The Good, the Bad, and the Social Media: Financial Implications of Social Media Reactions to Firm-Related News

Abstract

Firms and investors often react to financial news on social media. However, how they react to news of different nature and whether their reactions influence the stock market are far from clear. Employing data from multiple sources, we examine how the tweeting intensities of firms and investors vary with the sentiment and uncertainty of financial news, as well as how the changes in their tweeting intensities affect stock returns. Our analysis reveals several interesting findings. First, firms are more responsive to news with positive sentiment and low uncertainty, whereas investors are more responsive to news with high uncertainty. Second, when the tweeting intensities of firms and investors surge, the stock returns of focal firms rise. Third, investors tweet more about smaller firms facing uncertain news, and the increased volume of investor tweets, in turn, has a stronger effect on the stock returns of smaller firms.

Keywords: news, social media, sentiment, uncertainty, stock return, information environment

INTRODUCTION

The advent of social media has significantly upended communication channels between firms and their investors. The Securities and Exchange Commission (SEC) quickly caught up to the era of social media by ruling that firms could treat social media as legitimate, nonexclusive outlets for communication. Consequently, the public no longer needs to rely exclusively on traditional news media or analysts to consume or produce relevant information. Firms now use social media as direct channels to disseminate information and interact with interested parties [18,19,21,32,41].

Undoubtedly, Twitter is one of the most popular social media channels used by firms to disclose information to the public. By posting tweets on their official Twitter accounts, firms can make their information environment more transparent. We call the tweets posted by firms “firm tweets”, which fall into the realm of professionally generated content (PGC). Firm tweets can reduce information asymmetry between firms and investors [4,38]. Accordingly, Blankespoor et al. [4] find that the dissemination of firm-initiated news through Twitter is associated with lower abnormal bid-ask spreads and greater abnormal depths. Firms may also respond to news strategically on social media to influence their stock performance. Prior studies suggest that firms should make strategic decisions on the usage of social media under different circumstances, such as managing product-harm crises [21] and disseminating earnings announcements [19].

Social media users, including investors, also share their opinions about financial news and stock performance [7,12]. Taking Twitter as an example, investors often tweet about firms with cashtags [3]. The cashtag of a firm is a dollar sign (“$”) followed by the firm’s stock ticker, such as “$GOOG”. We call these tweets “investor tweets” as they are from people who follow stocks

[12]. Investor tweets are a form of user-generated content (UGC). Extant literature finds that the volume and sentiment of UGC are predictive of the stock price movement [3,7,12,23,42]. However, little is known about how investors on social media respond to financial news that may affect the stock prices of firms.

Prior studies investigating the impact of financial news on stock price movements focus on how the volume and sentiment of news predict the financial performance of firms [2,11,39]. Beyond the volume and the sentiment of news, the uncertainty in news is also a critical aspect to consider in the financial market [43], whereas its role in influencing the financial performance of firms is overlooked in extant research. Moreover, firms and investors, two key participants in the financial market, may both respond to financial news on social media and influence the stock price movements.[1] For example, the information posted on social media by firms (PGC) and investors (UGC) may resolve the uncertainty in the news. While news, PGC, and UGC may affect stock prices collectively, existing studies about how news and social media affect the financial performance of firms are highly fragmented and largely ignore the interactive nature of the social media channel. Specifically, existing studies typically examine either how news affects stock price movements or how social media content by firms/investors affects stock prices, but not how firms and investors react to the uncertainty in news on social media or how their social media reactions influence stock prices.

This paper seeks to provide a fuller picture of how these three types of information (i.e., news, UGC, and PGC) collectively influence stock returns. In doing so, it is critical to address the endogeneity of social media activities rather than assuming it away to provide practically meaningful insights for firms and investors. We investigate the simultaneity among stock price

[1] In Online Appendix A, we provide two examples showing how firm and investors react to news on social media, as well as the abnormal stock returns on the same day.

movements, firm tweets, and investor tweets, and use instruments and event studies to establish causality among news, UGC, PGC, and stock price movements, rather than discovering association or Granger causality (i.e., a type of relationship useful for prediction but does not necessarily imply true causality), a focus of most prior research. We also provide an understanding of how the effect of social media activities on stock price movements is moderated by the sizes of firms. Large firms have higher visibility and a more transparent information environment than small firms. When the uncertainty of firm-specific news is high, social media posts may help small firms overcome the lack of coverage from traditional media. The potential differences between small and large firms in this regard are not well-understood yet and we attempt to fill in the gap with this study.

To overcome the limitations of extant research and also to provide a full picture of the nuanced interactions among firm tweets, investor tweets, and the stock market, in terms of how they react to financial news, we address the following questions:

1) *How do firms and investors react to the sentiment and uncertainty of financial news on social media?* This question allows us to better understand the interactive nature of social media and shed light on the neglected role of uncertainty in financial news.

2) *How do the social media reactions of firms and investors affect the stock performance of the focal firms?* The impact of firms' and investors' social media reactions on the stock market has not been studied yet and this study may generate novel insights in that regard.

3) *Do firms, investors, and the stock market react differently to financial news related to large and small firms?* This question intends to investigate the potential asymmetric role of social media for large and small firms, in terms of how information is generated and absorbed, which have important implications for both managers and investors.

To address these questions, we compile a rich dataset from multiple sources, including the financial news about firms, the tweets of firms and investors, and the daily stock prices of the firms. We use a simultaneous equations model with instruments to analyze the data at the news event level. Several intriguing findings emerge from this research. (1) Firms and investors react to news sentiment and uncertainty differently on social media, i.e., firms are more responsive to positive news and news with low uncertainty, while investors are more responsive to news with high uncertainty. (2) When the volume of firm tweets or that of investors tweets increases, the stock returns of focal firms improve. (3) Investors are more responsive to uncertain news related to small firms than that related to large firms and their increased tweeting intensities on social media, in turn, have a larger effect on the stock returns of small firms. In fact, the increased tweeting intensities of investors only improve the stock returns of small firms, but not those of large firms.

We make the following contributions to the literature. First, this paper not only provides a comprehensive picture of how firms, investors, and the stock market react to the information in financial news, but also reveals the nuanced interactions among these three parties. In particular, we show that two key market participants (firms and investors) respond to financial news on social media and their social media reactions can influence stock price movements. Second, this paper heightens the overlooked role of uncertainty in financial news and further demonstrates that firms and investors have opposite reactions to the uncertainty in news, shedding light on the differential social media strategies and objectives of these two parties. Third, this paper demonstrates the value of social media in improving firms' information environment and illustrates intriguing differences between large and small firms, in terms of how information is generated on social media and how such information is absorbed by the stock market.

RELATED WORK

The financial performance of a firm can be influenced by multiple sources of public information related to the firm, including news from conventional media (e.g., newspapers and TV), UGC on social media, and PGC on social media. Accordingly, there are three streams of research regarding the relationship between public information and firm performance.

The first and earliest stream of research studies the impact of news on firm performance. Earlier studies along this line investigate how the overall stock market responds to macroeconomic news, such as monthly unemployment rates [5] and bond rating changes [16]. Later work began to examine how the volume and/or sentiment of firm-specific news, such as press mentions and annual earnings announcements, affect the stock price movements of the individual firms. Conrad et al. [11] find that the stock market responds to negative earnings surprises of firms more strongly as the market level rises, while the response to positive earnings surprises is not affected by the market level. Tetlock et al. [39] find that the fraction of negative words in firm-specific news forecasts low earnings, especially when the news focus on fundamentals. Alanyali et al. [2] show that the daily stock transaction volume is positively associated with a firm's daily number of mentions in the *Financial Times*. These studies suggest that the volume and sentiment of firm-specific news are predictive of the stock price movement.

The second stream of research examines the relationship between UGC and firm performance. Luo et al. [23] show that social media content is a significant predictor of firms' stock returns and has a faster and stronger predictive power than other online behavioral metrics such as Google search trends and web traffic. Chen et al. [7] further demonstrate the volume and sentiment of user comments on social media articles are also predictive of firms' future stock returns. Yu et al. [42] find that social media content (e.g., blogs, forums, and Twitter) has a stronger association with stock returns than conventional media (e.g., newspapers and TV), and different types of social media interrelate with conventional media in influencing stock returns. These studies demonstrate that the magnitude and sentiment of UGC on social media are also predictive of the stock price movement. Deng et al. [12] study the bidirectional relationship between social media sentiment and stock returns. They find that the effect of microblog sentiment on stock returns is both statistically and economically significant and stock returns have a larger effect on negative microblog sentiment than on positive microblog sentiment.

The third stream of research focuses on firms' strategic release or dissemination of PGC on social media to influence the stock price movement. Gordon et al. [17] find that the voluntary disclosure of items concerning information security is positively associated with the market value of a firm. Schniederjans et al. [32] show that the impression management strategies on social media are positively associated with firms' quarterly earnings per share. Blankespoor et al. [4] show that the additional dissemination of firm-initiated news via Twitter is associated with lower abnormal bid-ask spreads and greater abnormal depths, which implies that social media may help reduce information asymmetry. Lee et al. [21] show that the usage of social media attenuates the negative reaction of the stock market to product recalls, but the effect is weaker for interactive social media (e.g., Twitter) that firms have less control over. Jung et al. [19] find that firms are less likely to disseminate earnings announcements via social media when the news is bad and when the magnitude of earnings miss is greater. These studies demonstrate that strategic disclosure or dissemination of firm information on social media can influence the stock market.

Despite the significant research interest in how different types of information influence stock returns, there are several major limitations to extant research. (1) While news, UGC, and PGC may influence the stock return of a firm collectively, studies in this area are highly fragmented and typically focus on one or two types of information. More importantly, the interactive nature of social media (i.e., UGC and PGC may react to financial news) is largely ignored in the stock price movement process. This paper seeks to provide a full picture of how these three types of information collectively influence stock returns. (2) While causality is essential for managers and investors, extant research typically does not address the endogeneity of social media content and focuses on either association or Granger causality [12,23]. This paper strives to establish causality using instruments to provide practically meaningful insights to the participants of the financial market. (3) Extant research primarily focuses on the volume and sentiment of information and ignores the uncertainty in information, which plays a critical role in the stock market. This paper attempts to fill this gap by delving into the uncertainty in financial news. (4) The role of social media as a potential equalizer for large and small firms with different information environments is largely overlooked in extant research. This paper tackles this problem by studying how the reactions from social media and the stock market vary with the size of a firm.

