Market reaction to news and investor attention in real time

Deniz Erdemlioglu[a], Roland Gillet[b,c], and Thomas Renault*[a,c]

[a]IÉSEG School of Management and CNRS, France
[b]Université Libre de Bruxelles (ULB), Belgium
[c]Université Paris 1 Panthéon Sorbonne, France

This Version: August 10, 2017

Abstract

This paper develops a new framework to study investor attention in real time at high frequency. Using information retrieval approach, we construct a proxy for attention from the Twitter messages of financial experts, hedge funds and portfolio managers around the release of unscheduled news announcements. We then examine how markets react to new information in the absence and presence of attention. On implementing our methodology with high-frequency data for large-cap U.S. stocks, we find evidence that news events receiving attention on social media lead to large and persistent changes in trading activity, volatility and price jumps. When the attention is limited, however, the news effects on such trading patterns tend to be smaller and vanish quickly. With respect to reaction timing, we find that approximately one fourth of the news stories arrive first on Twitter before being reported by Bloomberg newswire. This result suggests that movements prior to news releases may not be explained only by private information, but could also be related to timestamp delays. We control for such potential biases by incorporating attention and correcting newswire timestamps. This adjustment considerably eliminates the pre-announcement effects in the data.

Keywords: Investor attention, News announcements, Stock returns, High-frequency data, Big data, Volatility, Jumps, Social media, Textual analysis, Information retrieval

JEL classification: D83, G12, G14.

*Electronic address: `thomas.renault@univ-paris1.fr`; Corresponding author: Thomas Renault. PRISM Sorbonne - Université Paris 1 Panthéon-Sorbonne, 17 rue de la Sorbonne, 75005 Paris, Tel.: +33(0)140463170.

1. Introduction

The flow of information plays a central role in financial markets. Macroeconomic announcements and firm-specific news often affect trading activity, market volatility and price dynamics. To understand the link between trading patterns and news flows, it is important to study how investors process the news and react accordingly by filtering value-relevant information from noise. In this paper, we propose a new framework for examining market reaction to news through investor attention to information in real time. Developing an attention network and using intraday one-minute data on large-cap U.S. stocks, we find evidence that news strongly influences trading activity only when investors pay attention to news announcements. In the absence of such attention, however, the effect of news on returns, volume, volatility and price jumps appears to remain weak and short-lived.

Measuring attention to financial news is challenging in continuous time. To cope with this difficulty, we use the flow of tweets from financial experts around the release of firm-specific unscheduled news. Although most of the 500 million messages sent every day on the micro-blogging platform Twitter are noisy, visual inspection typically suggests that certain Twitter posts—conveying value-relevant information—lead to large movements in financial markets. For example, on March 30, 2015, Tesla's stock price jumped by nearly 2% after a tweet from Tesla CEO Elon Musk announcing a new product line (Figure 1). On April 28, 2014, Twitter's stock price dropped by 5% after a tweet from Selerity, a FinTech company, leaking Twitter earnings results one hour earlier (Figure 2). On June 24, 2015, Netflix's stock price decreased sharply following a tweet from the investor Carl Icahn, who announced that he had sold his stake in the company (Figure 3). All these events, characterized by extremely active posting activity on social media, are associated with a large and long-lasting impact on trading volume, volatility and asset prices.

Even though these examples primarily reveal the reaction of individual stocks, tweets can also amplify marketwide movements. For example, on April 23, 2013, at 1:08 p.m., the Dow Jones Industrial Average index plunged by nearly 145 basis points in one minute after a fake tweet announced the bombing of the White House. Despite the fact that Dow Jones recovered by going back to its previous (pre-announcement) level—a few minutes just after the news proved to be false, this exceptional event illustrates the speed at which information can be shared and disseminated on social media as well as the link between investor attention and information in financial markets.

[**Insert Figures 1–3 about here**]

In light of these examples, the objective of this paper is to examine the intraday response of stocks to news announcements by uncovering investor attention to information. Our analysis has two main steps. First, we develop an *information retrieval method* (IRM) that utilizes Tweet flow as a proxy for investor attention. In *real time*, this method tracks all messages posted by financial experts and disentangles (value-relevant) signals from (irrelevant) noise. Second, we use the attention-based metric to characterize the effects of news events on stock market activity in various forms, such as abnormal returns, trading volume, price volatility and price jumps.[1] We compare market reactions in the presence (and absence) of attention and show that the reaction of stocks to news significantly depends on whether or not investors pay attention to news.

Contrary to conventional approaches on textual analysis and sentiment, our method extracts direct communication information from a micro-blogging platform rather than traditional newswires. This provides several advantages. First, Twitter may *break the news* because existing value-relevant information can be shared on social media before being re-

[1] Although the high-frequency market reaction to intraday stock-specific news using data from traditional newswires (Reuters or Dow Jones Newswire Services) has already been studied in the literature (Groß-Klußmann and Hautsch, 2011; Boudt and Petitjean, 2014), this paper is, to the best of our knowledge, the first to use data from Twitter for this purpose.

ported by mainstream newswires (Kwak et al., 2010). Second, Twitter may even *create the news* by itself. Companies can use the Twitter platform for announcing key information to investors in compliance with the Regulation Fair Disclosure, since the Security and Exchange Commission (SEC) reports on the use of social media by companies and markets participants.[2] Third, by targeting a wider number of "content providers" instead of focusing on a specific newswire, Twitter enables the construction of a transparent measure of news relevance and investor attention. As "news stories are not all created equal" (Barber and Odean, 2008), exploiting social media activity associated with a specific news release can help researchers and practitioners divide market reaction into (attention-grabbing) value-relevant news and noisy signals.

At low trading frequencies (such as daily or weekly), prior research documents mixed evidence on the power of search engines, social media and Internet communication in predicting asset returns (see Nardo et al., 2016 for a survey). Although those findings could be explained through market efficiency, other factors are also likely to justify the absence of predictability. First, information may be rapidly incorporated into asset prices, which requires an intraday analysis to better understand the causal relationship between online user-generated content and market dynamics. Second, value-relevant messages may, in fact, be lost in a massive flow of noisy content, leading to, on average, noisy signals. In this respect, accounting for users' credibility and reputation could help disentangle news from noise. Third, investor attention (based on online discussions) strongly correlates with traditional newswires stories. Therefore, combining online messages with traditional news flows could permit to identify "news that matters" from noisy stories and routine news coverages. In this paper, we provide a new methodology and empirical analysis to explore these three competing explanations.

[2] The Regulation Fair Disclosure applies to social media and other emerging means of communication used by public companies in the same way it applies to company websites. See "SEC Says Social Media OK for Company Announcements If Investors Are Alerted" (April, 2, 2013).

To disentangle news from noise, our methodology relies on information retrieval through network structure and identifies *experts in the crowd*. Specifically, we first start with a list of influential *Twitterers* (i.e., contraction of "Twitter" and "user") sharing opinions, views and news about the stock market on Twitter. Then, we implement an iterative algorithm based on directed relationships (friendships) between *Twitterers* to characterize a network of thousands of financial experts including hedge funds, asset/portfolio managers and investors. After identifying experts in the crowd, we consider five listed U.S. companies consisting of Apple, General Electrics, Walmart, Johnson & Johnson, and IBM. We extract all messages sent on Twitter by our list of financial experts.

When implementing our approach, we combine Twitter data with firm-specific "Hot Headlines" (HH) from the Bloomberg Terminal, focusing on unscheduled news published during the market's opening hours.[3] For each of the HH, we utilize a *similarity measure* to examine whether or not the news was available on Twitter before being reported by Bloomberg. This measure also allows us to automatically compute a statistical proxy for investor attention by assessing the similarity between HH content and messages published on Twitter around the release of the news.

Relying on our investor attention measurement, our empirical analysis reveals several distinct patterns and regularities in the data. First, for a large number of news events, we find that the content of the unscheduled news was already available on Twitter before being published on Bloomberg. The evidence—documented for other types of events, such as the death of Osama bin Laden or the plane crash in the Hudson River—confirms that Twitter can also *break* financial news. One implication of this result is that combining Twitter flow with traditional newswires' data can help locate the exact timestamp when public information is available to market participants. This, in turn, allows for better assessment of the impact

[3] For instance, our unscheduled news announcements are related to product announcements, firm-events announcements, activist investors communication and legal issues, among others.

of unscheduled news on financial markets' movements. Along these lines, we use one-minute intraday data and newswire-corrected timestamps to investigate market reactions to news that receive high-attention versus low-attention from financial experts. We find evidence that attention-grabbing news are followed by large and persistent changes in trading volume, volatility and price jumps. When investors pay no attention to news, however, the impact on such measures of market reaction appears to be low, and the effect of news vanishes very quickly at intraday levels. For both high-attention and low-attention cases, we find no evidence in favor of price predictability after news releases, reflecting information-induced efficiency. More broadly, our results suggest that Twitter can help trace news events that matter in a continuous flow of information, without relying on "black box" pre-processed data or subjectively-picking-seemingly-relevant news. By combining Twitter flow with high-frequency news events from traditional newswires, researchers can disentangle the effects of pre-announcement private information from ambiguous timestamp identification and avoid underestimating the impact of unscheduled news caused by noisy stories or rumors.

The remainder of this paper is organized as follows. In Section 2, we discuss the related literature. Section 3 describes our identification approach to retrieve information from Twitter. This section further presents the intraday data on stocks and news announcements. Section 4 introduces our methodology to construct a proxy for investor attention in real time. Section 5 outlines our empirical analysis and presents the results. Section 6 concludes.

2. Related literature

Modern finance theory suggests that "news"—defined as textual information from traditional media—should not influence stock markets, unless the news events contain value-relevant information about the discounted value of future cash flows. Under the assumption that information revealed by traditional media is stale due to the publication lag, media should thus play no role in the price discovery process.

A few decades ago, before the advent of the Internet and the availability of high-frequency data, one could argue that information disseminated by traditional daily morning and afternoon newspapers was indeed stale when made public. This conclusion is, however, questionable in today's financial markets, where an almost continuous flow of news can be exploited by fast-moving traders (Foucault et al., 2016) and by machines reading the news (Groß-Klußmann and Hautsch, 2011). Furthermore, news is no longer the monopoly of traditional media, that is, every user can now be a media outlet by publishing content on blogs, message boards, or social media (Shirky, 2008). U.S. companies can now directly use social media to disseminate key information to investors, in compliance with the Regulation Fair Disclosure. Traditional media are still among the main news providers, but their business model has evolved from a daily newspaper to a continuous flow of online information, where breaking news often plays a significant role in developing online traffic. These recent technological, organizational, and regulatory changes reinforce the need for empirical research on the informational efficiency of financial markets.

The literature on the high-frequency market impact of scheduled macroeconomic releases (Andersen et al., 2007; Bollerslev et al., 2016) and Federal Open Market Committee announcements (Faust et al., 2007; Wongswan, 2009) is rather vast. High-frequency scheduled news (surprise) has a significant impact on asset prices and, typically, explains a substantial fraction of the increase in price volatility and trading volume following the news (Balduzzi et al., 2001). Recently, Bernile et al. (2016) also document substantial informed trading before the official release time of scheduled announcements, consistent with information leakage from the news media or from insiders. Research on the impact of unscheduled news arrivals is, however, relatively scarce. While macroeconomic events often affect the movements of individual stocks, sudden and unexpected firm-specific information can also impact asset pricing and market liquidity (Boudt and Petitjean, 2014). For example, analyzing the intraday market dynamics of firm-specific announcements at the Paris Bourse, Ranaldo (2008) finds a significant increase in liquidity and higher adverse selection costs around news arrivals. Riordan et al. (2013) utilize the Thomson Reuters newswire messages and provide similar evidence for the Toronto Stock Exchange.

When analyzing the impact of unscheduled news on stock returns, liquidity, or volatility, researchers and financial economists reconstruct news databases by searching news events about a given company on Factiva, Thomson Reuters, Dow Jones News Services, Lexis-Nexis, or the Wall Street Journal (Das et al., 2014). However, unlike the precisely timed macroeconomic announcements, the identification of the exact timestamp of unscheduled news events is non-trivial. Moreover, as the number of news articles published steadily increases, tracing value-relevant news in an overwhelming number of articles published every day is a major challenge. Pre-processed news data such as the Reuters NewsScope Sentiment Engine (Riordan et al., 2013) or the RavenPack News Analytics (Smales, 2014) can help solve issues related to the identification of value-relevant news. Proprietary "black box" algorithms developed by a few private companies automatically assess novelty, relevancy, and sentiment scores for all articles from a list of news providers. For example, Groß-Klußmann and Hautsch (2011) use intraday pre-processed news from Reuters to analyze the high-frequency trading impact of unscheduled news releases by separating announcements into relevant and non-relevant news. They find evidence of significant positive (negative) price movements prior to positive (negative) relevant news releases, but notice only weak return responses after the announcements. Volatility, liquidity, and trading volume typically start increasing about 60 minutes prior to the news release, peak at the exact time of the release, and decrease later in the day. Although pre-event movement could be explained by private information, Groß-Klußmann and Hautsch (2011) argue that the availability of other sources of information and an induced clustering of news items are mainly responsible for pre-announcement effects. This finding and the absence of transparency of the algorithms used to derive articles' scores (novelty, relevancy, and sentiment), encourage further research in the area, not only to improve the timestamp of news detection but also to disentangle value-relevant news from noisy content with a more robust methodology.

From a different perspective, recent studies on computational science focus on methodologies to automatically detect "breaking stories" in a continuous flow of messages from social media (Mathioudakis and Koudas, 2010; Petrovic et al., 2013; Ifrim et al., 2014). The approaches in these works are built on the tools in the area of natural language processing (named *entity recognition*) and topic classification. The basic intuition behind event detection is as follows. When value-relevant breaking news arrives, users on social media will change their posting activity and start talking very actively about the event. By analyzing the tweet flow in real time, looking for surge in absolute posting activity, a burst in the frequency of certain keywords, or the appearance of new topical clusters, practitioners can identify value-relevant news even in the absence of a specific news provider or another measure of relevancy. Historical databases of messages are available (albeit expensive), so that discovering the precise timestamp at which information was public is possible (at the ex-post level) by identifying the first mention of a news article on social media.

Theoretical models and empirical studies often suggest that investor attention to marketwide news plays a central role in changing asset prices and volatility (Li and Yu, 2012; Andrei and Hasler, 2015; Yuan, 2015). At the company level, Barber and Odean (2008) provide evidence that retail investors are net-buyers of attention-grabbing stocks, and Solomon et al. (2014) document that media coverage attracts investor attention and affects investors' capital allocations to mutual funds. Recently, Boulland et al. (2017) demonstrate that investor attention, measured by the use of an English-language electronic wire service, significantly affects market reaction to earning surprises.

