	29,692
Canadian Corporate News	27,161
Hugin Globe Newswire	6,922
DJ Global Press Release Wire	6,812
DGAP News	1,155
JSE News	752
Cision News	433
Actus News	240
OSLO BORS News	236
News Aktuell	122
HK Exchange News	121
Tensid News	93
IRW Press News	15
Karachi Stock Exchange News	12
Realwire	8
Pressetext News	5
BSE News	1

As we can see, the top four press release wires, namely, PR Newswire, Business Wire, Globe Newswire, and Marketwired, contain most of the press releases in the sample. These four wires also take the majority of the market share in the US for public firms (Solomon and Soltes, 2012). For this paper, I restrict the sample to all the press releases from these four major press release wires.

In Table 15, I list all the data cleaning steps that I have taken to construct the final regression sample for this paper. For each step, I also list the number of press releases and number of unique firms in the sample.

Table 15: Sample selection

The table below shows the steps I take to compile the sample of press releases used in this paper.

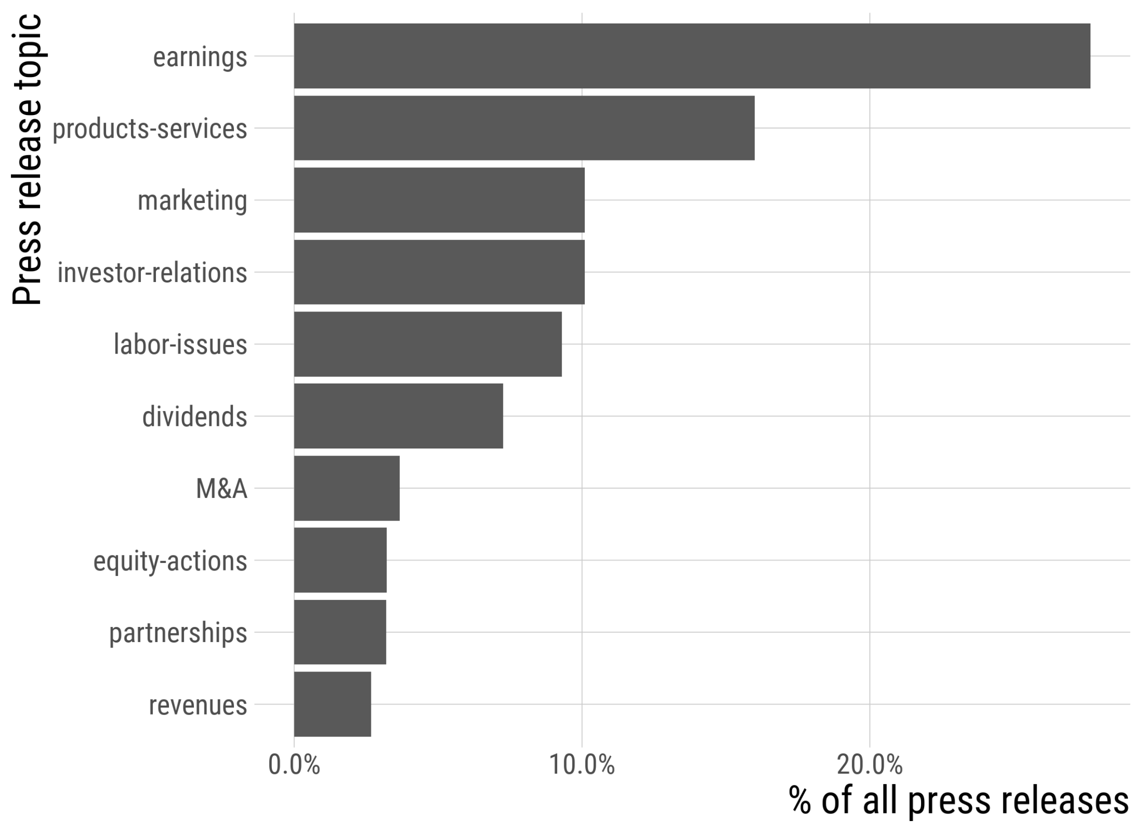
Filtering criteria	# of press releases	# of unique firms
Keep if RELEVANCE = 100 and in CRSP/Compustat	1,620,046	9,406
Keep Top 4 press release wires	1,502,021	9,386
Remove duplicated releases (ENS = 100)	1,068,148	9,373
Keep only one press release per firm-day	909,874	9,368
Keep only trading days	901,774	9,366
Keep if after April 1, 2006	738,196	8,756
Keep if issued in the first 30 seconds of an hour	188,981	7,911
Keep if issued in the first 10 seconds of an hour	131,683	7,503
Keep if issued in 7AM-9AM or 4PM	80,246	6,560

Table 16: Number of news articles by sources

This table shows the number of news articles from different sources, as provided in the RavenPack DJ Edition.

Source	Number of news articles
Dow Jones Newswire	9,561,578
Market Watch	149,021
Barrons	79,418
Wall Street Journal	74,106

Figure 4: Press release topic distribution



Online Appendix B. Details of generating EDGAR search variables

Please visit my website bruceyli.com for more updated details!