THEORETICAL BACKGROUND

Press news and social media are two key communication channels influencing the information environment of firms in the financial market. In this section, build on extant literature, we discuss how firms and investors may react to the sentiment and uncertainty of financial news on social media to improve the information environment, as well as the impact of their social media reactions on stock returns due to the informational value of social media posts. Given that large and small firms generally have very different information environments, we further discuss the potential moderating role of firm size in this process. Figure 1 summarizes our conceptual framework. Note, while social mediation reactions and stock returns can be interdependent, this paper focuses on the impact of the former on the latter, as the latter is what matters the most in the financial market.

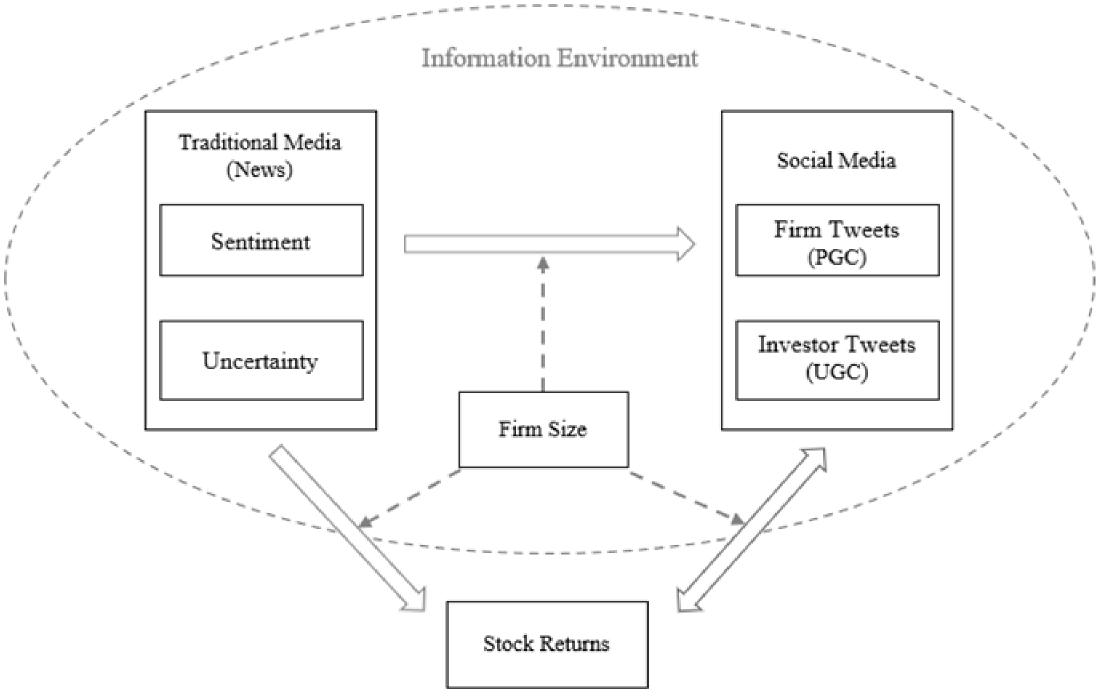

Figure 1. Conceptual Framework

Effects of News Sentiment on Social Media Reactions

Firms often use social media for impression management [32]. As such, their reactions to financial news may depend on the sentiment of news. When the sentiment of the news is positive, firms may seize this opportunity to spread the good news or to promote their products and services, rendering them more active than usual on social media. On the other hand, when the sentiment of the news is negative, remaining active on social media may invite unnecessary criticisms. Along this line, Jung et al. [19] find that firms are less likely to disseminate earnings news via social media when the news is negative. Based on these arguments, firms are more likely to be active in a positive than in a negative news cycle.

The sentiment of financial news may also affect the social media activities of investors. According to the loss aversion theory [20], investors are more sensitive to losses than to gains. Consequently, investors may pay more attention to negative news than to positive news. In support of this argument, Conrad et al. [11] find that the abnormal returns of stocks are more sensitive to negative earnings surprises than to positive earnings surprises. Deng et al. [12] further show that stock returns are responsive to negative microblog sentiment, but not to positive microblog sentiment. If investors' activities on social media are aligned with their behavior in the stock market, they may be more responsive to negative news than to positive news on social media and post more in a negative news cycle to exchange information that can help stakeholders minimize potential losses.

Effects of News Uncertainty on Social Media Reactions

When the uncertainty of financial news is high, firms may share their private information via social media to resolve the uncertainty. However, without fully understanding the long-term consequence of the news, this strategy can be risky as sharing inaccurate or inappropriate information can result in legal issues. Therefore, a safer option for firms is to stay quiet until the consequence of the news becomes clear. In fact, using a game-theoretical model, Suijs [37] shows that a firm may prefer not to disclose its private information when investors' responses to the information disclosure are unclear. Hence, firms are expected to be less active on social media when the uncertainty of news is high.

From the perspective of investors, the uncertainty in firm-related financial news entails both opportunities and risks. To decide whether they should buy, sell, or hold the focal firms' stocks, investors need to obtain as much information as possible in a timely manner. Social media is a perfect outlet for this purpose as it often beats mainstream media in tracking news events in real-time [34]. Therefore, when the uncertainty of financial news is high, investors may use social media more actively to seek and share information, essentially leveraging community intelligence to remove ambiguous information [29]. As such, investors' social media activities related to the focal firms are expected to increase with the uncertainty in financial news.

Effects of Social Media Reactions on Stock Returns

Social media represent an exciting and emerging source of information to the stock market [3]. Firms can improve their information environment by sharing information on social media. For example, Blankespoor et al. [4] find that the dissemination of PGC via Twitter can help reduce information asymmetry among investors. Further, Lee et al. [21] demonstrate that firms can attenuate the stock market's negative reaction to product recalls by posting more tweets than they normally would. By being more active on social media upon the release of financial news, firms can reduce the information asymmetry related to the news and associated risks, which in turn can improve their stock performance.

Similarly, investors can also provide valuable information to the stock market by posting information and insights on social media [3]. The informational value of the social media activities by investors may influence the stock market. Prior studies have shown that the volume and sentiment of UGC on social media are predictive of stock returns [7,12,23,42]. For instance, Luo et al. [23] show that an unexpected increase in positive blog posts is associated with a surge in the daily stock return. Thus, the increased activities of investors on social media may reduce the information asymmetries for the focal firms and improve their financial performance.

Moderating Effects of Firm Size

The information environment of a firm largely depends on the size of the firm. Compared to large firms, small firms have lower visibilities on conventional media and hence typically have a less transparent information environment [4,43]. Due to the higher information asymmetry for smaller firms, investors and firms may benefit more from exchanging information on social media, rendering them more responsive to financial news on smaller firms. In particular, when financial news with high uncertainty is released, investors and firms may use social media more actively than usual to exchange information to resolve the uncertainty. In this process, investors may pay more attention to small firms whose information is scanter than large firms, and smaller firms are also more incentivized to use social media to improve their information environment and reduce their stock volatility.

Owing to the limited information available for small firms, one piece of financial news may result in dramatic changes in what the public knows about these firms. Consequently, financial news with the same level of sentiment may lead to larger volatility in the stocks of smaller firms. Indeed, it is well-documented that small firms generally have larger stock volatility than larger firms [10]. As such, the effect of news sentiment on abnormal returns is expected to be weaker for large firms than for small firms.

Firm size may also moderate how the social media reactions of investors and firms influence the stock market. Since small firms tend to have a less transparent information environment than large firms, the informational value of investors' social media activities might be more pronounced for small firms. In line with this argument, Blankespoor et al. [4] find that the effectiveness of social media in reducing information asymmetry is limited to firms with low visibility. Thus, the increased social media activities of investors upon the release of financial news may have a stronger effect on the stock returns of small firms than on those of large firms.

DATA

Data Collection

To collect the data for this study, we continuously monitored the tweets posted by a sample of firms and the tweets mentioning these firms for one whole year. It is crucial to keep monitoring Twitter because some tweets might be removed after they were posted, and more importantly because the tweets mentioning firms were collected through the streaming API of Twitter and cannot be retrieved retrospectively. To make this time-consuming data collection task manageable, following prior work in the literature [12,23], we focused on firms in a specific sector, instead of all publicly traded firms. Specifically, we chose the technology sector because it is the largest and fastest-growing sector in the stock market.[2] We selected two major categories of technology firms in the North American Industry Classification System (i.e., 519130 – Internet & Web Search industry and 511210 – Software industry) so that our sample is not too narrowly focused. We obtain 348 firms from these two categories.

We download the tweets, stock prices, and accounting numbers of all the 348 firms. We also collect news and tweets that mentioned these firms. We focus on data about these firms from Jan 1, 2014, through Dec 31, 2014. News articles about the focal firms are downloaded from Factiva. We use the firm keys in Factiva (i.e., ticker symbols) to search for news about these firms, in which we also require the title or leading paragraph of a news article to contain the company name. We include all major news publishers in the Factiva dataset such as the Dow Jones Newswires and the Wall Street Journal. Since 70 out of 348 firms have no news in the year 2014, we are left with 278 firms for our analysis.

We collect the focal firms' Twitter account names and use the Twitter API to download all the tweets posted by each firm. These tweets are what we call "firm tweets", as they are the tweets posted by firms under their own Twitter accounts. We also collect tweets about the focal firms from potential investors. Following prior studies in the literature [3,12], we use cashtags to identify investors' tweets about the focal firms. In doing so, we focus on the overall population

[2] The technology sector accounts for 27% of S&P 500 in 2020, up from 19% in 2014. See https://seekingalpha.com/article/2319245-updated-s-and-p-500-sector-weightings and https://seekingalpha.com/article/4355143-nasdaq-tech-growth-keep-going.

of investors that actively discuss stocks on Twitter using cashtags and do not differentiate individual and institutional investors. Compared to using company names to identify firm-related posts, using stock symbols (or cashtags) "is much more accurate and is currently the best practice" [12]. We download all the tweets containing the cashtags of the focal firms. We call these tweets "investor tweets" as they are posted by users watching the stocks of firms.