Indirect measures, such as 52-week high (Driessen et al., 2013), the day of the week (Friday effect) (DellaVigna and Pollet, 2009), or the level of media coverage (Barber and Odean, 2008), have been used in the literature to proxy for investor attention. Despite this substantial progress, these proxies have at least two important limitations. First, market data (such as price and volume) contain idiosyncratic components that are unrelated to attention. Second, simply counting the number of news articles does not take into account the salience of news coverage and could be easily affected by routine company press release.

To overcome these issues related to proxy selection, Dimpfl and Jank (2016) and Da et al. (2011) examine the number of queries about a given company on Google to compute a more direct measure of investor attention. Although the Google search engine is of interest from a practical viewpoint, detecting investor attention through the evolution of online search behavior also has certain drawbacks. For instance, the Google data are available only on a daily basis with no information on the absolute level of search (scaled on a range of 0 to 100 based on a topic's proportion to all searches on all topics). Therefore, the precise assessment of the evolution in the Search Volume Index remains elusive. In this respect, combining news published on traditional media with attention toward the news on Twitter could resolve issues related to partial identification. It also allows the construction of a transparent high-frequency proxy for investor attention. We attempt to bring this resolution to research on news reaction analysis of financial markets.

Following the existing theoretical and empirical literature, we hypothesize that news arrivals cause price jumps and they are followed by a persistent increase in volatility and trading volume. We add to the literature by examining two further hypotheses. The first hypothesis focuses on movements prior to the news release and on the importance of using newswire-corrected timestamps for intraday studies. This hypothesis relies on the results in

Bradley et al. (2014), who document that timestamp delays could lead to incorrect inference. Consistent with Bradley et al. (2014), we argue that the pre-announcement effects are generally *overestimated* when researchers consider standard newswire timestamps. We find empirical evidence supporting this hypothesis and further show that correcting the timestamp delays significantly reduces movements prior to news releases. The second hypothesis is closely related to the theoretical framework of Andrei and Hasler (2015): high attention should induce high return volatility if attentive investors immediately incorporate new information into prices. Conversely, when investors pay little attention to news, information should only be gradually incorporated into prices resulting in low volatility. By empirically testing these theoretical features, we show that market reaction to information (in forms of volatility, price jumps, trading volume and return) is indeed much more pronounced when investor pay attention to news compared to when they do not.

3. Data

We proceed with the description of our databases. In Section 3.1, we first discuss how to use Twitter data and process communication information to construct our attention proxy. Section 3.2 then details the data on stocks and news announcements obtained from the Bloomberg Data Analytics.

3.1. Constructing a proxy for attention using Twitter data

On the micro-blogging platform Twitter, users can post short messages, called "tweets" (140-characters maximum), share messages sent by other users with their community of followers ("retweet"), or simply read "tweets" of users they choose to follow. Compared to other social networks, such as Facebook or LinkedIn, relationships between users on the platform are public and can be accessed through the Twitter Application Programming Interface (TAPI). We use TAPI to transform the Twitter information network into an adjacency matrix and identify specific clusters inside that matrix. The Twitter-directed relationships allow us to create a list of *financial influencers*. The list includes the tweets of important investors (through pre-filtering), financial journalists, investors/traders and experts working in financial markets or institutions.

Specifically, we take the following steps to identify market participants. First, we start with a list of 10 influential *Twitterers* (i.e, combination of "Twitter" and "user") sharing news and ideas about the stock market on Twitter. We impose the following four criteria to include a user in our initial list: (1) the user has a verified Twitter account, (2) the user has a dedicated Wikipedia page, (3) the user has at least 100,000 followers, and (4) the user has a professional job related to financial markets.[4] Table 1 presents our initial list of 10 users. We denote this set as N_0.

[**Insert Table 1 about here**]

We then consider that common friends of influential experts in finance should also be influential and tweet regularly about financial markets. To account for this, We use the TAPI to extract the friends list of each user in N_0.[5] We insert the unique identifier of all users followed by at least one user from N_0 into a MongoDB database, ending up with a list of 15,390 users. Finally, we create a new iterated list (i.e., N_1) by augmenting N_0 with the 50 most commonly followed *Twitterers* from the list of 15,390 users.[6] Appendix A details our setup as well as the implementation of the network algorithm used to characterize

[4] The final list of 3,010 users identified using our methodology is robust to the initial list with the choice of 10 users. We find a similarity of 85-95% when considering other lists of 10 financial experts.

[5] https://dev.twitter.com/rest/reference/get/friends/list.

[6] These *Twitterers* added during the first iteration include Elon Musk (Tesla CEO), ZeroHedge (financial media), Citron Research (financial analyst), Blackhorse Analytics (equity research), Joe Weisenthal (Bloomberg editor), John Carney (Wall Street Journal market editor), Fred Wilson (venture capitalist), the New York Times Business section (media), Dan Primarck (journalist at Fortune), Chris Sacca (venture investor), Henry Blodget (former equity research analyst, CEO Business Insider), Horace Dediu (industry analyst), and T. Boones Pickens (hedge fund manager).

attention. Figure 4 displays our constructed N_1 network based on this setup. We notice that the information network is highly connected, with a total of 973 directed links between the 60 users from N_1.

[**Insert Figure 4 about here**]

As in step 1, we iterate this algorithm by extracting the friends list of each user in N_1 and adding the 50 most commonly followed users to N_1.[7] Having generated 60 iterations, we obtain our final network N_{60} that consists of 3,010 users.[8] Our final list includes official media Twitter accounts (e.g., CNBC, Financial Times, Reuters, Bloomberg), journalists' personal accounts (e.g., Jim Cramer, Carl Quintanilla, Maria Bartorimo, David Faber), market participants and investors (e.g., Warren Buffet, Carl Icahn, Mark Cuban, Marc Andreesen), CEO and insiders (e.g., Tim Cook, Elon Musk, Satya Nadella, Marissa Mayer), institutions (CBOE, Federal Reserve, NASDAQ) and several celebrities, such as Taylor Swift, Ellen DeGeneres, Barack Obama, and Oprah Winfrey.

Given our constructed network, we are particularly interested in examining trading activity patterns around Twitter messages as reliable proxies for investor attention. To achieve this goal, we focus on five U.S. companies that are Apple (AAPL), Walmart (WMT), International Business Machine (IBM), Johnson & Johnson (JNJ) and General Electric (GE).[9]

---

[7] In this case, we consider a new list of 60 influencers instead of the initial list of 10. *Twitterers* identified during the second iteration include official financial media (e.g., WSJ, Bloomberg, CNBC), financial journalists (e.g., Jenn Ablan, Dennis K. Berman, Charles Gasparino), hedge fund and portfolio managers (e.g., Doug Kass, Mark Dow, Anthony Scaramucci), and traders/venture capitalists (e.g., Paul Kedrosky, Jon Najarian, Bill Gurley).

[8] For brevity, we do not report the full list yet it is available upon request.

[9] One challenge is that historical access to archives for keyword-related queries is rather limited. The TAPI allows registered applications to extract only the last 3,200 tweets sent by each user. To have a greater depth and retrieve all tweets since January 2013, we hence developed an application using Python. First, we relied on the new "advanced search tool" available on Twitter (since April 2014) and extracted the unique identifier of all tweets sent by users in N_{60} including the keywords related to companies in our sample. For the Apple company, for instance, we extracted the unique identifier of all tweets containing keywords "Apple," "$AAPL," "AAPL," "Tim Cook," "iPhone," "iPad," "iPod," "iTunes" and "Macbook" sent by experts from N_{60}. Next, we used the Twitter "GET statuses/show/:id" function to retrieve detailed information about each message and inserted all tweets into a MongoDB

Those companies are amongst the 10 companies with the highest market capitalization in the U.S. as of January 1, 2013. This classification also helps us avoid a sectoral bias observed typically on high-tech companies, such as Google, Microsoft, or AT&T, or in the oil industry (Chevron, Exxon).[10]

[Insert Table 2 about here]

Based on this data construction scheme, our Twitter database consists of 498,366 messages containing a keyword related to one of the five companies in the sample. For the messages of market participants from N_{60}, the adjusted data span the period from January 1, 2013 to December 31, 2015. The most widely covered company is Apple (414,844 tweets), followed by Walmart (32,872), IBM (25,444), General Electric (17,915), and Johnson & Johnson (7,191).[11] Table 2 shows a sample of messages published on January 2, 2013 (the first trading day of our period).

3.2. News announcements and stock data

We extract company-specific news announcements from the Bloomberg Professional Service (BPS).[12] When constructing our news database, we focus on Bloomberg "Hot Headlines" (HH) because HH are typically released very quickly by Bloomberg Analytics. In order to

NoSQL database. Given the TAPI limits, the data collection process is limited to one message every five seconds. To collect all data, we ran our Python script during one month.

10 In this study, we are particularly interested in investigating how markets respond to new information arrivals when investors pay attention to news and when they do not. Of course, future research can consider other asset classes to assess the link between attention and marketwise scheduled news announcements.

11 The large difference between social activity about Apple and other companies can be explained in several ways. First, Apple was the company with the highest market capitalization in the world at this time. Second, Apple is the most covered company by media and a well-known company to the general public. Third, high-tech companies are, on average, more covered on social media than industrial companies. Lastly, every new product released by Apple is followed by a wave of euphoria in the real world as fans, for instance, queuing in front of Apple stores. This pattern is also visible on social media and Twitter.

12 BPS is a platform through which financial professionals can monitor and analyze real time data, news and analytics.

alert practitioners about the release of a (potentially) value-relevant announcement (e.g., political event, macroeconomic event or company-specific news), HH are released as short statements (e.g., 10 words on average). These hot headlines can be further recognized in the flow of news available on Bloomberg as they are capitalized and highlighted in red. The exact timestamp of the news release (up to the second) is available on Bloomberg.

We manually extract all HH relative to Apple, General Electric, IBM, Johnson & Johnson and Walmart. Similar to the patterns of tweet flow, the most covered company is Apple, with a total of 1,528 HH between January 1, 2013, and December 31, 2015, followed by General Electric (977), Walmart (457), Johnson & Johnson (383) and IBM (323). Then, we manually filter all HH to remove duplicate events, irrelevant news, and announcements related only to variation in stock prices.[13] We further eliminate headlines about scheduled news or events. For example, a few times a year and during market trading hours, Apple organizes special events (keynote events), where the company makes announcements about new products or developments. Price volatility is especially high during those scheduled events, and, given the high number of news announcements provided in a small amount of time, isolating the effect of a specific announcement from others is difficult.[14] As in Bollerslev et al. (2016), we consider only day-trading sessions and hence neglect the entire news events published overnight and on non-trading days.

For all companies in our sample, we use one-minute data (transaction prices and volume) from January 1, 2013, to December 31, 2015.[15] As is standard in the literature, we also omit trading days that have too many missing values or low trading activity. Because trading

[13] For Apple, for instance, we removed the headline "Blackberry previews secure work space tech for Android, iOS" or the headline "Apple unchanged erasing gain of 1.4% at the open."

[14] When conducting robustness checks (unreported for brevity), we also include those news articles in our event-study. The results remain qualitatively the same.

[15] Of course, an analysis based on higher frequency (tick-by-tick or one-second) could shed more light on how market participants process information in a framework with fast-moving (or slow-moving) traders (see (Foucault et al., 2016)). We restrain our analysis to one-minute data due to data availability and, more importantly, to avoid potential frictions related to market microstructure noise.

volume and volatility exhibit strong intraday patterns (due to opening and closing hours), we use the procedure proposed by Erdemlioglu et al. (2015) and remove periodic patterns before conducting our empirical analysis.[16]

4. Methodology

In this Section, we introduce our main methodology for examining the high-frequency response of stocks to investor attention. For this objective, the next subsection presents first the underlying continuous-time model and then details how we characterize market fluctuations in various forms. In Section 4.2, we outline the identification of our attention measure based on information retrieval approach.

4.1. Reaction forms in continuous-time

Because our focus is to examine the market impact of attention at high frequency, we describe the behavior of stocks at short-time scales in continuous-time. Throughout, we thus assume that the log-price of a stock $p(t)$ follows a standard diffusion process with jumps. While the former component helps us characterize the smooth/diffusive volatility reaction to news, jumps reflect abnormal uncertainty or shocks. That is,

$$dp(t) = \underbrace{\mu(t)dt}_{\text{drift}} + \underbrace{\sigma(t)dW(t)}_{\text{volatility shocks}} + \underbrace{\kappa(t)dq(t) + h(t)dL(t)}_{\text{jump shocks}}, \tag{1}$$

where $dp(t)$ denotes the logarithmic price increment for $t \geq 0$, $\mu(t)$ is a continuous, locally bounded, variation process, $\sigma(t)$ is a strictly positive and càdlàg (right-continuous with left limits) stochastic volatility process, and $W(t)$ is a standard Brownian motion. In Equa-

16 For brevity, we do not report the estimated periodicity factors and diurnal patterns. These results are available upon request.

tion (1), $q(t)$ further denotes a counting process (e.g., compound Poisson process), $L(t)$ is a pure Lévy jump process (e.g., Cauchy process), $\kappa(t)$ and $h(t)$ denote the jump shock sizes of the counting and Lévy processes, respectively. Intuitively, the jump shocks of (1) potentially represent both finite- and infinite-activity. While finite-activity jumps capture rare and large abnormal reactions, the infinite-activity component tracks relatively small yet frequent jumps in asset prices.

Given this underlying model, we next estimate diffusive (spot) volatility and detect the arrivals of extreme price changes (i.e., intraday jumps). We use the truncation approach of Bollerslev et al. (2013) to identify the realized intraday jump shock increments of the assets. Let $JV_{t,i}$ denote the jump variation at time t of a trading day i. Then, we identify the intraday jumps as

$$JV_{t,i} = \{i \in [0,T] : |r_{t,i}| \geq u\}, \tag{2}$$

where $r_{t,i}$ is the intraday price increment (return) at time t of a trading day i, $u = \alpha\Delta^{\varpi}$ is the truncation threshold and α (> 0) is expressed in units of standard deviations of the continuous part of the process for a constant $\varpi \in (0, 1/2)$. This truncation approach in (2) can be used to detect large price changes (i.e., jumps), and hence its reverse version retains the diffusive (or continuous) volatility shock component such that

$$CV_{t,i} = \{i \in [0,T] : |r_{t,i}| < u\}, \tag{3}$$

where $CV_{t,i}$ is the estimated diffusive spot volatility of (1). As in Bollerslev et al. (2013), we set the truncation thresholds $\alpha = 3$ and $\varpi = 0.47$ for both jump and volatility estimations (i.e., in Equations (2) and (3), respectively).

In addition to volatility and jump reaction forms, we further consider volume as a proxy for trading activity. In this regard, we follow Groß-Klußmann and Hautsch (2011) and estimate abnormal trading volume by standardizing the process using the yearly average of the corresponding underlying one-minute intervals. That is,

$$V^*_{s,t} = \frac{V_{s,t}}{\frac{1}{250}\sum_{d=-250}^{-1} V_{d,s,t}}, \tag{4}$$

where $V^*_{s,t}$ denotes the abnormal trading volume on minute t for stock s, $V_{s,t}$ is the one-minute trading volume and $V_{d,s,t}$ is the trading volume of the corresponding underlying minute t on day d. Finally, in order to characterize the return reaction at high frequency, we assume a normal-return asset pricing model as in Groß-Klußmann and Hautsch (2011). Specifically, we define the abnormal return as the difference between the actual return and the estimated normal return given by

$$R_{s,t} = \alpha_s + \beta_1 Rm_t + \beta_2 R_{s,t-1} + \epsilon_{s,t}, \tag{5}$$

$$AR_{s,t} = R_{s,t} - \widehat{R}_{s,t}, \tag{6}$$

where $AR_{s,t}$ denotes the one-minute abnormal return of stock s, $R_{i,t}$ is the one-minute return of the individual stock and Rm_t is the one-minute return of the S&P500 (SPY Exchange Traded Fund). In line with Fama (1998), and because we focus our analysis on a short [-30:+30] minutes event window, the considered model of normal returns barely affects the inference about abnormal returns (i.e., the expected returns on a short event window are close to zero).[17] To pin down the effects of unscheduled news on intraday returns accurately, we implement a trading strategy by selling stocks on negative news and buying stocks on positive news. We define headline's sentiment manually since sentiment measurement based on standard dictionary-based approach (see e.g., Loughran and McDonald, 2011; Jegadeesh

[17] In various robustness checks, we confirm that our results are insensitive to the choice of asset pricing model (i.e., constant-mean or market-return). These results are unreported for brevity, but they are available upon request.

and Wu, 2013) is likely to be biased due to the insufficient number of words in hot headlines.

4.2. Quantifying attention with information retrieval scores

To characterize investor attention properly, we analyze all messages posted on Twitter in a [-15:0] minutes window before (and at the exact same minute) the release of each Bloomberg headline. We utilize a Term Frequency-Inverse Document Frequency (TF-IDF) cosine similarity measure to avoid considering messages posted around the release of the news but not related to the news.[18] Given the interactions between b_1 (a Bloomberg headline) and t_1 (a Twitter message), the cosine similarity—collapsed into two TF-IDF vectors B and T)—is given by

$$Cos^{sim}(b_1, t_1) = \frac{\sum_{i=1}^{n} B_i T_i}{\sqrt{\sum_{i=1}^{n} B_i^2}\sqrt{\sum_{i=1}^{n} T_i^2}}. \tag{7}$$

Because the TF-IDF value is always positive, the cosine similarity ranges between 0 and 1.[19] Higher cosine similarity implies a closer similarity between a given message published on Twitter and the Bloomberg headline. Table 3 reports some examples of cosine similarities between a Bloomberg headline and all messages sent on Twitter on a [-15:0] minutes window around the timestamp of the release of the headline on Bloomberg.[20]

[Insert Table 3 about here]

[18] Cosine similarity is a standard approach taken from natural language processing and information science literature to assess the similarity between two documents (see Loughran and McDonald, 2016).

[19] To improve the accuracy of the TF-IDF cosine similarity measure, we use a Porter stemmer to remove the commoner morphological and inflexional endings from the words in all messages and headlines. We also remove all stop-words, links, company names, and mentions from messages. For example, the headline "Apple PT cut to 530 from 660 at Nomura" became "pt cut 530 660 nomura." The tweet "It's one of the great conundrums of investing. What IS this stock? @JamesStewartNYT on whether $AAPL is growth or value. @CNBC" became "one great conundrum invest what stock whether growth valu."

[20] That is, "Einhorn drops suit against Apple over shareholder vote" (released on Bloomberg at 11:25:12 a.m. on March 1, 2013).

For each Bloomberg headline, we compute the attention score by adding the cosine similarity between the headline and all messages sent on Twitter in a [-15:+0] minutes window around the exact timestamp of the Bloomberg news release. Assuming that there are n messages published on Twitter within this window, we finally define $News_i^*$ as the level of attention to HH

$$News_i^* = \sum_{j=1}^{n} Cos^{sim}(b_i, t_{i,j}). \tag{8}$$

We define the *attention-grabbing* news from with a $News_i^*$ score greater than (or equal to) 0. Although we consider other threshold values (0.5 and 1) to separate attention-grabbing news from low-attention news, we find that results are robust to the different threshold values (see Appendix B for an example on trading volume). In our empirical analysis, we set the threshold value as 0.

5. Empirical analysis

5.1. Analyzing market reaction to news: an event-study

We conduct an event-study to investigate the high-frequency impact of unscheduled news announcements on trading patterns. As we discussed in Section 4.1, We consider four reaction forms: abnormal trading volume, abnormal returns, diffusive volatility and sudden price jumps. We account for investor attention using our information retrieval approach and—relying on this scheme—we examine the characteristics of market reaction for high- versus low-attention to news.

We set a [-30:+30] minutes event window. We follow Bollerslev et al. (2016) and eliminate news articles published during the first and the last 30 minutes of each trading day. Therefore, we derive all minutes in the event window from the same trading days, which allows us to cope with the identification issue due to overnight news and the sharp opening variation at 9:30 a.m. on each trading day. We also impose a minimum length of 30 minutes between two events for a given company to avoid problems related to overlapping or timing. Taken together, we analyze the duration, news timing, and persistence of all news-implied reactions. These considerations provide us 547 events in total. To assess the significance of our variable of interest (return/volatility/jumps/volume), we compare the estimates on the event window with those obtained from the last trading days without any unscheduled news events during market opening hours. For example, for the Bloomberg HH "Apple gets 30M iPad deal from LA unified school district" published at 1:10:24 p.m. on June 19, 2013, we consider an event window from 12:40 p.m. to 1:40 p.m. (61 minutes) on that day, and we compare the event window results by considering the last previous trading days without any unscheduled news as our estimation window (i.e., 330 minutes of trading on June 18, 2013, from 10 a.m. to 3:30 p.m.). For statistical inference, we carry out the non-parametric rank test of Corrado (1989). In the next section, we present and discuss our main empirical results. This section will then proceed by presenting a correction method to identify the timestamps of news releases. In Section 5.3, we decompose market news responses into high- versus low-attention components and compare the patterns in the data.

5.2. Reaction and timing: does Twitter break the news?

For all 547 pre-identified Bloomberg news events, we start by comparing the exact timestamp of the HH (up to the second) with the first mention of the same news on Twitter.[21]

[21] It is worth mentioning that Twitter provides incentives for users to attempt to "break the news". The main incentive is related to the fact that publishing information on Twitter before the release on traditional newswires could increase the credibility and reputation of financial experts. Even official media Twitter accounts (e.g., CNBC, Reuters) and journalists associated with traditional media tend to publish "breaking news" on Twitter before reporting the news on their websites or platforms. By increasing their reputation and their number of followers, the media (and journalists) can increase readership and maximize future revenues derived from traffic acquired through Twitter. Investors and market participants also benefit from sharing breaking news on social media in order to increase their

Analyzing all messages with a positive cosine similarity sent on a [-15:0] minutes window around the release of each HH, we find that Twitter effectively *breaks* 127 news events out of 547 (23.22%). In previous example shown in Table 3, we identify a mention of the news on Twitter two minutes before Bloomberg release. At 11:23:45 a.m., Kaja Whitehouse, a New York Times, reporter covering crime and corruption, published the following message on Twitter: "David Einhorn withdraws lawsuit against Apple. Manhattan federal court approves." Bloomberg headline was then posted at 11:25:12 a.m. "Einhorn drops suit against Apple over shareholder vote." Table 4 reports examples of such cases for which Twitter does break the news.[22]

[**Insert Table 4 about here**]

The delay between newswire-reported timestamps and the very first moment at which news arrives on social media (and hence becomes public) tend to support the conclusion of Groß-Klußmann and Hautsch (2011) on high-frequency news-implied market reactions. Price movements prior to news releases may not be solely attributed to private pre-release information, but could be explained by biased (or imperfect) timestamps. As also shown by Bradley et al. (2014) for analysts' recommendations, identifying the exact minute at which an unscheduled news event was made public is crucial for a high-frequency analysis because a failure to do so can lead to (buy/sell) timing implications for traders.

Before assessing whether the degree of attention influences how markets respond to news, we first assess if combining traditional newswire data with Twitter helps disentangle the effects of private information from misspecification of the exact timestamp of news releases. More precisely, we compare market reaction to news considering (1) all HH using Bloomberg reported timestamp as the event minute and (2) all HH considering the first mention of the

own reputation or influence other investors.

[22] The duration of release time ranges between few seconds and a few minutes.

news on Twitter (when social media "breaks the news") and Bloomberg reported timestamp otherwise. Table 5 reports the results for each 5-minutes interval around the release of HH.

[**Insert Table 5 about here**]

As expected, we find strong and significant increase in volatility, price jump, trading volume and abnormal return around the release of unscheduled news announcements. Perhaps more interestingly, the results show that the use of Bloomberg reported timestamp tends to exploit the magnitude of the movements prior to the news release, especially for price jumps and trading volume. These two reactions forms are overestimated by around 10% due to poor timestamp identification five minutes prior to the news release. Figure 5 illustrates such patterns for the pre-event period when abnormal trading volume estimated using newswire timestamps exceeds the quantity with the correction. The reaction differential between two magnitudes is statistically significant and economically large for an intraday news impact evaluation. It is, however, important to note that return reactions are rather swift and relatively insensitive to timestamp considerations.

[**Insert Figure 5 about here**]

Having corrected the timestamp delays using Twitter, we find that the pre-announcement effects disappear for volatility and trading volume, and such effects also shrink for price jumps. Consistent with this evidence, combining the Twitter messages of financial experts with traditional media news stories provides a precise identification of the exact minute at which news is public. This in turn allows to better understand the role of private information (if any) in the price formation process and trace the magnitude and duration of market reaction to unscheduled news announcements. Throughout, we use newswire-corrected timestamps for our analysis.

5.3. *Does attention change the reaction?*

We now turn to assess whether the degree of attention (high versus low) influences how markets respond to news announcements. Tables 6 and 7 report the results for each 1-minute and 5-minute surrounding the announcement releases, respectively. We further plot in Figures 6-10 the patterns for volatility, price jumps, abnormal trading volume, abnormal returns and cumulative abnormal returns.

[Insert Tables 6 and 7 about here]

Consistent with the findings of Andrei and Hasler (2015), the results indicate that spot diffusive volatility is significantly high after the release of attention-grabbing news whereas its reaction is relatively low when investors do not pay attention to events. Specifically, the reaction is statistically significant up to 15 minutes following the news release, peaking around 10 minutes after the release of news before it slowly decreases. Price volatility is, on average, 50% higher following high-attention news than following low-attention news.

Turning to jump-type tail reactions, we also find a large impact differential between high- and low-attention news announcements. Price jumps are significantly more frequent from two minutes prior the release up to five minutes following the release of high-attention news. In sharp contrast, there is, however, no significant increase in the magnitude of price jumps for low-attention news (except at the exact minute of the release). Overall, when considering a [+1:+5] minutes interval after the release, the probability of observing a jump is at least four times higher for high-attention news than that for low-attention news. More broadly, these results are in line with those reported in Dewachter et al. (2014) on the impact of Euro area official communication on the foreign exchange market. Dewachter et al. (2014) find that central bank unscheduled communication triggers large jumps and significant increase in volatility for approximately an hour after the news release. We provide evidence for similar effects in the U.S. equity market: unscheduled company-specific news announcements trigger market uncertainty in the form of price jumps. We show that such patterns are particularly noticeable when investors pay attention to news.

[**Insert Figures 6-10 about here**]

Trading volume exhibits patterns similar to volatility and jumps. For instance, we notice that volume is, on average, two times higher following high-attention rather than low-attention news. Furthermore, trading volume remains statistically abnormal for up to 30 minutes after the release of high-attention news whereas the effects die out very quickly (within five minutes) for news events that do not receive attention from market participants. These results are broadly consistent with Groß-Klußmann and Hautsch (2011), who utilize the same measure of abnormal volume and find that the money value trading is around 2.6 times higher following the release of high-relevance news, but only 1.5 times higher for low-relevance news. Tweet flow and investor attention thus may help disentangle relevant news events from those having only noisy signals. The similar patterns could be due to a correlation between our measure of investor attention and the "relevancy" indicator of the the Reuters Newscope Sentiment Engine used by Groß-Klußmann and Hautsch (2011).

With respect to regularities associated with abnormal returns (Figure 9), our event-study analysis reveals an increase of 0.047% (0.013%) at the exact minute of the release of attention-grabbing (versus low-attention) unscheduled news releases. The size difference is statistically significant and fairly large as a small (high-frequency scale) return magnitude. Nevertheless, we find no clear evidence for return predictability and there is neither momentum nor price reversal effect after the announcements. The U.S. stock market thus appears to be efficient enough, in the sense of Jensen (1978), such that a trader cannot generate profits based on widely disseminated news without acting almost immediately. This finding is in line with the intraday event-study results of Busse and Green (2002) related to the analysts' views broadcasted on CNBC TV.

Unlike these features related to interval-based abnormal returns, the cumulative abnormal return reactions (displayed in Figure 10) appear to be more systematic, particularly when the news events attract the attention (dashed versus solid lines). Two interesting patterns emerge in this analysis. First, if the information content is *unobtrusive* (failing to catch attention), then the event impact line remains mostly the same and does not change its path before and after the news release (red-solid line). Second, when investors pay attention to news (so that the content is considered as relevant), cumulative abnormal returns are negative (positive) during certain periods prior (following) the announcements (blue-solid line). The size of the news effect in the presence of attention is substantially large compared two alternative cases that either show low attention or neglect entirely the attention to news (dashed line).

These results overall suggest that the degree of attention significantly influences how markets respond to news announcements. According to Hirshleifer and Teoh (2003), the immediate reaction to news within a short event windows implies some investors turn their attention very rapidly to relevant announcements. Our findings may also help explain empirically the theoretical model of Andrei and Hasler (2015) on the role played by investors attention to news in measuring volatility. In this regard, we provide evidence such that unscheduled *attention-grabbing* news events are also significantly followed by large and persistent market impact on trading volume (for up to 30 minutes) and prices jumps (for up to 5 minutes). Moreover, despite the fact that a trading strategy based on unscheduled news releases may not be that profitable through one-minute data (as in Groß-Klußmann and Hautsch, 2011), the return trends might still be exploited by algorithmic (fast-moving) traders, who are able to trade at the exact second of the release of the news.[23]

[23] This can be achieved even after correcting for timestamp delays and accounting for the level of attention. We hence encourage further research in this area. It is also of interest to uncover the impact of attention to news for smaller companies for which the level of attention might have more impact on the price

6. Conclusion

This paper develops a new measure of investor attention by combining the news flows from conventional newswires with the tweet flow of market participants and financial experts. We find evidence that investor attention can help identify the exact time at which a news event becomes publicly available. Firm-specific announcements often break on Twitter before being reported on newswires, which in turn allows researchers to better understand the role of private information in affecting trading processes prior to the official news releases. Market reaction to news in pre-announcement spells could be related to bias in the news release times rather than information leakages.