We download the accounting fundamentals of the focal firms from COMPUSTAT and daily stock prices from CRSP. The firm-level summary statistics of the raw data from different information sources are provided in Table 1. Among the 278 sample firms, 93 firms did not post any tweets in the year 2014, including 34 firms without a Twitter account. Except for 10 firms, all the other firms were mentioned by investors at least once on Twitter. The daily stock prices are available for 235 firms in the CRSP database.

Table 1. Firm-Level Summary Statistics of Raw Data

Data Source	**Mean**	**SD**	**Min**	**Median**	**Max**	**#Firms**
Press Media						
News	28.6	25.0	1	25	155	278
PGC on Social Media						
Firm tweets	948.0	1362.4	1	422	7,111	185
UGC on Social Media						
Investor tweets	1100.8	1980.9	1	503.5	17,100	268
COMPUSTAT & CRSP						
Market capitalization (billion USD)	3.73	19.97	0.01	0.53	262.11	235

Variable Construction

To investigate how investors and firms respond to news and how their responses affect the stock performance of the respective firms, we construct and analyze our data at the news event level, following the event study approach [24,26]. Specifically, for each piece of news about a focal firm, we construct variables to measure the abnormal tweeting intensities of investors and the firm, as well as the abnormal stock return for the firm, on the same day. We also consider the cumulative abnormal return in a two-day window (i.e., the day the news was released and the following day) in our robustness checks.

To eliminate the potential interference among different news about the same firm on adjacent days, we focus on the sets of isolated news that occurred on the same day for each firm. Here, a set of the news posted on the same day is considered isolated if there is no other news about the same firm 48 hours before (after) the earliest (latest) news in the set. We use 48 hours as a cutoff because the stock market is known to be rather efficient to absorb the information in news, often in a matter of hours, minutes, or even seconds [1,26,33]. Hence, we believe a 48-hour buffer is large enough to ensure that different news events on the same firm do not interfere with each other. In addition to eliminating interference, focusing on isolated news can mitigate the over-representation issue of large firms in the resulting data, since they generally receive much more attention from the press than small firms do.

After removing non-isolated news, 100 news on non-trading days (e.g., weekends), and 98 news in the first two weeks (as will be explained soon, the abnormal tweeting intensities of firms and investors are not defined in this period), we are left with 3,927 unique sets of news on 277 firms, among which 3,285 sets include only one piece of news. For simplicity, we refer to each set of news on a given firm as one event and the date of the event as the event day. Next, we discuss the dependent variables (abnormal return and abnormal tweeting intensities of firms and investors), focal independent variables (sentiment and uncertainty of news), the moderator (firm size), and the control variables used in this study.

Abnormal Return: We use the abnormal return of a firm's stock on the event day to measure the financial performance of the firm. Abnormal return is a crucial measure to understand the impact of events on stock performance, which is generally computed as the difference between the firm's actual and expected stock returns on a given day [24]. The expected return can be estimated using a variety of models. However, since the choice of expected return model has little effect on the estimation of abnormal returns in a short time window [24], we choose the classical market model which is used in the vast majority of event studies in the finance literature.[3] The market model predicts the expected firm return using a linear regression model [24].[4] Since the distribution of abnormal returns is highly skewed, we use the inverse hyperbolic sine (IHS) function to transform the abnormal returns [6]. The IHS function (i.e., $\log\left(x+\sqrt{x^2+1}\right)$) behaves similarly to a log function but is an odd function that extends to negative values, which has been widely used in income and wealth-related studies as one's net wealth can be negative [15]. Figure 2 shows the histogram of abnormal returns before and after the IHS transformation, which demonstrates the effectiveness of the IHS transformation in reducing skewness.

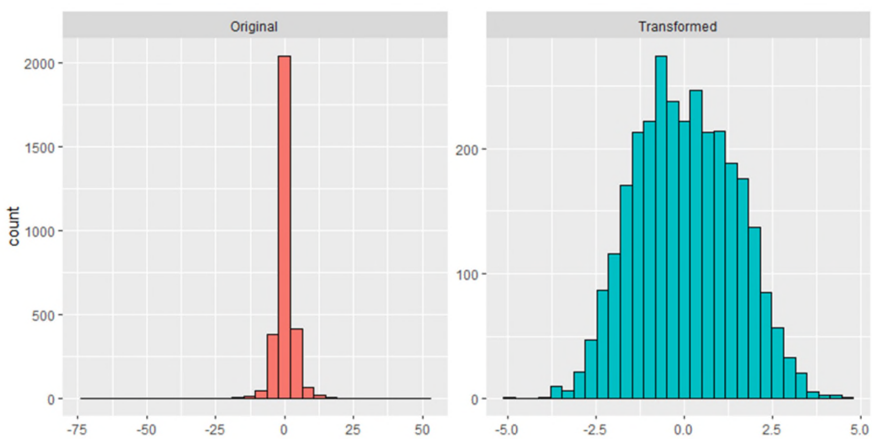

Figure 2. Histograms of Abnormal Returns before and after IHS Transformation

[3] https://www.eventstudytools.com/expected-return-models.
[4] We use the SAS code provided by WRDS to calculate the abnormal return, with a 180-day estimation window (i.e., [-190, -11]) to train the expected return model. We exclude the 10 days prior to the event from the estimation window to prevent the expected return model from picking up any information that was leaked to the market beforehand. Please see https://wrds-www.wharton.upenn.edu/pages/support/applications/event-studies/event-study-research-application/ for more details. As a robustness check, we also consider the extended Fama-French four-factor model [13,23] and the resulting abnormal returns have a correlation of 0.98 with those generated by the market model.

Abnormal Investor and Firm Tweets: We use the number of abnormal investor tweets about the focal firm on the event day to measure investors' reactions to the news. Specifically, the number of abnormal investor tweets is defined as the difference between the number of investor tweets mentioning the focal firm on the event day and the average number of investor tweets mentioning the focal firm in the past two weeks.[5] This approach is highly similar to how Lee et al. [21] measure firms' abnormal tweeting activities around product recalls. We apply the log transformation on these two numbers before taking the difference as their distributions are highly skewed. Similarly, the variable *abnormal_firm_tweets* measures the change in the log-transformed number of firm tweets on the event day. Using abnormal tweets allows us to remove potential time-varying firm-level effects, which cannot be captured by firm-level fixed effects.

Sentiment and Uncertainty of News: While there are many general-purpose tools to analyze the sentiment of texts (e.g., VADER, sentimentr, and Google Natural Language API), their performances are not satisfying on financial news as we test them on our data. This is not surprising because the polarities of many words are dramatically different in normal contexts and the financial market [22]. For example, "beat" is generally considered a negative word, but it has a positive meaning in financial news (e.g., "beats Wall Street estimates"). "Probe" is a neutral word in generic news, but is highly negative in financial news (e.g., "SEC probes Facebook"). To obtain a reliable measure of the sentiment, we recruit an average of 16 workers from Amazon Mechanical Turk (AMT) to manually assess the impact of each piece of news on the stock price on a 5-point Likert scale, from highly positive to highly negative.[6] Then, the sentiment of each

---

[5] The results are similar if we use the average daily tweet volume of investors in the past four or eight weeks as a proxy for the normal tweet volume of investors on a firm. Comparatively, a two-week window allows us to better capture the increase or decrease in the normal tweet volume over time.

[6] We invite AMT workers who reside in US and are knowledgeable about the stock market to rate the news in multiple batches. Each piece of news is rated by at least 15 workers and an average of 17.8 workers. For quality control, we remove 10% of workers whose ratings deviate the most from the average ratings of other workers on the same news. After that, we are left with 15.8 workers for each piece of news on average. As to inter-rater reliability, the

news article is computed as the average sentiment ratings of all workers. Furthermore, we use the variance in the sentiment ratings of all workers to measure the uncertainty in the sentiment of the news article. Naturally, the larger the variance in the sentiment ratings among workers, the greater the uncertainty the news entails.

We do not consider the sentiment of tweets that is defined only when the number of tweets is nonzero. In our data, we do not observe any investor (firm) tweets on the event day for 19% (48%) of events. Removing these events not only substantially reduces our sample size but also undermines the representativeness of our sample.

Moderators and Control Variables: The moderator of interest in this study is firm size. We use a dummy variable *large_cap* to denote firm size, which is 1 for firms whose market capitalizations are over 10 billion US dollars on the prior trading day. Large firms have a more transparent information environment than small firms and the impact of news events on large firms is expected to be different from that on small firms. As to control variables, we consider a number of news- and firm-level characteristics. The news level control variables include *words* (the number of words in the leading paragraph[7]), *title_words* (the number of words in the title), and *news* (the number of news related to the focal firm on the event day). *words* and *title_words* capture the depth of news, while *news* represents the breadth of news coverage. In terms of firm-level characteristics, we consider the stock price prior to the event day (i.e., *price*), and the number of outstanding stock shares prior to the event day (i.e., *share_out*). We only consider time-variant firm-level characteristics because we use firm-level fixed effects to account for the effects of the observed and unobserved time-invariant characteristics of firms. In addition, we

Krippendorff's alpha for the labeled sentiments is 0.34, which is comparable with the values reported for subjective measures [31,40]. The Gwet's AC1, a metric that is penalized less by subjectivity [31], is rather high (0.84), increasing our confidence in the quality of the labeled sentiments.

[7] The news data we obtained from Factiva only includes the title and leading paragraph for each piece of news.

use month and weekday dummies to capture the seasonality in tweeting activities. Table 2 summarizes all the variables constructed at the event level for this study.

Table 2. Variable Description

Variable	Description
abnormal_firm_tweets	The abnormal volume of tweets by the focal firm on the event day, as defined by log(1+firm_tweets) – log(1+ firm_past_tweets).
abnormal_investor_tweets	The abnormal volume of tweets by investors on the event day, as defined by log(1+investor_tweets) – log(1+ investor_past_tweets).
abnormal_return	The excess stock return of the focal firm over the expected return on the event day (measured in percentage).
firm_past_tweets	The average number of tweets per day by the focal firm in the past two weeks.
firm_tweets	The number of tweets by the focal firm on the event day.
investor_past_tweets	The average number of tweets per day about the focal firm by investors in the past two weeks.
investor_tweets	The number of investor tweets about the focal firm on the event day.
large_cap	Whether the market capitalization of the focal firm on the prior trading day exceeds 10 billion US dollars.
month	The month of the event day (1-12).
news	The number of news about the focal firm on the event day.
price	The stock price of the focal firm on the prior trading day.
sentiment	The average sentiment of the news as rated by AMT workers.
sentiment_var	The variance in the sentiment rated by AMT workers.
share_out	The number of outstanding stock shares of the focal firm on the prior trading day (measured in thousand shares).
title_words	The number of words in the title of the news article on the focal firm on an event day.
weekday	The weekday of the event day (1-5).
words	The number of words in the news article on the focal firm on an event day.

Note: If the event day has multiple pieces of news, the news-level characteristics (i.e., words, title_words, sentiment, and sentiment_var) are averaged over all news. Our empirical results remain consistent if we focus on events with only one piece of news.

The descriptive statistics of the variables for the event study are presented in Table 3. There are various numbers of missing values for different variables due to three issues. First, 96 firms did not post any tweets two weeks before any of the events, resulting in 1,073 missing values for three variables related to firm tweets. Second, 23 firms were not mentioned by investors two weeks before or on the event day, which leads to 114 missing values for three variables related to investor tweets. Third, the daily stock price information is missing or incomplete for part of firms in our observation window, which results in 916 missing values on *abnormal_returns*, 541 missing values on price and *share_out*, and 492 missing values on *large_cap*.[8] The missingness in our data primarily results from the first and third issues. The first issue does not undermine the representativeness of our sample, since there is no point to consider firms that do not use Twitter when studying how firms respond to news on social media. The third issue is a limitation of the CRSP database. There are 15 large-cap firms in total, including 5 firms whose market capitalizations exceeded 10 billion USD only for part of the time in the year 2014. Large-cap firms account for 6% of events in our data. The correlations among the constructed variables are shown in Table 4.

Table 3. Descriptive Statistics

Variable	**Mean**	**SD**	**Min**	**Median**	**Max**	**#Missing Obs**	**#Obs**
abnormal_firm_tweets	0.15	0.49	-2.31	0.13	2.91	1073	3927
abnormal_investor_tweets	0.27	0.65	-2.31	0	3.25	114	3927
abnormal_return	0.15	3.90	-72.07	-0.01	50.51	916	3927
firm_past_tweets	2.95	4.05	0	1.43	41.43	1073	3927
firm_tweets	4.07	5.54	0	2	68	1073	3927
investor_past_tweets	2.12	5.36	0	0.21	55.50	114	3927
investor_tweets	3.89	7.48	0	1	64	114	3927
large_cap	0.06	0.24	0	0	1	492	3927
news	1.23	0.62	1	1	8	0	3927
price	33.93	52.80	0.30	19.48	589.47	541	3927
sentiment	0.39	0.47	-1.64	0.40	1.69	0	3927
sentiment_var	0.38	0.24	0	0.33	2.07	0	3927
share_out	154361	824068	309	38161	8300724	541	3927
title_words	10.85	4.09	1	10	39	0	3927
words	819.17	911.94	14	576	9605	0	3927

Note: The descriptive statistics for observations without missing values are reported in Online Appendix B.

[8] There are considerably more missing values for *abnormal_returns* because the training of the expected return model requires the stock price information of a firm to be observed in a long pre-event window. There are relatively fewer missing values for *large_cap* as this variable can be imputed based on partial stock price information.

Table 4. Correlation Matrix

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) abnormal_firm_tweets	1.00										
(2) abnormal_investor_tweets	0.03	1.00									
(3) abnormal_return	0.01	0.15	1.00								
(4) large_cap	-0.03	-0.06	-0.02	1.00							
(5) news	0.01	0.16	0.01	0.01	1.00						
(6) price	0.00	-0.09	-0.04	0.41	0.03	1.00					
(7) sentiment	0.14	0.00	0.09	-0.08	-0.08	-0.05	1.00				
(8) sentiment_var	-0.03	0.10	-0.01	0.17	0.08	0.09	-0.20	1.00			
(9) share_out	-0.03	-0.04	-0.01	0.48	-0.01	0.05	-0.07	0.11	1.00		
(10) title_words	0.08	-0.11	0.02	0.00	-0.13	-0.01	0.20	-0.05	-0.01	1.00	
(11) words	0.05	0.23	0.02	-0.05	0.01	-0.06	0.01	0.17	-0.03	-0.15	1.00

EMPIRICAL ANALYSIS

Empirical Model

We aim to estimate how investors, firms, and the stock market respond to the sentiment and uncertainty of news, as well as how the reactions of investors and firms (on social media) influence the stock market. However, there might be simultaneity among the reactions of firms, investors, and the stock market that are observed on the same day. For example, abnormal firm/investor tweets and abnormal returns may mutually affect each other.

Simultaneous Equations Model: Both the vector autoregression and simultaneous equations models are popular models in the presence of simultaneity. The simultaneous equations model, often estimated using the three-stage least squares (3SLS) method, addresses simultaneity with instrument variables that only affect a subset of dependent variables, while the vector autoregression model circumvents the endogeneity by estimating time-lagged influences on one another (i.e., using lagged variables as regressors). The time-lagged influences can only be interpreted as Granger causality [12,23], which is known to be different from true causality. Moreover, although using lagged variables can avoid simultaneity, these variables may well represent outdated information in that the stock market can absorb information in hours, minutes, or even seconds [1,26,33], which can lead to measurement errors, another type of endogeneity. As a remedy for the measurement error problem while using the vector autoregression model, we may consider constructing the variables at an hourly level instead of daily. Nevertheless, as shown in Table 1, on average, only 28.6 news articles on each firm are released in one year. The resulting data constructed for the vector autoregression model will be too sparse to be useful. Thus, we use the simultaneous equations model, which is a better tool to address our research questions and more appropriate for event studies. The simultaneous equations model is particularly suitable to establish causality in our context in that it allows outcome variables to be interdependent and has built-in support for instrument variables. Specifically, we use the following three-equation model on the constructed cross-sectional data at the news event level:

$$\begin{aligned}
ihs(abnormal_return_i) &= \alpha_0 + \alpha_1 abnormal_user_tweets_i + \alpha_2 abnormal_firm_tweets_i + \alpha_3 sentiment_i \\
&+ \alpha_4 sentiment_var_i + \alpha_5 \log(words_i) + \alpha_6 \log(title_words_i) + \alpha_7 \log(news_i) \\
&+ \alpha_8 \log(price_i) + \alpha_9 \log(share_out_i) + \theta_{1,f(i)} + \varepsilon_{1i}
\end{aligned}$$

$$\begin{aligned}
abnormal_user_tweets_i &= \beta_0 + \beta_1 ihs(abnormal_return_i) + \beta_2 abnormal_firm_tweets_i + \beta_3 sentiment_i \\
&+ \beta_4 sentiment_var_i + \beta_5 \log(words_i) + \beta_6 \log(title_words_i) + \beta_7 \log(news_i) \\
&+ \beta_8 \log(user_past_tweets_i) + \theta_{2,f(i)} + \delta_{2,m(i)} + \eta_{2,d(i)} + \varepsilon_{2i}
\end{aligned}$$

$$\begin{aligned}
abnormal_firm_tweets_i &= \gamma_0 + \gamma_1 ihs(abnormal_return_i) + \gamma_2 abnormal_user_tweets_i + \gamma_3 sentiment_i \\
&+ \gamma_4 sentiment_var_i + \gamma_5 \log(words_i) + \gamma_6 \log(title_words_i) + \gamma_7 \log(news_i) \\
&+ \gamma_8 \log(firm_past_tweets_i) + \theta_{3,f(i)} + \delta_{3,m(i)} + \eta_{3,d(i)} + \varepsilon_{3i}
\end{aligned}$$

where the subscripts $f(i)$, $m(i)$, and $d(i)$, respectively, indicate the firm, month, and weekday corresponding to the event i. Accordingly, $\theta_{j,f(i)}$, $\delta_{j,m(i)}$, and $\eta_{j,d(i)}$ respectively capture the fixed effects of firm, month, and weekday in the j-th equation. The error terms in the three equations (i.e., ε_{1i}, ε_{2i}, ε_{3i}) are allowed to be correlated with each other. We use the IHS transformed abnormal return because linear regression is highly sensitive to outliers.

In the above model, the abnormal return depends on the abnormal volumes of investor and firm tweets and vice versa. When the three dependent variables are used as independent variables, they are likely endogenous due to simultaneity. The identification of the simultaneous equations model requires each equation to have at least as many instrument variables as endogenous independent variables. In a simultaneous equations model, the instruments for the endogenous (independent) variables in a focal equation are exogenous independent variables that are excluded from the focal equation but included in the equations where the endogenous variables are used as dependent variables. For example, an instrument for the endogenous variable *abnormal_firm_tweets* in the first equation needs to be excluded from the first equation but included in the third equation.

Instruments: We consider the following three sets of instruments that appear in only one or two of the three simultaneous equations. (1) We use month and weekday fixed effects for the abnormal investor and firm tweets because tweeting activities often exhibit some sort of seasonality.[9] However, in an efficient stock market, it is highly unlikely that the abnormal return exhibits any seasonal patterns persistently, otherwise, investors would take advantage of the seasonal patterns to invest, which in turn washes away the pattern. This is also known as the self-destruction property of anomalies [25]. For example, the earlier discovered January effect [30]

[9] http://pages.trackmaven.com/rs/trackmaven/images/Retweet_Report_.pdf.

has disappeared in both US and international markets [14,27]. (2) Naturally, firms' past tweeting intensities are expected to affect their abnormal tweeting intensities because active and inactive firms may respond to news differently. However, there is no reason to believe a priori that investors and the stock market will respond to firms' average tweeting intensities in the past two weeks in an abnormal way [21]. Similarly, investors' past tweeting intensity on a firm is expected to influence their abnormal tweeting intensity on the event day, but it is unlikely to lead to abnormal stock movements or firm behavior. Given that investors' average tweeting intensity in the past two weeks is not new information to the market, the information reflected in this variable should have already been absorbed by the stock market over the past two weeks. Meanwhile, while it is possible that firms may adjust their tweeting intensities based on investors' tweeting intensities in the past two weeks, the adjustment is more likely to be made during the past two weeks, instead of abruptly on the event day. (3) Since the price and shares outstanding one day prior to the event are generally very close to their levels two or three days prior to the event, the price and shares outstanding prior to the event day are not new information at all and hence are very unlikely to lead firms or investors to tweet abnormally on the event day, especially after controlling for firm-level fixed effects.[10] While price and shares outstanding of the prior day are not novel information to the market, they can influence abnormal returns because they jointly determine the market capitalization of a firm. The stock volatility is generally larger for firms with smaller market capitalizations [10]. Consequently, the abnormal returns are typically larger for firms with smaller market capitalizations. In support of this, the

[10] We have also considered using the price and shares outstanding *one week* prior to the event day as instruments. These two variables represent information that has been available to the public for a long time (i.e., one week) and hence is highly unlikely to lead firms or investors tweet abnormally on the event day. The results using these two alternative instruments for the abnormal return are essentially the same. Please refer to Online Appendix D for more details.

mean absolute abnormal returns on the event day are 0.9% and 2.3% for large-cap and small-cap firms, respectively, in our dataset. Therefore, price and shares outstanding are expected to be predictive of abnormal returns. Given that the number of instruments for each equation exceeds the number of endogenous variables in every equation, the above model can be identified despite the endogeneity problem.