The results also suggest that the degree of attention on Twitter about particular news changes the trading activity of stocks. While unscheduled attention-grabbing news are significantly followed by large and persistent market impact (on volume, volatility and price jumps), low-attention news flow fails to move trading. The price impact of new information is large only if investors give close attention to news. Studies on empirical asset pricing and market structure may hence incorporate the attention factor which captures how investors view and interpret the information content of news announcements.

Our study offers several directions for future research. One important direction is to examine the interaction between attention and trading patterns at ultra high-frequency (UHF) scales such milliseconds or nanoseconds. Studying UHF price dynamics may hence permit to uncover whether market reactions to news published on Twitter are driven by algorithmic traders, who can use textual analysis to automatically derive trading signals from tweet flows.[24] Another line would be to investigate the heterogeneity in market reaction depending on the credibility or reputation of the user who sends the tweet. For example, in

dynamics as shown for example by Huberman and Regev (2001).

[24] Unlike a machine, an investor often needs at least few seconds to read a tweet and pass an order. Therefore, the speed of reaction to news could also help disentangle pure algorithmic trading from human trading.

the case of Twitter earnings leak by a FinTech company called "Selerity Corp", the impact on financial markets was partially muted as practitioners had been debating online about the veracity of the message and figures provided by the Fintech company.[25] When Reuters Twitter account confirms Selerity information, market reacts strongly, which reflects how opinions and credibility could affect the speed of adjustment to news events. Last but not least, we believe that an interesting path for future research would be to analyze the role of dissemination in market liquidity at the intraday level, extending previous findings from Blankespoor et al. (2013) on the relation between firm-initiated news via Twitter, bid-ask spreads and abnormal depth. In this line, combining Twitter data with firm-initiated traditional press releases could help researchers understand the relation between the level of attention and information asymmetry in financial markets.

Acknowledgments

We thank Mardi Dungey, Michael Rockinger, Rogier Quaedvlieg and participants at the the 10th Annual SoFiE Conference and Panthéon-Sorbonne Research Seminar for valuable comments and discussions.

Appendix A - Implementation of the network algorithm

We consider all active users m in Twitter. As of January 1, 2016, m≈300 million. Relationships in Twitter are formalized on an $m \times m$ matrix, where $a_{i,j} = 1$ if user i follows user j, and $a_{i,j} = 0$ otherwise ($i \in m$, $j \in m$). We proceed as follows.

Step 1. We select 10 users i (i_1, i_2, ..., i_{10}) and we denote this list N_0. For each user $j \notin N_0$, we compute a variable of influence by defining $c_j = \sum_{i=i_1}^{i_{10}} a_{i,j}$.

Step 2. We sort users in descending order of influence c'_j. We add the first 50 users to N_0, and we denote this list N_1.

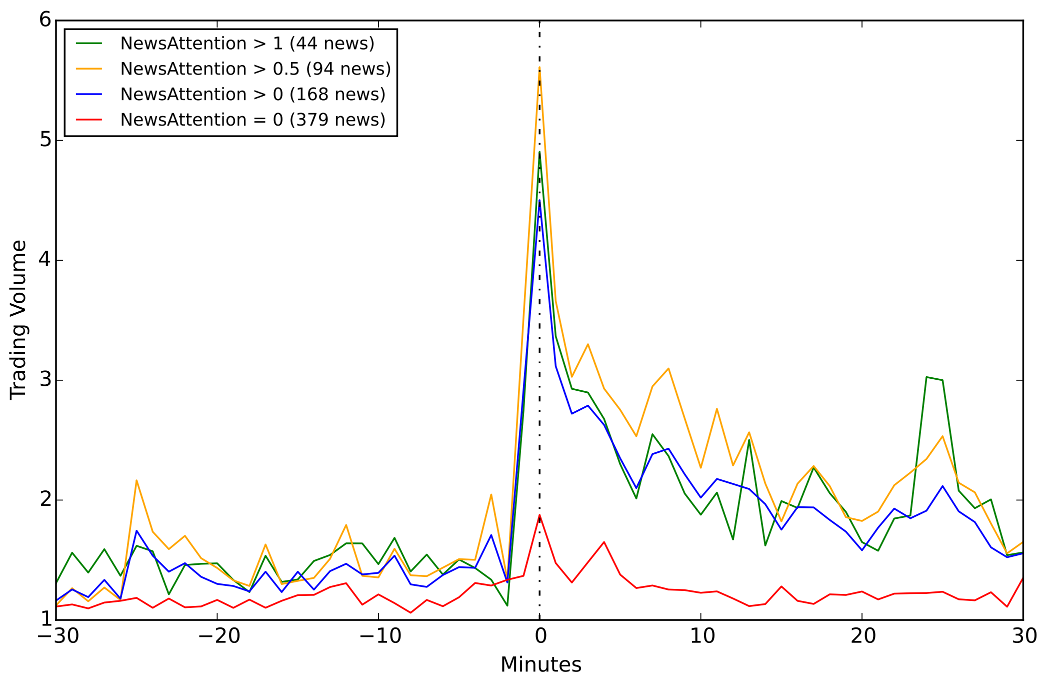
[25] As we discussed in Section 1 and illustrated on Figure 2.

Step 3. We select 60 users (i_1, i_2, ..., i_{60}) from N_1. For each user $j \notin N_1$, we compute a new variable of influence by defining $c'_j = \sum_{i=i_1}^{i_{60}} a_{i,j}$.

Step 4. We sort users in descending order of influence c_j. We add the first 50 users to N_1, and we denote this list N_2.

Step 5. We replicate step 3 and 4 until reaching network N_{60} composed of 3,010 users i (i_1, i_2,...i_{3010}).

Appendix B - Varying threshold values for retrieving attention



Notes: The figure presents the evolution of abnormal volume around the release of unscheduled HH with newswire-corrected timestamps. We consider HH when investors do not pay attention to news ($News_i^* = 0$) and for various threshold values to define attention-grabbing-news ($News_i^* > 0$; $News_i^* > 0.5$; $News_i^* > 1$).

References

Andersen, T. G., Bollerslev, T., Diebold, F. X., Vega, C., 2007. Real-time price discovery in global stock, bond and foreign exchange markets. Journal of International Economics 73, 251–277.

Andrei, D., Hasler, M., 2015. Investor attention and stock market volatility. The Review of Financial Studies 28, 33.

Balduzzi, P., Elton, E. J., Green, T. C., 2001. Economic news and bond prices: Evidence from the US treasury market. Journal of Financial and Quantitative Analysis 36, 523–543.

Barber, B. M., Odean, T., 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.

Bernile, G., Hu, J., Tang, Y., 2016. Can information be locked up? Informed trading ahead of macro-news announcements. Journal of Financial Economics 121, 496–520.

Blankespoor, E., Miller, G. S., White, H. D., 2013. The role of dissemination in market liquidity: Evidence from firms' use of Twitter. The Accounting Review 89, 79–112.

Bollerslev, T., Li, J., Xue, Y., 2016. Volume, volatility and public news announcements. Working Paper, Duke University .

Bollerslev, T., Todorov, V., Li, S. Z., 2013. Jump tails, extreme dependencies, and the distribution of stock returns. Journal of Econometrics 172, 307–324.

Boudt, K., Petitjean, M., 2014. Intraday liquidity dynamics and news releases around price jumps: Evidence from the DJIA stocks. Journal of Financial Markets 17, 121–149.

Boulland, R., Degeorge, F., Ginglinger, E., 2017. News dissemination and investor attention. Review of Finance 21, 761–791.

Bradley, D., Clarke, J., Lee, S., Ornthanalai, C., 2014. Are analysts recommendations informative? Intraday evidence on the impact of time stamp delays. The Journal of Finance 69, 645–673.

Busse, J. A., Green, T. C., 2002. Market efficiency in real time. Journal of Financial Economics 65, 415–437.

Corrado, C. J., 1989. A nonparametric test for abnormal security-price performance in event studies. Journal of Financial Economics 23, 385–395.

Da, Z., Engelberg, J., Gao, P., 2011. In search of attention. The Journal of Finance 66, 1461–1499.

Das, S. R., et al., 2014. Text and context: Language analytics in finance. Foundations and Trends (R) in Finance 8, 145–261.

DellaVigna, S., Pollet, J. M., 2009. Investor inattention and Friday earnings announcements. The Journal of Finance 64, 709–749.

Dewachter, H., Erdemlioglu, D., Gnabo, J.-Y., Lecourt, C., 2014. The intra-day impact of communication on euro-dollar volatility and jumps. Journal of International Money and Finance 43, 131–154.

Dimpfl, T., Jank, S., 2016. Can Internet search queries help to predict stock market volatility? European Financial Management 22, 171–192.

Driessen, J., Lin, T.-C., Van Hemert, O., 2013. How the 52-week high and low affect option-implied volatilities and stock return moments. Review of Finance pp. 369–401.

Erdemlioglu, D., Laurent, S., Neely, C. N., 2015. Which continuous-time model is most appropriate for exchange rates? Journal of Banking and Finance 61, 256–268.

Fama, E. F., 1998. Market efficiency, long-term returns, and behavioral finance. Journal of Financial Economics 49, 283–306.

Faust, J., Rogers, J. H., Wang, S.-Y. B., Wright, J. H., 2007. The high-frequency response of exchange rates and interest rates to macroeconomic announcements. Journal of Monetary Economics 54, 1051–1068.

Foucault, T., Hombert, J., Roşu, I., 2016. News trading and speed. The Journal of Finance 71, 335–382.

Groß-Klußmann, A., Hautsch, N., 2011. When machines read the news: Using automated text analytics to quantify high frequency news-implied market reactions. Journal of Empirical Finance 18, 321–340.

Hirshleifer, D., Teoh, S. H., 2003. Limited attention, information disclosure, and financial reporting. Journal of Accounting and Economics 36, 337–386.

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

Ifrim, G., Shi, B., Brigadir, I., 2014. Event detection in Twitter using aggressive filtering and hierarchical tweet clustering. In: *Second Workshop on Social News on the Web (SNOW), Seoul, Korea, 8 April 2014*, ACM.

Jegadeesh, N., Wu, D., 2013. Word power: A new approach for content analysis. Journal of Financial Economics 110, 712–729.

Jensen, M. C., 1978. Some anomalous evidence regarding market efficiency. Journal of Financial Economics 6, 95–101.

Kwak, H., Lee, C., Park, H., Moon, S., 2010. What is Twitter, a social network or a news media? In: *Proceedings of the 19th international conference on World wide web*, ACM, pp. 591–600.

Li, J., Yu, J., 2012. Investor attention, psychological anchors, and stock return predictability. Journal of Financial Economics 104, 401–419.

Loughran, T., McDonald, B., 2011. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. The Journal of Finance 66, 35–65.

Loughran, T., McDonald, B., 2016. Textual analysis in accounting and finance: A survey. Journal of Accounting Research 54, 1187–1230.

Mathioudakis, M., Koudas, N., 2010. Twittermonitor: trend detection over the Twitter stream. In: *Proceedings of the 2010 ACM SIGMOD International Conference on Management of data*, ACM, pp. 1155–1158.

Nardo, M., Petracco, M., Naltsidis, M., 2016. Walking down wall street with a tablet: a survey of stock market predictions using the web. Journal of Economic Surveys 30, 356–369.

Petrovic, S., Osborne, M., McCreadie, R., Macdonald, C., Ounis, I., Shrimpton, L., 2013. Can Twitter replace newswire for breaking news? In: Kiciman, E., Ellison, N. B., Hogan, B., Resnick, P., Soboroff, I. (eds.), *ICWSM*, The AAAI Press.

Ranaldo, A., 2008. Intraday market dynamics around public information arrivals. In: Wiley, J., Sons, H. N. (eds.), *Stock Market Liquidity: Implications for Market Microstructure and Asset Pricing*.

Riordan, R., Storkenmaier, A., Wagener, M., Zhang, S. S., 2013. Public information arrival: Price discovery and liquidity in electronic limit order markets. Journal of Banking and Finance 37, 1148–1159.

Shirky, C., 2008. Here comes everybody: The power of organizing without organizations. New York: Penguin .

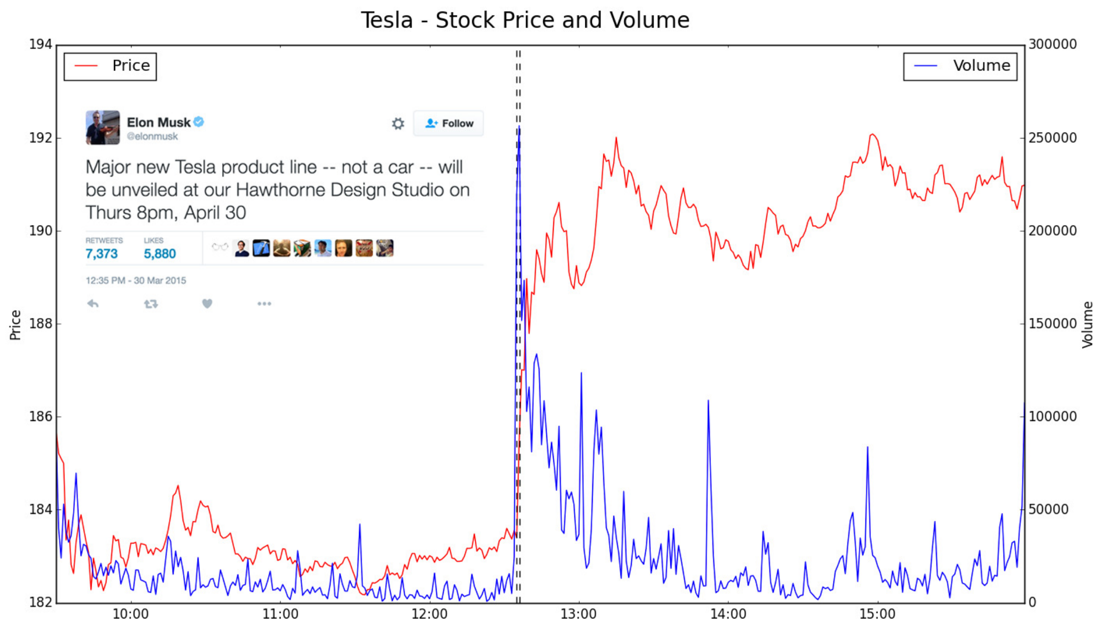
Smales, L. A., 2014. Non-scheduled news arrival and high-frequency stock market dynamics: Evidence from the australian securities exchange. Research in International Business and Finance 32, 122–138.