Main Results

We first report our main results using the simultaneous equations model, which allows us to test the relationships laid out in the conceptual framework (Figure 1). To operationalize this analysis, we focus on the subset of firms that were active on Twitter and whose abnormal returns are available. Table 5 summarizes the results. There are several interesting findings.

First, firms post more than usual when the sentiment of the news is positive (effect size of *sentiment* on abnormal firm tweets is 0.068, p-value = 0.014) and the uncertainty is low (effect size of *sentiment_var* on abnormal firm tweets is −0.086, p-value = 0.082), whereas investors post more than usual only when the uncertainty in news is high (effect size of *sentiment_var* on abnormal investor tweets is 0.252, p-value < 0.001). This finding reflects the different purposes of these two parties in using social media. Specifically, firms usually use social media for impression management [32], whereas investors often use social media for information exchange [12,28]. When the information in the news is positive and certain, the focal firm may take advantage of this opportunity to spread the good news; when the uncertainty of news is high, the focal firm tends to play safe and stay low-key until the consequence of the news becomes clear. On the other hand, when the uncertainty involved in news is high, investors are more incentivized to use social media and resolve the uncertain or ambiguous information through collective intelligence [8,29].

Table 5. The Responses of Investors, Firms, and Stock Market to News

VARIABLES	(1) ihs(abnormal_return)	(2) abnormal_investor_tweets	(3) abnormal_firm_tweets
ihs(abnormal_return)		-0.020	0.055
		(0.079)	(0.065)
abnormal_investor_tweets	0.180+		-0.010
	(0.096)		(0.053)
abnormal_firm_tweets	0.612*	0.138	
	(0.269)	(0.124)	
sentiment	0.172*	0.002	0.068*
	(0.080)	(0.034)	(0.028)
sentiment_var	0.123	0.252***	-0.086+
	(0.150)	(0.058)	(0.050)
log(words)	0.006	0.089***	0.034**
	(0.036)	(0.014)	(0.013)
log(title_words)	0.033	-0.222***	0.060+
	(0.094)	(0.036)	(0.032)
log(news)	0.061	0.261***	0.030
	(0.100)	(0.038)	(0.035)
log(price)	-0.920***		
	(0.173)		
log(share_out)	-0.459*		
	(0.200)		
log(investor_past_tweets)		-0.463***	
		(0.031)	
log(firm_past_tweets)			-0.351***
			(0.038)
Observations	2,110	2,110	2,110
R-squared	0.078	0.375	0.149
Firm FE	Yes	Yes	Yes
Month FE	No	Yes	Yes
Week FE	No	Yes	Yes

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.1. The model is estimated using 3SLS, which allows the error terms in the three equations to be correlated.

Second, both *abnormal_firm_tweets* and *abnormal_investor_tweets* have a positive effect on the abnormal return (though the latter is only marginally significant). One possible explanation is that the abnormal tweeting activities of firms and investors in the event window can increase the stock returns by reducing information asymmetry and associated risks [4]. In particular, the positive effect of *abnormal_investor_tweets* on stock return resonates with the prior finding that firms can alleviate the negative reaction of the stock market to product recalls by posting more than they normally would [21]. Our results demonstrate the value of social media in a broader context (i.e., the reaction of the stock market to financial news) and further show that social media is an effective communication channel for both firms and investors.

Third, there is no evidence that firms or investors react to the abnormal stock returns as the abnormal return has no significant effect on the abnormal tweeting intensities of either firms or investors. In addition, neither investors nor firms respond to the abnormal tweeting intensities of each other, which contradicts the prior finding that the volume of PGC on movies is predictive of the volume of UGC on movies [35]. The discrepancy likely results from the difference in contexts. For movies, the primary objective of PGC is to increase the volume of social media buzz so that more people are aware of the movies. Therefore, it is not surprising that the PGC volume is predictive of the UGC volume for movies. However, in our context, the main objective of posting PGC upon the release of financial news is to provide information to investors, rather than to stimulate social media buzz. Our finding suggests that the interplay between UGC and PGC may be context-dependent.

We summarize our key findings in Figure 3. Our findings add to the literature by showing that both firms and investors respond to the uncertainty in news and that the reactions of firms and investors on social media can influence the stock performance of firms. Of particular note is that firms and investors have opposite reactions to the uncertainty in news, which provides novel insights into the differential social media strategies and objectives of these two parties. Next, we further investigate whether our findings vary across large and small firms.

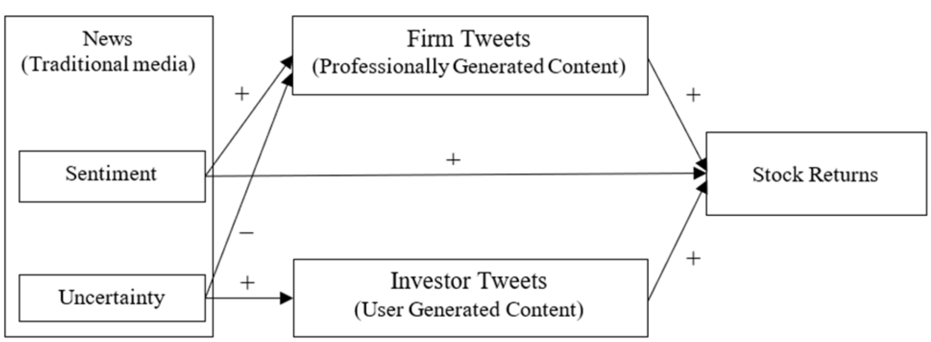

Figure 3. Main Results of Simultaneous Equation Modeling

Overall and Heterogeneous Effects

The simultaneous equations model gives us a full picture of the relationships among the two focal variables (sentiment and uncertainty) and the responses of investors, firms, and the stock market to financial news. However, the estimated effects of sentiment and uncertainty on each of the three parties do not represent their overall effects on the party, as the indirect effects through the other two parties are controlled for. From a practical point of view, the overall effects are also of great interest to market participants. In fact, our first research question is more concerned with the overall effects, and our third research question focuses on how the overall effects vary across large and small firms. To further answer our first and third research questions, we estimate the overall and heterogeneous effects of financial news' sentiment and uncertainty on abnormal investor tweets, abnormal firm tweets, and abnormal returns in turn. While analyzing the effects of sentiment and uncertainty on one dependent variable, we do not control for the other two dependent variables like we do in the simultaneous equations model, as otherwise the coefficients of sentiment and uncertainty cannot be interpreted as overall effects.

Abnormal Investor Tweets: We begin with studying how financial news' sentiment and uncertainty affect the abnormal tweeting activities of investors using data on all firms. The results on the overall and heterogeneous effects are summarized in Table 6. In all four columns, the effect of *sentiment* on abnormal investor tweets is positive (either significant or marginally significant), suggesting that investors tweet more in a positive news cycle than in a negative news cycle.[11] While prior studies have shown that the stock market is more responsive to negative news [11,12], our result suggests that investors are more responsive to positive news on

[11] This effect represents the overall effect of sentiment on abnormal investor tweets and is not at odds with the insignificant effect in Column 2 of Table 5, which only captures the direct effect of sentiment on abnormal investor tweets. The results in Table 5 do suggest that sentiment may indirectly increase abnormal investor tweets by increasing abnormal firm tweets.

social media, indicating that investors may respond to news differently in the stock market and on the social media. A potential explanation is that retail investors are more likely to chase stocks surrounded by good news on social media.

Table 6. Investor Response to News

	(1)	(2)	(3)	(4)
VARIABLES	abnormal_investor_tweets			
log(words)	0.077***	0.077***	0.076***	0.076***
	(0.013)	(0.013)	(0.013)	(0.013)
log(title_words)	-0.213***	-0.211***	-0.211***	-0.210***
	(0.032)	(0.032)	(0.032)	(0.032)
log(news)	0.307***	0.307***	0.306***	0.306***
	(0.034)	(0.034)	(0.034)	(0.034)
sentiment	0.055*	0.048+	0.051*	0.050+
	(0.024)	(0.027)	(0.024)	(0.027)
sentiment_var	0.196***	0.199***	0.224***	0.223***
	(0.048)	(0.048)	(0.050)	(0.050)
sentiment*large_cap		0.060		0.007
		(0.047)		(0.048)
sentiment_var*large_cap			-0.307*	-0.301*
			(0.125)	(0.119)
Observations	3,418	3,418	3,418	3,418
R-squared	0.299	0.299	0.300	0.300
Firm FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.1. Robust standard errors clustered by firms in parentheses.

In addition, the significantly positive main effect of *sentiment_var* suggests that investors are more active in tweeting about the focal firm when the uncertainty is high. Also, abnormal investor tweets increase with *log(news)* and *log(words)*, which respectively capture the breadth and depth of firm-related news. This is not surprising as publishers select and publish informative news that can capture public attention and investors often respond to such news articles on social media. Finally, in columns 3 and 4 of Table 6, the negative coefficient of *sentiment_var***large_cap* suggests that investors respond more actively to the uncertainty in the news for small firms than that for large firms. One rationale behind this finding is that it is often easy to obtain information on large firms, whereas less so for small firms, which makes social media an ideal platform to exchange related information. Consequently, investors are more likely to resort to social media when there is uncertainty about small firms.

Abnormal Firm Tweets: Now we examine how news affects abnormal firm tweets. As demonstrated in Table 7, the abnormal tweeting activities of firms increase with the sentiment of news but decrease with the uncertainty in news. This finding sheds light on the strategic behaviors of firms in using social media platforms for impression management [32]. When news is positive and certain, firms more actively interact with users on social media to further spread the good news; while when the news is negative or comes with high uncertainty, firms are less active on social media. This finding echoes the prior finding that firms tend to hold back the optional dissemination of earnings news on social media when the news is bad [19]. It is counterintuitive that firms do not respond to the news with high uncertainty or mixed sentiments, but it also makes sense that firms keep quiet to stay away from lawsuits or other legal matters when the consequence of the news is unclear.

Table 7. Firm Response to News

	(1)	(2)	(3)	(4)
VARIABLES		abnormal_firm_tweets		
log(words)	0.032**	0.032**	0.032**	0.032**
	(0.011)	(0.011)	(0.011)	(0.011)
log(title_words)	0.063+	0.064*	0.063+	0.064*
	(0.032)	(0.032)	(0.032)	(0.032)
log(news)	0.051+	0.052+	0.051+	0.052+
	(0.030)	(0.030)	(0.030)	(0.030)
sentiment	0.096***	0.092***	0.096***	0.092***
	(0.026)	(0.026)	(0.026)	(0.026)
sentiment_var	-0.091*	-0.090*	-0.091*	-0.092*
	(0.043)	(0.043)	(0.044)	(0.044)
sentiment*large_cap		0.074		0.084
		(0.112)		(0.108)
sentiment_var*large_cap			-0.001	0.066
			(0.165)	(0.119)
Observations	2,475	2,475	2,475	2,475
R-squared	0.120	0.120	0.120	0.121
Firm FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.1. Robust standard errors clustered by firms in parentheses.

Also, we do not find a significant difference between large and small firms in terms of how they respond to the sentiment and uncertainty of news. It is a bit surprising that small firms do not spread good news or disseminate firm information on social media platforms more often than big firms do. A possible explanation is that many small firms do not have a dedicated social media team due to limited resources. Considering the fact that small firms have less visibility than big firms, small firms are urged to improve their information environment by using social media platforms more actively, especially when there is good news about them. This finding suggests that there is still room for improvement in terms of firms' social media strategy.

Financial Performance: The overall and heterogeneous effects of the sentiment and uncertainty of news on the financial performance of firms are summarized in columns 1-3 of Table 8. As shown in column 1, the abnormal return increases with the sentiment of news but not the uncertainty of news, which makes sense because the uncertainty should not drive the abnormal return in a particular direction but the positive sentiment in news is expected to improve stock returns.

The coefficients of *sentiment* and *sentiment*large_cap* in column 2 suggest that the impact of sentiment is only significant for small firms (effect size = 0.240, p-value = 0.001), but not for large firms (effect size = $0.240 - 0.119 = 0.121$, p-value = 0.158). A potential explanation for this finding is that there is much more information available to the public for large firms than for small firms [4,43], which makes large firms less sensitive to individual news. Meanwhile, the insignificant effects of *sentiment_var* and *sentiment_var*large_cap* in column 3 are not surprising as the uncertainty in news is not expected to have a directional effect on abnormal returns, regardless of firm size.

Table 8. Impact of News and Social Media Reactions on Financial Performance

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	ihs(abnormal_return)					
log(words)	0.038	0.038	0.038	0.018	0.040	0.026
	(0.032)	(0.032)	(0.032)	(0.031)	(0.037)	(0.036)
log(title_words)	0.081	0.076	0.080	0.133+	0.027	0.084
	(0.080)	(0.080)	(0.080)	(0.078)	(0.094)	(0.093)
log(news)	0.098	0.098	0.098	0.014	0.133	0.066
	(0.095)	(0.095)	(0.095)	(0.096)	(0.115)	(0.117)
log(price)	-0.876***	-0.875***	-0.876***	-0.766***	-1.023***	-0.921***
	(0.177)	(0.177)	(0.177)	(0.173)	(0.181)	(0.177)
log(share_out)	-0.478	-0.478	-0.478	-0.488+	-0.525	-0.521
	(0.312)	(0.313)	(0.312)	(0.292)	(0.337)	(0.318)
sentiment	0.227***	0.240**	0.227***	0.204**	0.246**	0.227**
	(0.067)	(0.074)	(0.067)	(0.067)	(0.082)	(0.085)
sentiment_var	0.089	0.083	0.084	-0.009	0.148	0.040
	(0.116)	(0.117)	(0.124)	(0.114)	(0.163)	(0.162)
sentiment*large_cap		-0.119				0.031
		(0.114)				(0.159)
sentiment_var*large_cap			0.051			0.573
			(0.301)			(1.223)
abnormal_investor_tweets				0.295***		0.250***
				(0.050)		(0.059)
abnormal_investor_tweets*large_cap				-0.294*		-0.941***
				(0.123)		(0.245)
abnormal_firm_tweets					0.007	0.009
					(0.068)	(0.068)
abnormal_firm_tweets*large_cap					0.246	0.194
					(0.394)	(0.247)
Observations	3,007	3,007	3,007	2,994	2,121	2,110
R-squared	0.096	0.096	0.096	0.111	0.104	0.115
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.1. Robust standard errors clustered by firms in parentheses.

Columns 4-5 of Table 8 further investigate whether the social media reactions of investors and firms have differential impacts on the abnormal returns of large and small firms. The significantly negative coefficient of *abnormal_investor_tweets*large_cap* in column 4 demonstrates that the effect of abnormal investor tweets on abnormal stock return is stronger for small firms than for large firms. Moreover, the coefficients of *abnormal_investor_tweets* and *abnormal_investor_tweets*large_cap* combined show that the abnormal tweeting activities of investors have a significant effect on the financial performance of small firms (effect size =

0.295, p-value < 0.001), but not that of large firms (effect size = 0.295 – 0.294 = 0.001, p-value = 0.994). A likely explanation is that the tweets of investors can help resolve the uncertainty of news related to small firms, which are known to have a less transparent information environment than large firms. The coefficient of the interaction term in column 5 shows that the social media reactions of firms, in terms of abnormal tweets, do not have significantly different effects on the financial performances of firms, large or small. Finally, the results in column 6 show that the findings are similar when all interaction terms are included at the same time.

Robustness Checks

We conduct extensive robustness checks to verify the reliability of our findings. First, to address the concern that our proposed instruments for the simultaneous equations model can be weak or invalid, we evaluate the strength and validity of the instruments with the weak- and over-identification tests (Online Appendix C). Second, we use two alternative instruments for the abnormal return and find that the results of the simultaneous equations model remain consistent (Online Appendix D). Third, to further alleviate the concern regarding the potential interference among the adjacent news events on the same firms, we conduct an additional analysis in which we exclude all news events that occurred within seven days of the last news events on the same firms (Online Appendix E). Fourth, we show that the effects of abnormal investor (firm) tweets on the abnormal returns of large and small firms remain consistent after controlling for the total numbers of likes and retweets on the tweets posted by investors and firms on the event day, suggesting that our findings are not driven by potential cascading effects (Online Appendix F). Fifth, to increase the sample size, we remove the firm equation and estimate the remaining two equations simultaneously on all firms whose abnormal returns are available, including those that were inactive on Twitter (Online Appendix G). Sixth, to address the concern that investors and firms might be slow to react to the news or the news may have a sustaining effect on financial performance, we also consider the responses of the three parties in a two-day window (Online Appendix H). Finally, we conduct additional heterogeneity analyses in which we use market capitalization as a continuous measure of firm size (Online Appendix I). The results from these robustness checks are all consistent with our main findings.

MANAGERIAL IMPLICATIONS

Increasingly creating a supply of social media content benefits firms and enhances their financial performance. We find that firms respond to positive news on social media more actively than negative news. We also find that firms may improve their financial performance by posting more social media content than they normally would. Managers leverage social media platforms to enhance the financial performance of their firms. We demonstrate a clear benefit that motivates managers to actively engage in social media platforms and post social media content. Besides, we find that the impact of firm tweets on financial performance is significant within one day window but diminishes on the following day (see Appendix H). It suggests managers respond to news quickly on social media and seize on opportunities to further boost their firms' financial performance in a positive news cycle. When leveraging social media platforms for business, managers should act more strategically on the timing of their social media posts. Social media platforms are an effective tool to increase firm publicity and brand awareness in general, but firms can generate their social media content and disseminate firm information to further create additional values for them.

Managers may consider advancing their strategies and tactics on influencing people who have adopted social media platforms. We find that UGC contributed by investors is significantly and positively associated with the financial performance of the firm, regardless of whether the firm is active on social media. We also find that, when there is uncertainty or ambiguity in news about a firm, investors become more active in using social media to discuss the firm and exchange information, whereas firms tend to reduce their activities on social media. Managers may consider proactively engaging investors and interacting with them on social media platforms. Engaging investors on social media platforms can be beneficial to firms as it may further stimulate investor responses to the uncertain news on social media. The increase in investor responses provides additional information about the firms and can positively and significantly influence the financial performance of the firms.

Managers should bridge the gap between the social media activities of firms (i.e., PGC) with those of investors (i.e., UGC). We find that abnormal social media activities of investors on firms have no significant impact on the social media activities of firms and vice versa. Our findings show that the increase in firm tweets and investor tweets both affect the financial performance of firms, however, firm tweets and investor tweets do not interact with each other. It should not be the case. Managers should consider implementing strategies that align investors' social media activities and firms' social media activities. Managers may create additional informational value by bridging UGC and PGC, which may allow firms to influence investors and thus increase their financial performance.

The information environment of small firms is less transparent compared to large firms. Accordingly, we find that investors respond to uncertain news about small firms more actively than large firms. When there is uncertainty in news stories, it creates a discussion on social media platforms. This social media content contributed by investors helps alleviate the problem of information asymmetry for small firms. We also find that the increased tweeting activities of investors only increase the financial performance of small firms, but not that of large firms, which suggests that social media usage generates more information and provides more value to small firms than to large firms.

CONCLUDING REMARKS

In this paper, we document several intriguing causal relationships among traditional media (e.g., news), social media (e.g., UGC and PGC), and firm performance. Specifically, we find that both firms and investors react to financial news on social media and their reactions can influence the stock price movement, which deepens our understanding of the value of social media to the stock market as an interactive communication channel. As to how firms and investors respond to news on social media, we find that firms are more responsive to good news and news with low uncertainty, whereas investors are more responsive to uncertain news. This finding reflects the different purposes of these two parties in using social media, namely, impression management [32] and information exchange [12,28].