Solomon, D. H., Soltes, E., Sosyura, D., 2014. Winners in the spotlight: Media coverage of fund holdings as a driver of flows. Journal of Financial Economics 113, 53–72.

Wongswan, J., 2009. The response of global equity indexes to us monetary policy announcements. Journal of International Money and Finance 28, 344–365.

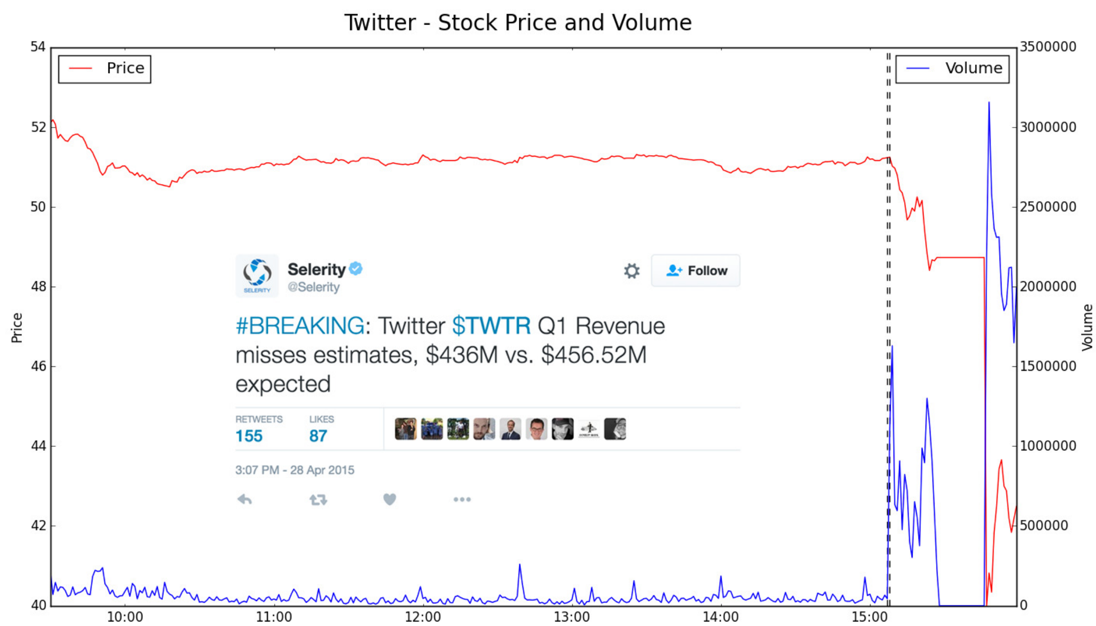
Yuan, Y., 2015. Market-wide attention, trading, and stock returns. Journal of Financial Economics 116, 548–564.

Fig. 1. Elon Musk tweet - Impact on Tesla stock price and trading volume



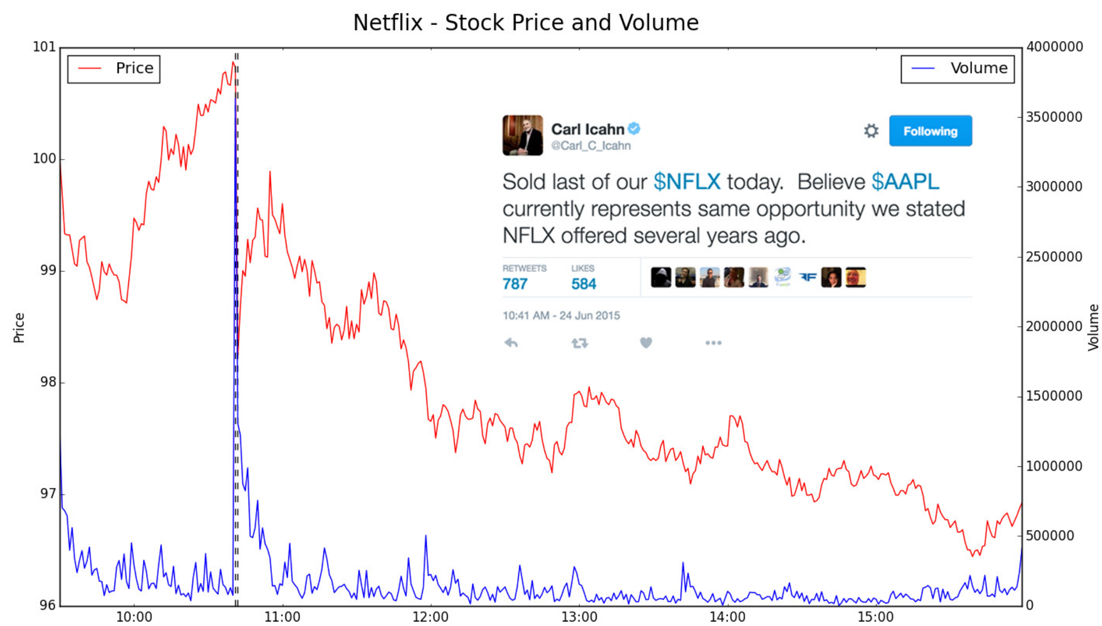
Notes: The figure illustrates the large increase in Tesla's stock price and trading volume following Elon Musk tweet announcing a new product line on March 30, 2015 at 12:35 p.m. For a complete story, read "Elon Musk tweet about new product line boosts Tesla shares" (MarketWatch, March 30, 2015).

Fig. 2. Selerity tweet - Impact on Twitter stock price and trading volume



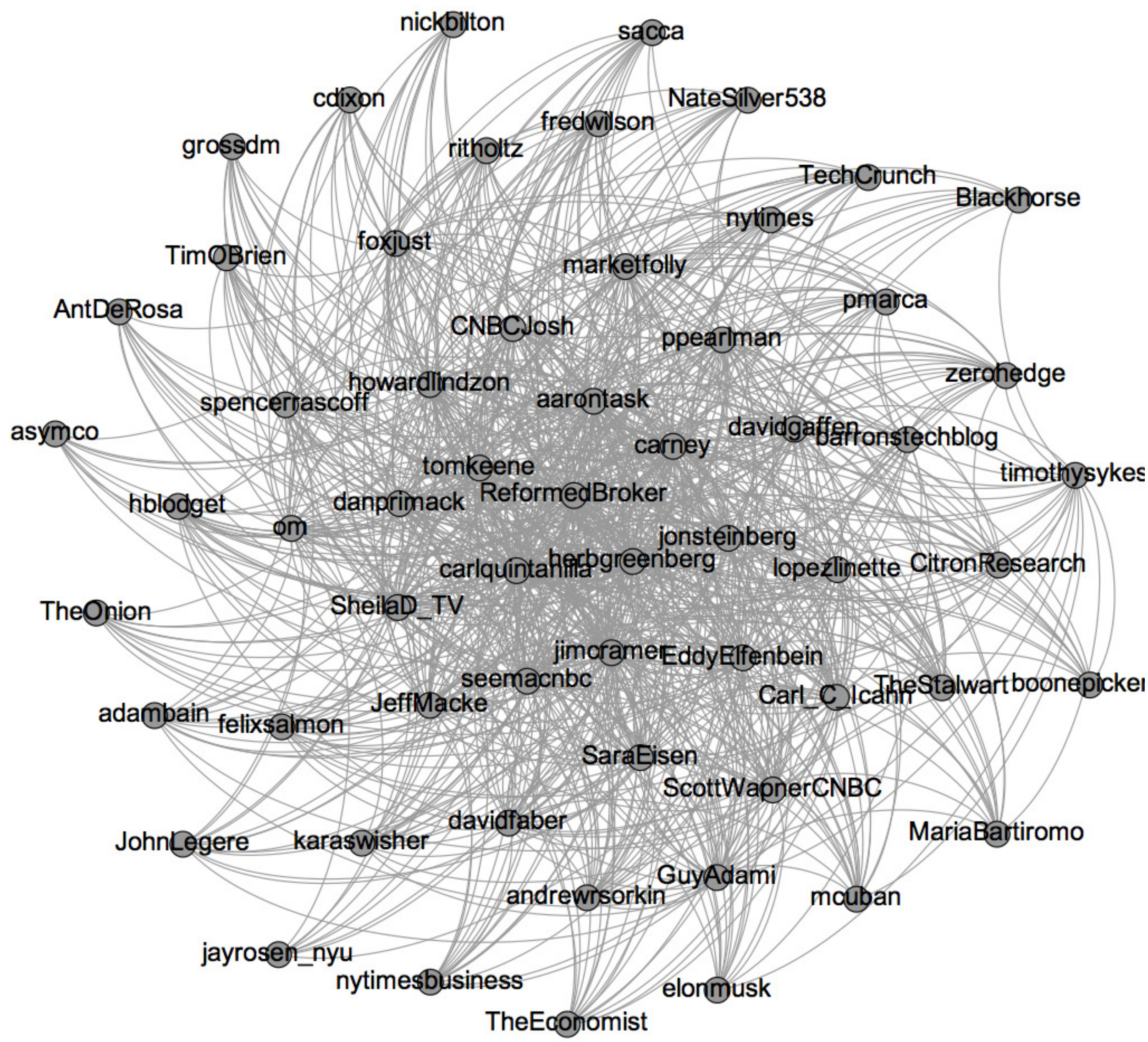
Notes: The figure shows the large decrease in Twitter's stock price and large increase in trading volume following Selerity tweet leaking Twitter earnings results on April 28, 2015 at 3:07 p.m. The leak prompted a NYSE trading halt for "news pending" starting at 3:27 p.m. For a complete story, read "The tweets that made Twitter stock crash" (MarketWatch, April 28, 2015).

Fig. 3. Carl Icahn tweet - Impact on Netflix stock price and trading volume



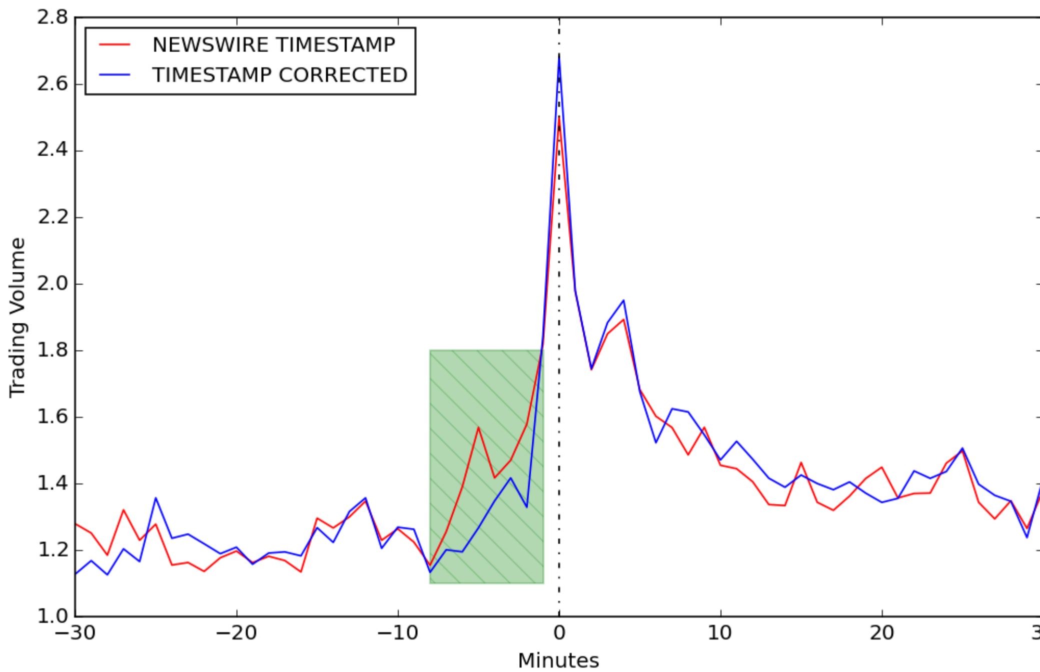
Notes: The figure illustrates the large decrease in Netflix stock price and large increase in trading volume following Carl Icahn tweet announcing that he sold his stake in Netflix on June 24, 2015 at 10:41 a.m. For a complete story, read “Carl Icahn sells his Netflix stock near record highs” (MarketWatch, June 24, 2015).

Fig. 4. Network N_1 - Twitter financial influencers



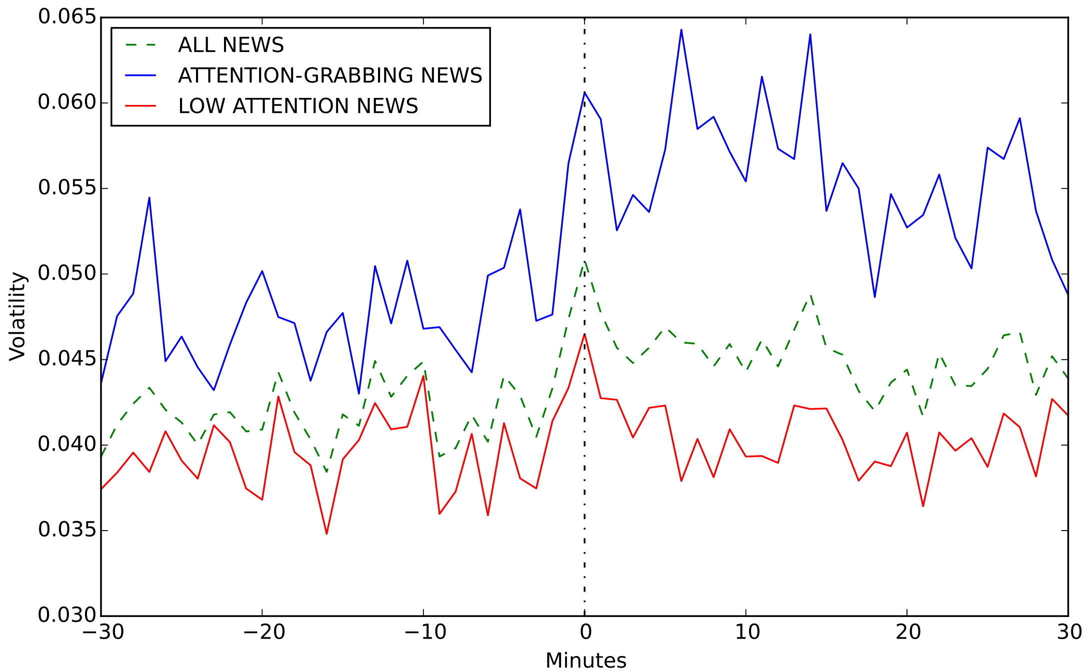
Notes: The figure shows the network structure N_1 (60 users). Each node represents a user and each link a directed friendships between two users. The graph is generated using Gephi, an open-source network analysis and visualization software.

Fig. 5. High-frequency volume patterns with (and without) newswire-corrected timestamps



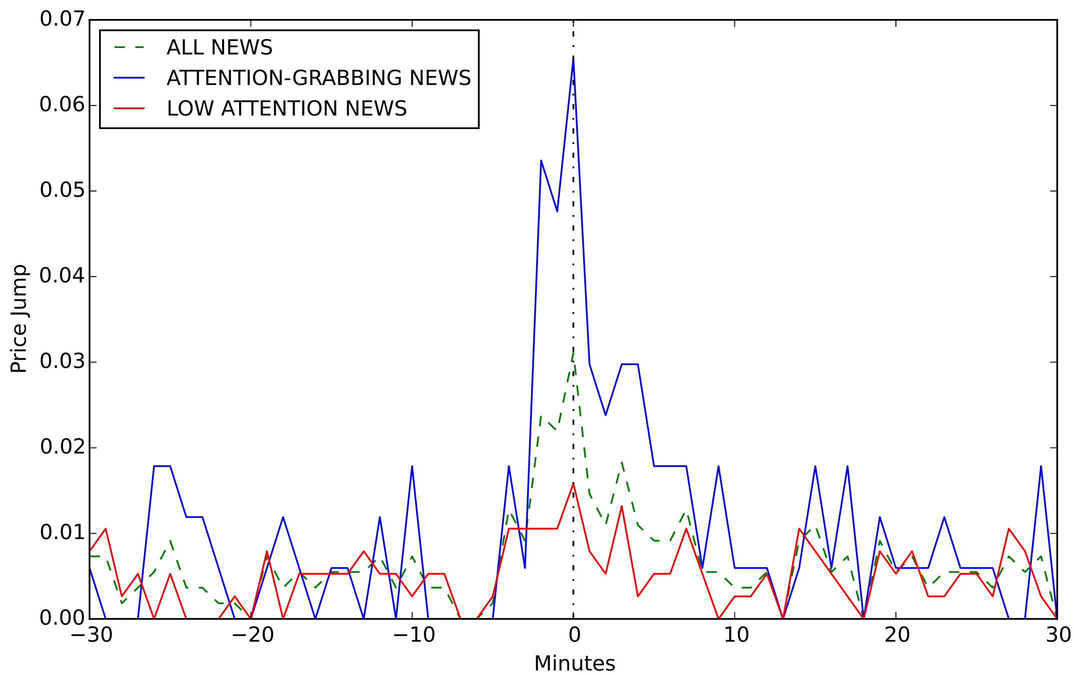
Notes: The figure displays the evolution of abnormal trading volume on a [-30:+30] minutes event-window around the release of unscheduled HH for "All news" (547 news stories) with and without corrected newswire timestamps. In green, we highlight the period during which abnormal trading volume is overestimated if timestamp delays are not corrected.

Fig. 6. High-frequency volatility patterns around information and attention



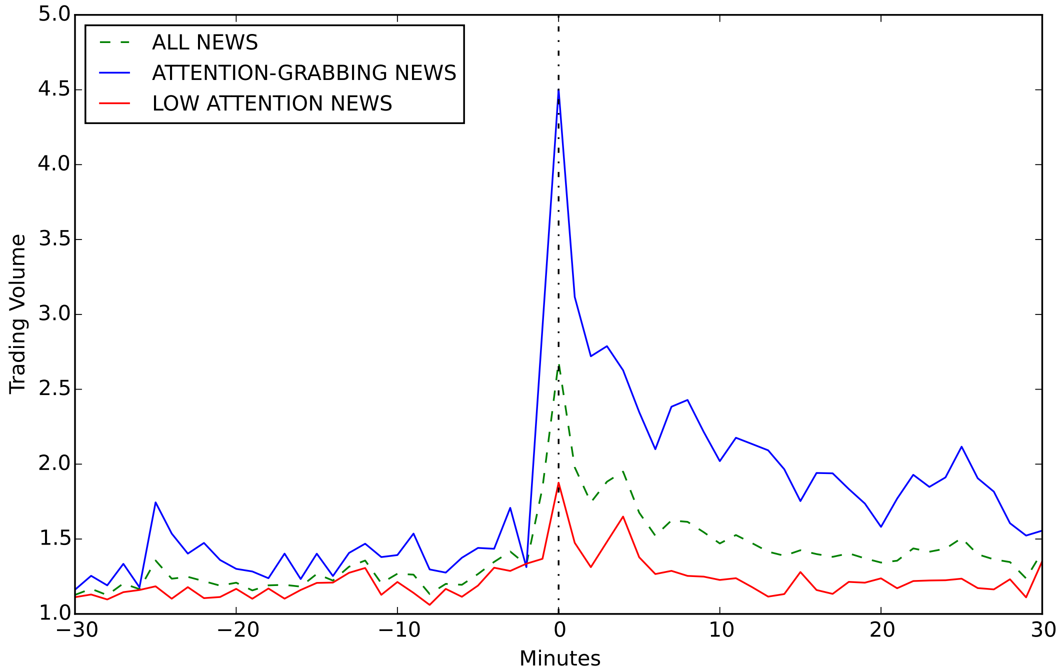
Notes: The figure shows the evolution of volatility on a [-30:+30] minutes event-window around the release of unscheduled HH with newswire-corrected timestamps, for "All news" (547 news stories), "Attention-grabbing news" ($News_i^* > 0$; 168 news stories) and "Low-attention news" ($News_i^* = 0$; 379 news stories).

Fig. 7. High-frequency jump patterns around information and attention



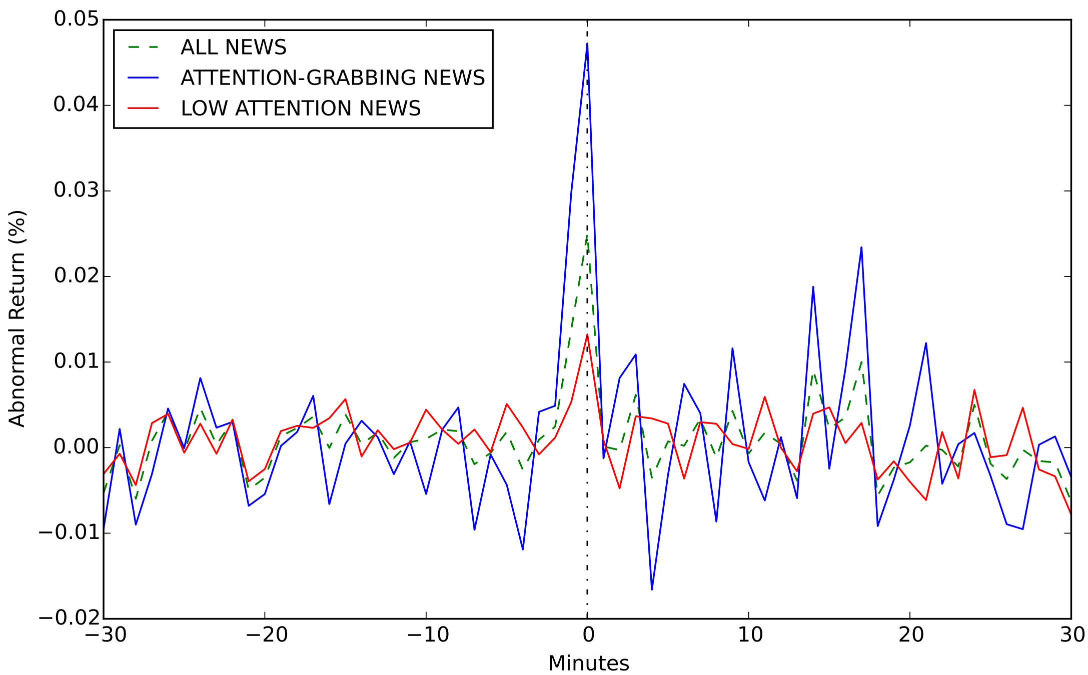
Notes: The figure shows the evolution of the average number of jumps on a [-30:+30] minutes event-window around the release of unscheduled HH with newswire-corrected timestamps, for "All news" (547 news stories), "Attention-grabbing news" ($News_i^* > 0$; 168 news stories) and "Low-attention news" ($News_i^* = 0$; 379 news stories).

Fig. 8. High-frequency volume patterns around information and attention



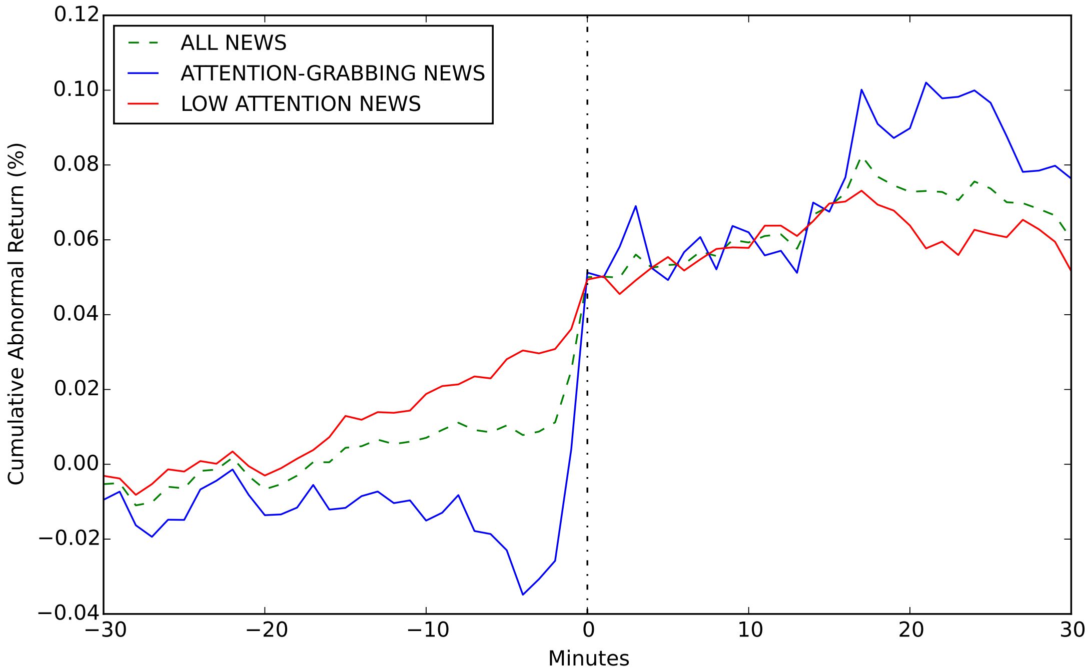
Notes: The figure illustrates the evolution of abnormal trading volume on a [-30:+30] minutes event-window around the release of unscheduled HH with newswire-corrected timestamps, for "All news" (547 news stories), "Attention-grabbing news" ($News_i^* > 0$; 168 news stories) and "Low-attention news" ($News_i^* = 0$; 379 news stories).

Fig. 9. High-frequency abnormal returns patterns around information and attention



Notes: The figure shows the evolution of abnormal returns on a [-30:+30] minutes event-window around the release of positive (long) and negative (short) unscheduled HH with newswire-corrected timestamps, for "All news" (358 news stories), "Attention-grabbing news" ($News_i^* > 0$; 124 news stories) and "Low-attention news" ($News_i^* = 0$; 234 news stories). HH with a neutral sentiment are removed from the analysis.

Fig. 10. High-frequency cumulative return patterns around information and attention



Notes: The figure shows the evolution of cumulative abnormal returns on a [-30:+30] minutes (rolling) event-window around the release of positive (long) and negative (short) unscheduled HH with newswire-corrected timestamps, for "All news" (358 news stories), "Attention-grabbing news" ($News_i^* > 0$; 124 news stories) and "Low-attention news" ($News_i^* = 0$; 234 news stories). HH with a neutral sentiment are removed from the analysis.

Table 1: The initial list of 10 influential users in N_0

user	username	description	follower	friend	message
@carlquintanilla	Carl Quintanilla	Fin. journalist, CNBC	119,306	3,799	4,696
@Carl_C_Icahn	Carl Icahn	Activist investor	285,997	107	267
@herbgreenberg	Herb Greenberg	Fin. journalist, The Street	240,000	702	25,903
@howardlindzon	Howard Lindzon	Fin. analyst, fmr hedge fund manager	253,405	1,682	104,860
@jimcramer	Jim Cramer	TV journalist, fmr hedge fund manager	948,924	428	65,442
@MariaBartiromo	Maria Bartiromo	Fin. journalist, FOX	186,602	1,048	12,011
@om	Om Malik	Venture capitalist, entrepreneur	1,532,690	1,242	44,494
@pmarca	Marc Andreesen	Investor, entrepreneur	546,562	7,465	96,096
@ReformedBroker	Joshua Brown	Fin. advisor	143,361	3,296	102,210
@timothysykes	Timothy Sykes	Stock trader, penny-stock expert	122,404	6,408	63,864

Notes: The table reports the descriptive statistics for the 10 *Twitterers* in N_0 (initial list). *"Follower"* represents the number of *Twitterers* who opt in to see tweets of each expert. *"Friend"* represents the number of *Twitterers* who have opted in to follow each expert. *"Message"* represents the total number of messages sent by each expert since the creation of the Twitter account.