Furthermore, we find that the social media reactions of investors to news and their influence on stock returns depend on the size of the focal firm. On the one hand, investors are more responsive to uncertain news on small firms than on large firms, implying that investors may pay more attention to the uncertainty associated with firms that have a less transparent information environment. On the other hand, we find that the social media reactions of investors to news, in terms of abnormal volume of tweets, have a stronger effect on the financial performance of small firms than on that of large firms. This finding suggests that small firms with low visibility on traditional media may benefit asymmetrically more from social media. Surprisingly, while social media level the playing field for small firms, we do not find small firms to respond to news more actively than large firms do on social media. Our findings suggest that small firms should be more proactive in using social media to improve their information environment.

Overall, our findings suggest that social media are an effective tool that firms and investors can leverage to affect the stock movement. However, this important channel of information communication must be strategically managed, with an eye towards the nature of the news and the size of a firm.

Our paper has its limitations. First, our analysis focuses on data from only one social media channel (i.e., Twitter) and our findings may not extend to other social media platforms (e.g., Facebook and Reddit). However, given that Twitter is one of the most prominent social media channels closely followed by investors, we believe that our findings on this channel alone still have important implications. Second, since our sample is limited to firms in the technology sector, our findings may not generalize to firms in other sectors of the stock market. It might be fruitful to investigate whether and, if so, how the interplay among firms, investors, and the stock market varies across sectors. Third, we do not consider the sentiment of tweets because the sentiment of investor (firm) tweets is missing for 19% (48%) of observations. Future studies along this direction may focus on a subset of firms that are highly active on Twitter and yet also frequently tweeted by investors. Fourth, we assume that the sentiment and uncertainty of news are exogenous, which might not be true for all types of news. We alleviate this problem to some extent by focusing on isolated news that does not have any follow-ups or precursor events two days before and after. If investors can anticipate or predict the sentiment and uncertainly of events, their effects on stock returns are likely underestimated. Lastly, we do not consider whether investor and firm tweets are directly responding to specific news, which might be an interesting direction that warrants a separate study. However, given the large volumes of news and tweets, the feasibility of such a study hinges on the availability of a machine learning tool that can automatically recover the conversation structure through texts.

REFERENCE

1. Adams, G., McQueen, G.R., and Wood, R. The effects of inflation news on high frequency stock returns. *Journal of Business*, *77*, 3 (2004), 547–574.
2. Alanyali, M., Moat, H.S., and Preis, T. Quantifying the relationship between financial news and the stock market. *Scientific Reports*, *3*, (2013).
3. Bartov, E., Faurel, L., and Mohanram, P.S. Can Twitter help predict firm-level earnings and stock returns? *Accounting Review*, *93*, 3 (2018), 25–57.
4. Blankespoor, E., Miller, G.S., and White, H.D. The role of dissemination in market liquidity: Evidence from firms' use of Twitter. *Accounting Review*, *89*, 1 (2014), 79–112.
5. Boyd, J.H., Hu, J., and Jagannathan, R. The stock market's reaction to unemployment news: Why bad news Is usually good for stocks. *Journal of Finance*, *60*, 2 (2005), 649–672.
6. Burbidge, J.B., Magee, L., and Robb, A.L. Alternative transformations to handle extreme values of the dependent variable. *Journal of the American Statistical Association*, *83*, 401 (1988), 123–127.
7. Chen, H., De, P., Hu, Y., and Hwang, B.H. Wisdom of crowds: The value of stock opinions transmitted through social media. *Review of Financial Studies*, *27*, 5 (2014), 1367–1403.
8. Chen, J., Xu, H., and Whinston, A.B. Moderated online communities and quality of user-generated content. *Journal of Management Information Systems*, *28*, 2 (2011), 237–268.
9. Chen, Y., Ganesan, S., and Liu, Y. Does a firm's product-recall strategy affect its financial value? An examination of strategic alternatives during product-harm crises. *Journal of Marketing*, *73*, 6 (2009), 214–226.
10. Cheung, Y.-W. and Ng, L.K. Stock price dynamics and firm size: An empirical investigation. *The Journal of Finance*, *47*, 5 (1992), 1985–1997.
11. Conrad, J., Cornell, B., and Landsman, W.R. When is bad news really bad news? *Journal of Finance*, *57*, 6 (2002), 2507–2532.
12. Deng, S., Huang, Z., Sinha, A.P., and Zhao, H. The interaction between microblog sentiment and stock returns: An empirical examination. *MIS Quarterly*, *42*, 3 (2018), 895–918.
13. Fama, E.F. and French, K.R. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*, 1 (1993), 3–56.
14. Fountas, S. and Segredakis, K.N. Emerging stock markets return seasonalities: The January effect and the tax-loss selling hypothesis. *Applied Financial Economics*, *12*, 4 (2002), 291–299.
15. Friedline, T., Masa, R.D., and Chowa, G.A.N. Transforming wealth: Using the inverse hyperbolic sine (IHS) and splines to predict youth's math achievement. *Social Science*

Research, *49*, (2015), 264–287.

16. Goh, J.C. and Ederington, L.H. Is a bond rating downgrade bad news, good news, or no news for stockholders? *The Journal of Finance*, *48*, 5 (1993), 2001–2008.
17. Gordon, L.A., Loeb, M.P., and Sohail, T. Market value of voluntary disclosures concerning information security. *MIS Quarterly*, *34*, 3 (2010), 567–594.
18. Hu, Y., Xu, A., Hong, Y., Gal, D., Sinha, V., and Akkiraju, R. Generating business intelligence through social media analytics: Measuring brand personality with consumer-, employee-, and firm-generated content. *Journal of Management Information Systems*, *36*, 3 (2019), 893–930.
19. Jung, M.J., Naughton, J.P., Tahoun, A., and Wang, C. Do firms strategically disseminate? Evidence from corporate use of social media. *Accounting Review*, *93*, 4 (2018), 225–252.
20. Kahneman, D. and Tversky, A. Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 2 (1979), 263–291.
21. Lee, L.F., Hutton, A.P., and Shu, S. The role of social media in the capital market: Evidence from consumer product recalls. *Journal of Accounting Research*, *53*, 2 (2015), 367–404.
22. Loughran, T. and Mcdonald, B. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *Journal of Finance*, *66*, 1 (2011), 35–65.
23. Luo, X., Zhang, J., and Duan, W. Social media and firm equity value. *Information Systems Research*, *24*, 1 (2013), 146–163.
24. MacKinlay, A.C. Event studies in economics and finance. *Journal of Economic Literature*, *35*, 1 (1997), 13–39.
25. Marquering, W., Nisser, J., and Valla, T. Disappearing anomalies: a dynamic analysis of the persistence of anomalies. *Applied Financial Economics*, *16*, 4 (2006), 291–302.
26. Mcwilliams, A. and Siegel, D. Event studies in management research: Theoretical and empirical issues. *Academy of Management Journal*, *40*, 3 (1997), 626–657.
27. Mehdian, S. and Perry, M.J. Anomalies in US equity markets: A re-examination of the January effect. *Applied Financial Economics*, *12*, 2 (2002), 141–145.
28. Moqri, M., Mei, X., Qiu, L., and Bandyopadhyay, S. Effect of “following” on contributions to open source communities. *Journal of Management Information Systems*, *35*, 4 (2018), 1188–1217.
29. Oh, O., Agrawal, M., and Rao, H.R. Community intelligence and social media services: A rumor theoretic analysis of tweets during social crises. *MIS Quarterly*, *37*, 2 (2013), 407–426.
30. Rozeff, M.S. and Kinney, W.R. Capital market seasonality: The case of stock returns. *Journal of Financial Economics*, *3*, 4 (1976), 379–402.

31. Salminen, J.O., Al-Merekhi, H.A., Dey, P., and Jansen, B.J. Inter-rater agreement for social computing studies. In *2018 5th International Conference on Social Networks Analysis, Management and Security, SNAMS 2018*. 2018, pp. 80–87.

32. Schniederjans, D., Cao, E.S., and Schniederjans, M. Enhancing financial performance with social media: An impression management perspective. *Decision Support Systems*, *55*, 4 (2013), 911–918.

33. Scholtus, M., Van Dijk, D., and Frijns, B. Speed, algorithmic trading, and market quality around macroeconomic news announcements. *Journal of Banking and Finance*, *38*, 1 (2014), 89–105.

34. Shi, Z., Rui, H., and Whinston, A.B. Content sharing in a social broadcasting environment: Evidence from Twitter. *MIS Quarterly*, *38*, 1 (2014), 123–142.

35. Song, T., Huang, J., Tan, Y., and Yu, Y. Using user- and marketer-generated content for box office revenue prediction: Differences between microblogging and third-party platforms. *Information Systems Research*, *30*, 1 (2019), 191–203.

36. Stock, J.H. and Yogo, M. Testing for weak instruments in Linear IV regression. In *Identification and Inference for Econometric Models*. Cambridge University Press, 2005, pp. 80–108.

37. Suijs, J. Voluntary disclosure of information when firms are uncertain of investor response. *Journal of Accounting and Economics*, *43*, 2–3 (2007), 391–410.

38. Tetlock, P.C. Does public financial news resolve asymmetric information? *Review of Financial Studies*, *23*, 9 (2010), 3520–3557.

39. Tetlock, P.C., Saar-Tsechansky, M., and Macskassy, S. More than words: Quantifying language to measure firms' fundamentals. *Journal of Finance*, *63*, 3 (2008), 1437–1467.

40. Wong, K., Paritosh, P., and Aroyo, L. Cross-replication reliability - an empirical approach to interpreting inter-rater reliability. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*. 2021, pp. 7053–7065.

41. Xie, K. and Lee, Y.J. Social media and brand purchase: Quantifying the effects of exposures to earned and owned social media activities in a two-stage decision making model. *Journal of Management Information Systems*, *32*, 2 (2015), 204–238.

42. Yu, Y., Duan, W., and Cao, Q. The impact of social and conventional media on firm equity value: A sentiment analysis approach. *Decision Support Systems*, *55*, 4 (2013), 919–926.