Table 2: A sample of Twitter messages published on January 2, 2013

Date	Twitter	Content
2013-01-02 12:00:55	KeeneOnMarket	Reversals to the downside $AAPL $NFLX $CRM
2013-01-02 12:02:15	business	FLASH: Amazon wins dismissal of Apple s false advertising claim
2013-01-02 12:02:17	jyarow	AMAZON WINS DISMISSAL OF APPLE S FALSE ADVERTISING CLAIM over use of App Store.
2013-01-02 12:04:59	bespokeinvest	A number of triple digit priced growth stocks are already down 1%+ from the open. $AAPL $CMG $CRM $ISRG $LNKD $PCLN $PNRA
2013-01-02 12:06:08	sspencer_smb	pleased w/ how smb handled $AAPL today. focused on long to resist @555. then played reversal below 548. flexible
2013-01-02 12:06:13	bespokeinvest	$AAPL tested its 50-DMA at the open but is now at its lows for the day.
2013-01-02 12:10:26	SAI	Amazon Gets Apple s Lawsuit Over App Store Tossed Out $AMZN $AAPL by @jyarow http://t.co/Mrmv8R3l
2013-01-02 12:20:18	daveweigel	RT @keshasuxx: ANIMALS!! i m on the cover of @seventeenmag! check it amp get a copy on 1/8 [...]
2013-01-02 12:20:33	tlmontana	@bespokeinvest: $AAPL tested its 50-DMA at the open but is now at its lows for the day.
2013-01-02 12:20:35	Techmeme	Court dismisses Apple s case against Amazon for use of App Store name @markgurman / 9to5Mac [...]
2013-01-02 12:21:02	talkingbiznews	New York Times deputy investigations editor discussed Wal-Mart coverage: http://t.co/dLY53m39
2013-01-02 12:21:37	Reuters	Judge rejects Apple false advertising claims vs. Amazon http://t.co/f3O1DDky $AAPL $AMZN
2013-01-02 12:21:47	FastMoneyLydia	Cooperman: Sold down $AAPL position - didn t like the way it was acting: doesn t like cash policy. Prefers Qualcomm. @cnbcfastmoney
2013-01-02 12:23:13	MarketCurrents	Johnson amp Johnson JNJ declares $0.61/share quarterly dividend in line with previous. Forward yield ... http://t.co/k9DA4m5M $JNJ
2013-01-02 12:26:00	verge	Court rules Amazon s App Store isn t false advertising but full trademark lawsuit goes on http://t.co/8JvdLpdw
2013-01-02 12:27:25	arohan	@jtruman5 JJ tried but it requires your email address. Could you send me a request at http://t.co/Bas1X6zS
2013-01-02 12:27:42	murphyrosecliff	@FastMoneyLydia @cnbcfastmoney I m hoping those comments from Leon Cooperman re $AAPL cash policy on @cnbcfastmoney will reach [...]
2013-01-02 12:29:53	business	Amazon wins dismissal of Apple s false advertising claims http://t.co/VvD5shw2
2013-01-02 12:31:17	ReutersBiz	Judge rejects Apple false advertising claims vs. Amazon http://t.co/170gupcv
2013-01-02 12:33:31	RedDogT3	@TrueChartTrader @ep_capital no one is 100%. I try and be consistent. As far as $aapl. I have a lot of happy followers there
2013-01-02 12:35:13	TechCrunch	Court Rejects Apple s False Advertising Claim In App Store Trademark Lawsuit http://t.co/0WDYOUV8 by @sarahintampa
2013-01-02 12:37:39	sharkbiotech	just went through my trades from last year 56% winners avg gain $1.09 avg loss .88 20% of my gains were $AAPL related 1 down month April
2013-01-02 12:38:26	jonnajarian	fine with the call sold $AAPL $VMW amp $NFLX this am 4 a trade RT @Pete_Romano: Najarian just cant admit he was wrong to go all cash.
2013-01-02 12:39:21	Benzinga	Smartphone Preference: 38% of Gamers Choose Google s Android Over Apple s iOS: http://t.co/WkJQfr37 $AAPL $GOOG $YHOO
2013-01-02 12:40:20	business	Samsung loses bid to seal sales data in Apple dispute http://t.co/SzS5TLwh
2013-01-02 12:45:18	bizrpt	Judge rejects Apple false advertising claims vs. Amazon http://t.co/pdhJko6A @bizrpt $AAPL $AMZN
2013-01-02 12:46:49	savitz	Acacia Research In Settlement With Apple In Patent Case: The patent licensing firm Acacia Research this morning ... http://t.co/mavLxG0B
2013-01-02 12:53:02	Chris_Ciaccia	How many of you will care less about the iPhone 6 than you do Ray Lewis or Ed Reed? http://t.co/JzbtrAfr
2013-01-02 12:56:21	CNBC	U.S. judge granted Amazon s bid to end part of Apple lawsuit over Amazon s use of the term APP STORE - http://t.co/udu5fAWj $AMZN $AAPL
2013-01-02 12:59:04	CNBCTopStories	Judge Rejects Apple s False Advertising Claims Against Amazon http://t.co/QOC1dRTZ

Notes: The table provides all messages sent on Twitter by experts from N_{60} on January 2, 2013 between 12 p.m. and 1 p.m.

Table 3: Cosine similarity example between a Bloomberg news and tweet flow

Source	Time	Content	Similarity
BBG	11:25:12	EINHORN DROPS SUIT AGAINST APPLE OVER SHAREHOLDER VOTE	
TWT	11:21:40	Ex-Apple marketing guru Guy Kawasaki now advising Motorola http://t.co/NWc1tpoZ4F	0.0
TWT	11:21:50	I'd like to get up to about 400 contracts on the sell side of these $AAPL calls assuming she stays below $435 for rest of the day basically	0.0
TWT	11:22:54	BREAKING NEWS: Government sources confirm that horsemeat has been found in a significant percentage of $AAPL products.	0.0
TWT	11:23:32	See also http://t.co/FfDdSFj4pL MT @juliakmarsh: Thank you PO Walke and Sgt Whitley at Transit Dt 32 for finding my iPhone this morning!	0.0
TWT	11:23:36	And then of course if I had short 400 contracts I d be poppin molly and sweatin ... WOOOOOO! $AAPL	0.0
TWT	11:23:45	David Einhorn withdraws lawsuit against Apple. Manhattan federal court approves	0.14
TWT	11:24:05	Can Apple and Google Win This New Market? http://t.co/h0wEPStwIe	0.0
TWT	11:25:01	Risk/reward of $AAPL settlement at any time IMHO heavily favors $VHC and that doesn't even include that I think they win against $CSCO.	0.05
TWT	11:25:27	Einhorn Drops Suit Against Apple Over Shareholder Vote $AAPL	1.0

Notes: The table presents an example of cosine similarity between the Bloomberg HH "Einhorn drops suit against Apple over shareholder vote" and all messages published on Twitter about Apple on a [-15:0] minutes around the release of the news event. We mark the tweets related to Bloomberg HH in red.

Table 4: News release times for Twitter versus Bloomberg

Date	News Provider	Content
2013-08-13 14:22:26	Bloomberg	ICAHN HAS LARGE POSITION IN APPLE ICAHN SAYS ON TWITTER
2013-08-13 14:21:29	Carl_C_Icahn	We currently have a large position in APPLE. We believe the company to be extremely undervalued. Spoke to Tim Cook today [...]
2014-10-08 12:04:55	Bloomberg	APPLE SENDS OUT INVITATIONS FOR OCT 16 EVENT IN CUPERTINO WSJ
2014-10-08 12:01:56	geoffreyfowler	Invites out for Apple event in Cupertino on October 16. http://t.co/gQRSp7YsMY
2015-02-13 13:23:55	Bloomberg	APPLE SAID TO HIRE AUTO EXPERTS TO WORK IN RESEARCH LAB FT
2015-02-13 13:22:26	tim	About those Apple car rumours... Apple is hiring automotive experts to work in a secret research lab @FT sources say [...]
2013-07-12 14:28:36	Bloomberg	GE SAYS ENGINES NOT TIED TO BOEING FIRE AT HEATHROW REUTERS
2013-07-12 14:24:41	firstadopter	It s NOT US!!! $GEsays$BA fire doesn t involve engines
2015-04-20 12:04:35	Bloomberg	GE SAID TO BE IN TALKS TO SELL COMMERCIAL LENDING BUSINESS DJ
2015-04-20 12:02:26	d4ytrad3	$GE in discussion to sell commercial lending biz
2015-05-06 15:30:08	Bloomberg	BLACKSTONE SAID TO BE AMONG BIDDERS FOR GE LENDING UNIT FT
2015-05-06 15:29:01	ftfinancenews	Blackstone joins race for GE lending unit http://t.co/1p2vviyuzH
2014-07-18 10:13:54	Bloomberg	IBM CUT TO SELL VS HOLD AT SOCIETE GENERALE EARLIER
2014-07-18 09:59:48	zozotrader	IBM cut at SocGen
2014-11-24 14:16:29	Bloomberg	ICAHN HAS ABSOLUTELY NO INVOLVEMENT IN IBM CNBC S WAPNER
2014-11-24 14:13:43	ScottWapnerCNBC	Sources tell me @Carl_C_Icahn has absolutely no involvement in $IBM. Stock had moved earlier on rumor that he did.
2015-10-27 13:51:24	Bloomberg	IBM LEARNED IN AUG SEC CONDUCTING PROBE ON REVENUE RECOGNITION
2015-10-27 13:47:24	Livesquawk	IBM says that the SEC is investigating the company in relation to revenue recognition - Rtrs $IBM
2014-01-15 12:53:34	Bloomberg	CARLYLE PAYING 4B 4 2B FOR J J BLOOD TESTING UNIT WSJ
2014-01-15 12:40:36	MikeSpectorWSJ	@OneCarlyle paying between $4 billion and $4.2 billion for J amp J blood-testing unit. @WSJ scoop: http://t.co/Uho1t5AUC6
2014-03-20 10:21:00	Bloomberg	JNJ 1.2B JUDGMENT OVERTURNED BY ARKANSAS SUPREME COURT AP
2014-03-20 10:18:50	YahooNews	Arkansas Supreme Court overturns $1.2B judgment against Johnson & Johnson over drug marketing @AP
2015-04-23 11:00:12	Bloomberg	J&J BOOSTS QTRLY DIV TO 75C SHR FROM 70C EST 74C
2015-04-23 10:59:43	OpenOutcrier	$JNJ Increasing div
2013-03-08 10:25:30	Bloomberg	LESLIE DACH OF WAL MART TO LEAVE COMPANY POLITICO REPORTS
2013-03-08 10:18:30	mikeallen	Wal-Mart CEO Mike Duke tells associates: Leslie Dach executive vice president corporate affairs leaving in June after 7 yrs [...]
2013-10-15 12:15:23	Bloomberg	WAL MART TO CLOSE UNDER PERFORMING STORES IN BRAZIL AND CHINA
2013-10-15 12:15:19	sbanjo	Walmart closing stores in Brazil China that aren't profitable revises down square footage from 20-22 million to 14 million
2015-12-11 12:48:08	Bloomberg	Walmart COM BEGINS SELLING APPLE WATCH TECHCRUNCH
2015-12-11 12:46:09	TechCrunch	Walmart com Begins Selling The Apple Watch https://t.co/KLZ3gbnuep by @sarahintampa

Notes: The table reports the 15 (selected) cases where Twitter effectively "breaks the news". For each news, the first line represents Bloomberg reported timestamp with the associated Bloomberg headline. The second line presents the first mention of the news on Twitter, the user who "breaks the news" and the tweet content.

Table 5: Event-study results without (with) timestamp correction

	Volatility		Jump		Volume		Return (%)	
TIME	B-TIME	C-TIME	B-TIME	C-TIME	B-TIME	C-TIME	B-TIME	C-TIME
[−30 : −26]	0.417	0.417	0.006	0.005	1.253	1.158	-0.002*	-0.001
[−29 : −25]	0.421	0.421	0.006	0.005	1.253	1.204	-0.001	-0.000
[−28 : −24]	0.420	0.418	0.005	0.005	1.233	1.217	0.000	0.001
[−27 : −23]	0.419	0.417	0.004	0.005	1.229	1.242	0.001	0.002
[−26 : −22]	0.417	0.414	0.004	0.005	1.192	1.245	0.001	0.002**
[−25 : −21]	0.416	0.412	0.003	0.004	1.182	1.249	0.000	0.001
[−24 : −20]	0.409	0.411	0.002	0.002	1.165	1.220	0.000	-0.000
[−23 : −19]	0.417	0.419	0.003	0.003	1.167	1.204	-0.000	-0.001
[−22 : −18]	0.421	0.420	0.003	0.003	1.171	1.193	0.001	-0.000
[−21 : −17]	0.419	0.416	0.003	0.004	1.177	1.188	0.000	-0.000
[−20 : −16]	0.415	0.412	0.003	0.004	1.169	1.187	0.001	0.001
[−19 : −15]	0.418	0.413	0.004	0.005	1.188	1.198	0.002**	0.002**
[−18 : −14]	0.410	0.407	0.004	0.005	1.209	1.212	0.002	0.002
[−17 : −13]	0.412	0.413	0.004	0.005	1.233	1.236	0.002	0.002
[−16 : −12]	0.417	0.418	0.006	0.005	1.268	1.269	0.002	0.001
[−15 : −11]	0.428	0.429	0.006	0.005	1.288	1.273	0.001	0.001
[−14 : −10]	0.435	0.436	0.005	0.006	1.281	1.274	0.001	0.001
[−13 : −9]	0.429	0.432	0.005	0.005	1.273	1.282	0.002	0.001
[−12 : −8]	0.422	0.422	0.005	0.005	1.244	1.245	0.002	0.001
[−11 : −7]	0.421	0.420	0.004	0.004	1.225	1.214	0.002	0.001
[−10 : −6]	0.414	0.412	0.004	0.003	1.257	1.212	0.002	0.001
[−9 : −5]	0.415	0.410	0.004	0.002	1.318	1.211	0.001	0.001
[−8 : −4]	0.423	0.417	0.007	0.004	1.357	1.228	0.001	-0.000
[−7 : −3]	0.424	0.419	0.009*	0.005	1.420	1.285	0.001	-0.000
[−6 : −2]	0.431	0.422	0.012***	0.010**	1.485	1.311	0.002	0.000
[−5 : −1]	0.444	0.436	0.016***	0.014***	1.570*	1.440	0.004	0.003
[−4 : 0]	0.457*	0.450	0.020***	0.020***	1.757**	1.724**	0.008***	0.008***
[−3 : 1]	0.466**	0.460*	0.020***	0.020***	1.869***	1.850**	0.007**	0.008***
[−2 : 2]	0.470**	0.470**	0.019***	0.020***	1.923***	1.916**	0.006***	0.008***
[−1 : 3]	0.467**	0.473**	0.018***	0.019***	1.977***	2.027***	0.006***	0.009***
[0 : 4]	0.469**	0.470**	0.016***	0.017***	1.828***	1.847***	0.004	0.006**
[1 : 5]	0.463**	0.462**	0.012***	0.013***	1.754**	1.755***	-0.000	0.001
[2 : 6]	0.457**	0.458*	0.010**	0.012***	1.719**	1.731**	0.000	0.001
[3 : 7]	0.462**	0.459**	0.010**	0.012***	1.646**	1.678**	0.001	0.001*
[4 : 8]	0.459*	0.458*	0.008	0.010**	1.581**	1.597**	0.002	-0.000
[5 : 9]	0.459**	0.459*	0.007	0.008	1.536**	1.556**	0.001	0.001
[6 : 10]	0.453*	0.453*	0.006	0.007	1.504*	1.556*	0.000	0.001
[7 : 11]	0.454*	0.454*	0.005	0.006	1.472*	1.526*	0.001*	0.002
[8 : 12]	0.451	0.451	0.004	0.005	1.442*	1.486*	0.001*	0.001
[9 : 13]	0.460**	0.455*	0.002	0.004	1.395	1.455*	0.001	0.000
[10 : 14]	0.457*	0.461**	0.005	0.004	1.396*	1.446*	0.002**	0.001
[11 : 15]	0.462**	0.464**	0.006	0.006	1.376*	1.420*	0.004**	0.002**
[12 : 16]	0.466**	0.462**	0.007	0.006	1.359*	1.402*	0.003*	0.002**
[13 : 17]	0.458**	0.459**	0.007	0.007	1.364*	1.400*	0.003	0.004**
[14 : 18]	0.450*	0.450*	0.007	0.007	1.380*	1.396*	0.003	0.004*
[15 : 19]	0.443	0.440	0.006	0.007	1.378*	1.380*	0.002	0.002*
[16 : 20]	0.439	0.437	0.005	0.005	1.380*	1.371*	-0.001	0.001
[17 : 21]	0.430	0.430	0.006	0.006	1.390*	1.382*	-0.001	0.000
[18 : 22]	0.438	0.434	0.006	0.005	1.392*	1.384*	-0.001	-0.002
[19 : 23]	0.439	0.437	0.007	0.006	1.401*	1.397*	-0.001	-0.001
[20 : 24]	0.441	0.437	0.005	0.005	1.411*	1.430*	-0.001	0.000
[21 : 25]	0.435	0.437	0.005	0.005	1.409*	1.439*	0.000	0.000
[22 : 26]	0.440	0.446	0.004	0.005	1.393*	1.424*	-0.000	-0.001
[23 : 27]	0.443	0.449*	0.005	0.005	1.389*	1.410**	-0.000	-0.001
[24 : 28]	0.445*	0.448*	0.005	0.005	1.350*	1.370*	0.000	-0.000
[25 : 29]	0.443*	0.451*	0.005	0.006	1.326*	1.352*	-0.001	-0.002
[26 : 30]	0.446*	0.450**	0.004	0.005	1.324*	1.322*	-0.001	-0.003
Event	547	547	547	547	547	547	547	547

Notes: The table reports the significance of volatility, jump, abnormal volume, and abnormal returns for each 5 minutes interval during a [-30:+30] event window around the release of unscheduled news announcements. We compare the results when considering Bloomberg reported timestamp as the event minute ("B-TIME") and when considering the first mention of the news on Twitter (when social media "breaks the news") and Bloomberg reported timestamp otherwise ("C-TIME"). *, ** and *** denote significance respectively at the 10% level, 5% level and 1% level. Significance is assessed using the non-parametric Corrado rank test.

Table 6: Event-study results with high- versus low-attention (1-minute intervals)

	Volatility		Jump		Volume		Return (%)	
	HIGH	LOW	HIGH	LOW	HIGH	LOW	HIGH	LOW
-30	0.044	0.037	0.006	0.008	1.162	1.112	-0.009	-0.003
-29	0.048	0.038	0.000	0.011	1.255	1.130	0.002	-0.001
-28	0.049	0.040	0.000	0.003	1.191	1.097	-0.009	-0.004
-27	0.054	0.038	0.000	0.005	1.335	1.145	-0.003	0.003
-26	0.045	0.041	0.018	0.000	1.178	1.160	0.005	0.004
-25	0.046	0.039	0.018	0.005	1.745	1.184	-0.000	-0.001
-24	0.045	0.038	0.012	0.000	1.537	1.102	0.008	0.003
-23	0.043	0.041	0.012	0.000	1.403	1.179	0.002	-0.001
-22	0.046	0.040	0.006	0.000	1.474	1.105	0.003	0.003
-21	0.048	0.037	0.000	0.003	1.360	1.113	-0.007	-0.004
-20	0.050	0.037	0.000	0.000	1.301	1.167	-0.005	-0.003
-19	0.047	0.043*	0.006	0.008	1.284	1.101	0.000	0.002
-18	0.047	0.040	0.012	0.000	1.238	1.170	0.002	0.003
-17	0.044	0.039	0.006	0.005	1.403	1.102	0.006*	0.002
-16	0.047	0.035	0.000	0.005	1.233	1.160	-0.007	0.003
-15	0.048	0.039	0.006	0.005	1.402	1.207	0.000	0.006
-14	0.043	0.040	0.006	0.005	1.253	1.209	0.003	-0.001
-13	0.050	0.042	0.000	0.008	1.407*	1.275	0.001	0.002
-12	0.047	0.041	0.012	0.005	1.469	1.307*	-0.003	-0.000
-11	0.051	0.041	0.000	0.005	1.380	1.128	0.001	0.001
-10	0.047	0.044	0.018	0.003	1.393	1.213	-0.005	0.004
-9	0.047	0.036	0.000	0.005	1.537*	1.141	0.002	0.002
-8	0.046	0.037	0.000	0.005	1.297	1.060	0.005	0.000
-7	0.044	0.041	0.000	0.000	1.276	1.167	-0.010	0.002
-6	0.050	0.036	0.000	0.000	1.375*	1.115	-0.001	-0.001
-5	0.050	0.041	0.000	0.003	1.441	1.189	-0.004	0.005*
-4	0.054	0.038	0.018	0.011	1.435*	1.309	-0.012	0.002
-3	0.047	0.037	0.006	0.011	1.709	1.287	0.004	-0.001
-2	0.048	0.041	0.054***	0.011	1.313	1.335	0.005	0.001
-1	0.056	0.043	0.048***	0.011	2.914**	1.368	0.030*	0.005*
0	0.061	0.047	0.065***	0.016**	4.502***	1.877***	0.047***	0.013**
1	0.059	0.043	0.030***	0.008	3.116***	1.475*	-0.001	0.001
2	0.053	0.043	0.024*	0.005	2.720***	1.312	0.008	-0.005
3	0.055	0.040	0.030***	0.013*	2.788***	1.482*	0.011	0.004
4	0.054	0.042	0.030***	0.003	2.627***	1.650	-0.017	0.003
5	0.057	0.042	0.018	0.005	2.348***	1.378**	-0.003	0.003
6	0.064	0.038	0.018	0.005	2.100***	1.266	0.007	-0.004
7	0.058**	0.040	0.018	0.011	2.384**	1.288	0.004	0.003
8	0.059	0.038	0.006	0.005	2.429**	1.254*	-0.009	0.003
9	0.057	0.041	0.018	0.000	2.216**	1.249	0.012	0.000
10	0.055	0.039	0.006	0.003	2.020**	1.227	-0.002	-0.000
11	0.062*	0.039	0.006	0.003	2.176*	1.239	-0.006	0.006
12	0.057	0.039	0.006	0.005	2.135*	1.179	0.001	0.000
13	0.057*	0.042	0.000	0.000	2.092*	1.116	-0.006	-0.003
14	0.064*	0.042	0.006	0.011	1.966**	1.133	0.019	0.004
15	0.054	0.042	0.018	0.008	1.753	1.280**	-0.002	0.005
16	0.056*	0.040	0.006	0.005	1.941**	1.160	0.009*	0.001
17	0.055	0.038	0.018	0.003	1.939*	1.134	0.023	0.003
18	0.049	0.039	0.000	0.000	1.834**	1.214	-0.009	-0.004
19	0.055	0.039	0.012	0.008	1.737*	1.209	-0.004	-0.002
20	0.053	0.041	0.006	0.005	1.581*	1.238*	0.003	-0.004*
21	0.053	0.036	0.006	0.008	1.770**	1.171	0.012**	-0.006
22	0.056	0.041	0.006	0.003	1.929**	1.220	-0.004	0.002
23	0.052	0.040	0.012	0.003	1.848**	1.223	0.000	-0.004
24	0.050	0.040	0.006	0.005	1.912	1.225	0.002	0.007
25	0.057	0.039	0.006	0.005	2.117**	1.235	-0.003	-0.001
26	0.057	0.042	0.006	0.003	1.906***	1.173*	-0.009	-0.001
27	0.059*	0.041	0.000	0.011	1.817**	1.164	-0.010	0.005
28	0.054	0.038	0.000	0.008	1.606*	1.231	0.000	-0.003
29	0.051	0.043	0.018	0.003	1.523	1.111	0.001	-0.003
30	0.049	0.042	0.000	0.000	1.556	1.352*	-0.003	-0.008
Event	168	379	168	379	168	379	124	368

Notes: The table reports the significance of volatility, jump, abnormal volume, and abnormal returns for each minute during a [-30:+30] event window around the release of high-attention and low-attention news events. We consider (as minute 0) the newswire corrected timestamp. HIGH (LOW): the news receiving high (low) attention from market participants. *, ** and *** denote significance respectively at the 10% level, 5% level and 1% level. Significance is assessed using the non-parametric Corrado rank test.

Table 7: Event-study results with high- versus low-attention (5-minute intervals)

	Volatility		Jump		Volume		Return (%)	
	HIGH	LOW	HIGH	LOW	HIGH	LOW	HIGH	LOW
[−30 : −26]	0.048	0.039	0.005	0.005	1.224	1.129	-0.003	-0.000
[−29 : −25]	0.048	0.039	0.007	0.005	1.341	1.143	-0.001	0.000
[−28 : −24]	0.048	0.039	0.010	0.003	1.397	1.138	0.000	0.001
[−27 : −23]	0.047	0.040	0.012	0.002	1.439	1.154	0.002	0.002
[−26 : −22]	0.045	0.040	0.013	0.001	1.467	1.146	0.004	0.002
[−25 : −21]	0.046	0.039	0.010	0.002	1.504	1.137	0.001	0.000
[−24 : −20]	0.046	0.039	0.006	0.001	1.415	1.133	0.000	-0.000
[−23 : −19]	0.047	0.040	0.005	0.002	1.364	1.133	-0.001	-0.000
[−22 : −18]	0.048	0.039	0.005	0.002	1.331	1.131	-0.001	0.000
[−21 : −17]	0.047	0.039	0.005	0.003	1.317	1.131	-0.001	0.000
[−20 : −16]	0.047	0.039	0.005	0.004	1.292	1.140	-0.001	0.002
[−19 : −15]	0.047	0.039	0.006	0.005	1.312	1.148	0.000	0.003
[−18 : −14]	0.046	0.039	0.006	0.004	1.306	1.170	0.001	0.003
[−17 : −13]	0.046	0.039	0.004	0.006	1.340	1.191	0.001	0.002
[−16 : −12]	0.047	0.040	0.005	0.006	1.353	1.232	-0.001	0.002
[−15 : −11]	0.048	0.041	0.005	0.006	1.382	1.225	0.000	0.001
[−14 : −10]	0.048	0.042	0.007	0.005	1.380	1.226	-0.001	0.001
[−13 : −9]	0.048	0.041	0.006	0.005	1.437*	1.213	-0.001	0.002
[−12 : −8]	0.047	0.040	0.006	0.005	1.415	1.170	-0.000	0.001
[−11 : −7]	0.047	0.040	0.004	0.004	1.377	1.142	-0.001	0.002
[−10 : −6]	0.047	0.039	0.004	0.003	1.376	1.139	-0.002	0.002
[−9 : −5]	0.047	0.038	0.000	0.003	1.385	1.134	-0.002	0.002
[−8 : −4]	0.049	0.039	0.004	0.004	1.365	1.168	-0.004	0.002
[−7 : −3]	0.049	0.039	0.005	0.005	1.447*	1.213	-0.004	0.002
[−6 : −2]	0.050	0.039	0.015*	0.007	1.454*	1.247	-0.002	0.001
[−5 : −1]	0.051	0.040	0.025***	0.009*	1.762*	1.298	0.005	0.003
[−4 : 0]	0.053	0.041	0.038***	0.012***	2.374**	1.435*	0.015*	0.004
[−3 : 1]	0.054	0.042	0.040***	0.011**	2.711**	1.468**	0.017**	0.004
[−2 : 2]	0.055	0.043*	0.044***	0.010**	2.913***	1.473**	0.018***	0.003
[−1 : 3]	0.057	0.043*	0.039***	0.011**	3.208***	1.503**	0.019***	0.004*
[0 : 4]	0.055*	0.042	0.036***	0.009*	3.151***	1.559**	-0.000	0.001
[1 : 5]	0.056*	0.041	0.026***	0.007	2.720***	1.460**	0.001	0.000
[2 : 6]	0.058**	0.041	0.024***	0.006	2.517***	1.418*	0.001	0.002
[3 : 7]	0.059**	0.040	0.023***	0.007	2.449***	1.413*	-0.003	0.002
[4 : 8]	0.059**	0.040	0.018*	0.006	2.377***	1.367*	0.002	0.001
[5 : 9]	0.059*	0.039	0.015*	0.005	2.295***	1.287*	0.003	0.000
[6 : 10]	0.058**	0.040	0.013	0.005	2.230***	1.257	-0.000	0.002
[7 : 11]	0.058*	0.039	0.011	0.004	2.245**	1.251	-0.001	0.002
[8 : 12]	0.058**	0.040	0.008	0.003	2.195**	1.230	-0.000	0.001
[9 : 13]	0.059**	0.040	0.007	0.002	2.128**	1.202	0.001*	0.001
[10 : 14]	0.059**	0.041	0.005	0.004	2.078**	1.179	0.001	0.002
[11 : 15]	0.058**	0.041	0.007	0.005	2.025**	1.189	0.004**	0.001
[12 : 16]	0.057**	0.041	0.007	0.006	1.978**	1.173	0.009*	0.002
[13 : 17]	0.056*	0.040	0.010	0.005	1.938**	1.164	0.008*	0.002
[14 : 18]	0.054*	0.040	0.010	0.005	1.887**	1.184	0.003	0.001
[15 : 19]	0.054*	0.039	0.011	0.005	1.841**	1.199*	0.004	-0.001
[16 : 20]	0.053	0.039	0.008	0.004	1.807**	1.191	0.005	-0.003
[17 : 21]	0.053	0.039	0.008	0.005	1.772**	1.193	-0.000	-0.003
[18 : 22]	0.054*	0.039	0.006	0.005	1.770**	1.210	0.001	-0.003
[19 : 23]	0.053	0.040	0.008	0.005	1.773**	1.212	0.003	-0.001
[20 : 24]	0.054*	0.039	0.007	0.005	1.808**	1.215	0.001	-0.000
[21 : 25]	0.054*	0.040	0.007	0.005	1.915**	1.215	-0.003	0.001
[22 : 26]	0.055*	0.040	0.007	0.004	1.942**	1.215	-0.004	0.001
[23 : 27]	0.055*	0.040	0.006	0.005	1.920**	1.204	-0.004	0.001
[24 : 28]	0.056*	0.040	0.004	0.006	1.871**	1.206	-0.004	-0.001
[25 : 29]	0.054*	0.041	0.006	0.006	1.794**	1.183	-0.004	-0.002
[26 : 30]	0.446*	0.450**	0.005	0.005	1.682**	1.206	-0.001	-0.003
Event	168	379	168	379	168	379	124	368

Notes: The table reports the significance of volatility, jump, abnormal volume, and abnormal retursn for each 5-minute interval during a [-30:+30] event window around the release of high-attention and low-attention news events. We consider (as minute 0) the newswire corrected timestamp. HIGH (LOW): the news receiving high (low) attention from market participants. *, ** and *** denote significance respectively at the 10% level, 5% level and 1% level. Significance is assessed using the non-parametric Corrado rank test.