43. Zhang, X.F. Information uncertainty and stock returns. *Journal of Finance*, *61*, 1 (2006), 105–137.

Online Appendices

ONLINE APPENDIX A. EXAMPLES OF SOCIAL MEDIA REACTIONS

Scenario 1: On 9/29/2014, a news article about Google was released on Dow Jones Institutional News. The news was positive and said, "*Google (GOOG) stock is still a buy argues RBC Capital Markets analyst …raised price target to $730 from $650*". On the same day, Google tweeted its partnership with the software company Adobe. The firm tweet was "*Adobe Photoshop for Chromebooks. It's happening*". Investors tweet about the partnership on that day, for instance, "*$GOOG $ADBE Google, Adobe announces Creative Cloud for Chromebooks*" or "*$ADBE $GOOG "@verge: Adobe is bringing Photoshop to Chromebooks*".

Google's stock price went up from $568 (open price) to $574 (close price) on that day. Its abnormal return on that day was 0.19% and was 0.96% on the following day.

Scenario 2: On 8/6/2014, a news article about a lawsuit against Yelp Inc was released on Reuters News. The news mentioned, "*Yelp Inc and its executives were accused in a lawsuit of selling more than $81 million in stock while deceiving shareholders about the quality of consumer reviews on its website*". The company did not respond to the news on Twitter. There was high uncertainty about the potential consequences of the lawsuit filing. The examples of investor tweets on that day were "*Yelp Receives Positive Earnings Outlook* Update *from MKM Partners $YELP*", "*$YELP : Robbins Geller Rudman & Dowd LLP : Files Class Action Suit against Yelp*", "*$YELP doubling their space(152k sq feet) at 11 Madison Avenue, paying $13MM for the first year*", or "*Surprise, Yelp Rings In Profits*".

Yelp's stock price did not move downwards on that day. The open price was $67.65 and the close price was $67.78. No significant change in stock price movements.

ONLINE APPENDIX B. DESCRIPTIVE STATISTICS FOR COMPLETE CASES

Table B1 reports the descriptive statistics for observations without missing values. Overall, they are very similar to the descriptive statistics for the full sample. Therefore, we do not expect missing values to be a major issue.

Table B1. Descriptive Statistics for Observations without Missing Values

Variable	**Mean**	**SD**	**Min**	**Median**	**Max**	**#Obs**
abnormal_firm_tweets	0.16	0.48	-2.01	0.13	2.46	2110
abnormal_investor_tweets	0.30	0.66	-2.31	0	3.25	2110
abnormal_return	0.18	3.57	-22.97	0	50.51	2110
firm_past_tweets	2.80	3.60	0	1.5	22.79	2110
firm_tweets	3.89	4.95	0	2	38	2110
investor_past_tweets	1.82	4.91	0	0.36	55.5	2110
investor_tweets	3.60	7.12	0	1	64	2110
large_cap	0.02	0.14	0	0	1	2110
news	1.23	0.60	1	1	7	2110
price	32.32	51.58	0.30	18.93	579.95	2110
sentiment	0.37	0.47	-1.64	0.38	1.69	2110
sentiment_var	0.37	0.23	0	0.32	1.81	2110
share_out	59310	69933	2049	38583	615232	2110
title_words	10.77	4.11	1	10	29	2110
words	826.12	909.58	14	588.75	9605	2110

ONLINE APPENDIX C. WEAK- AND OVER-IDENTIFICATION TESTS

The identification of the simultaneous equations model relies on the validity of the instruments. While we provide detailed justifications regarding the validity of the instruments, there are no readily available procedures to evaluate the strength and validity of instruments in the simultaneous equations framework. To conduct the standard weak- and over-identification tests for instruments, we run instrumental regression for the three equations in the simultaneous equations model separately. For each equation, we use the same set of instruments for the endogenous independent variables as those used in the simultaneous equations model. Table C1 summarizes the model estimates, as well as the results for weak- and over-identification tests. The three instrumental regressions are estimated using the limited information maximum likelihood method because it is robust to potential weak instruments. The results are similar using the two-stage least squares method.

The results from the instrumental regressions are fairly consistent with what we find from the simultaneous equations model. The only noticeable difference is that the coefficient of *sentiment* in the first model becomes statistically insignificant, but the *p*-value is only slightly above 10% (*p*-value = 0.104). This finding is not surprising as estimating the three equations separately leads to a loss of efficiency, which results in slightly larger standard errors than the simultaneous equations model in general. Importantly, all the weak identification statistics are larger than the critical values of the Stock-Yogo weak identification test [36], suggesting that all the instruments used in the simultaneous equations model are strong instruments. Moreover, the instruments pass the over-identification test in all three models (i.e., *p*-value > 0.05), increasing our confidence in the validity of the instruments.

Table C1. Weak- and Over-Identification Tests for Instruments

VARIABLES	(1) ihs(abnormal_return)	(2) abnormal_investor_tweets	(3) abnormal_firm_tweets
ihs(abnormal_return)		-0.023	0.061
		(0.079)	(0.092)
abnormal_investor_tweets	0.170		0.029
	(0.106)		(0.052)
abnormal_firm_tweets	0.784*	0.126	
	(0.354)	(0.111)	
sentiment	0.154	0.004	0.068+
	(0.094)	(0.035)	(0.035)
sentiment_var	0.139	0.250***	-0.094*
	(0.169)	(0.064)	(0.047)
log(words)	0.001	0.089***	0.030*
	(0.042)	(0.018)	(0.015)
log(title_words)	0.022	-0.221***	0.069*
	(0.099)	(0.036)	(0.033)
log(news)	0.055	0.262***	0.020
	(0.116)	(0.044)	(0.039)
log(price)	-0.917***		
	(0.187)		
log(share_out)	-0.497		
	(0.307)		
log(investor_past_tweets)		-0.458***	
		(0.047)	
log(firm_past_tweets)			-0.364***
			(0.058)
Observations	2,110	2,110	2,110
Firm FE	Yes	Yes	Yes
Month FE	No	Yes	Yes
Week FE	No	Yes	Yes
Overidentification Statistic	22.220	0.637	1.036
Overidentification Test p-value	0.102	0.425	0.309
Weak Identification Statistic	4.060	6.795	8.248
Weak ID Test Critical Values	3.55	5.44	5.44

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.1. Robust standard errors clustered by firms are reported in parentheses. The statistics used for over- and weak-identification tests are Hasen J statistic and Kleibergen-Paap rk Wald F statistic, respectively, both of which allow the errors to be clustered. In the overidentification test, a *p*-value higher than 0.05 indicates that the instruments pass the overidentification test. The Stock-Yogo weak identification test critical values for 10% maximal Wald test size distortions are reported in the bottom row. A statistic larger than the critical value indicates that the strength of the instruments is acceptable at 10% maximal Wald test size distortion.

ONLINE APPENDIX D. ALTERNATIVE INSTRUMENTS FOR ABNORMAL RETURN

In case one believes that the price and shares outstanding prior to the event day may somehow lead investors or firms to tweet abnormally, we further consider two alternative instruments, namely, the price and shares outstanding *one week* before the event day. Since the correlations of these two variables with their counterparts the day prior to the event day are greater than 0.99, they are still expected to be correlated with the abnormal stock return on the event day. However, given that Twitter is commonly used as a channel to exchange most recent information and that trading is highly time-sensitive, it is highly unlikely that firms or investors would tweet abnormally in response to the stock price and shares outstanding a week ago. Therefore, these two variables should be valid instruments. Table D1 summarizes the results of the simultaneous equations model when we use these two alternative instruments. The results are highly consistent with those in Table 5, which demonstrates the robustness of our findings.

Table D1. Results of Simultaneous Equations Model Using Alternative Instruments

VARIABLES	(1) ihs(abnormal_return)	(2) abnormal_investor_tweets	(3) abnormal_firm_tweets
ihs(abnormal_return)		0.059	0.033
		(0.087)	(0.072)
abnormal_investor_tweets	0.180+		-0.004
	(0.097)		(0.053)
abnormal_firm_tweets	0.645*	0.094	
	(0.271)	(0.122)	
sentiment	0.160*	-0.011	0.073**
	(0.080)	(0.035)	(0.028)
sentiment_var	0.135	0.238***	-0.085+
	(0.151)	(0.058)	(0.049)
log(words)	0.004	0.087***	0.035**
	(0.036)	(0.014)	(0.013)
log(title_words)	0.047	-0.223***	0.063*
	(0.094)	(0.035)	(0.031)
log(news)	0.054	0.253***	0.031
	(0.100)	(0.038)	(0.035)
log(price_prior_week)	-0.853***		
	(0.176)		
log(share_out_prior_week)	-0.357+		
	(0.198)		
log(investor_past_tweets)		-0.454***	
		(0.030)	
log(firm_past_tweets)			-0.355***
			(0.037)
Observations	2,110	2,110	2,110
R-squared	0.072	0.391	0.166
Firm FE	Yes	Yes	Yes
Month FE	No	Yes	Yes
Week FE	No	Yes	Yes

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.1. The model is estimated using 3SLS, which allows the error terms in the three equations to be correlated.

ONLINE APPENDIX E. INTERFERENCE BETWEEN ADJACENT EVENTS ON THE SAME FIRM

To avoid potential interference among the news events on the same firm, we focus on isolated events with no other events two days before or after in our main analysis. As a robustness check, we further exclude all news events that occurred within 7 days of the last news events for the same firms. The estimates of the simultaneous equations model on the retained events are reported in Table E1. The overall findings are similar to those in Table 5, suggesting that our findings are unlikely driven by interference between adjacent news events.

Table E1. The Responses of Investors, Firms, and Stock Market to News (Excluding Firm Events Occurred within 7 Days of the Last Events)

VARIABLES	(1) ihs(abnormal_return)	(2) abnormal_investor_tweets	(3) abnormal_firm_tweets
ihs(abnormal_return)		-0.076	0.074
		(0.111)	(0.090)
abnormal_investor_tweets	0.054		-0.047
	(0.110)		(0.063)
abnormal_firm_tweets	0.954**	0.163	
	(0.333)	(0.162)	
sentiment	0.199*	0.004	0.077*
	(0.098)	(0.047)	(0.039)
sentiment_var	0.313+	0.278***	-0.074
	(0.178)	(0.075)	(0.063)
log(words)	0.001	0.087***	0.031*
	(0.042)	(0.017)	(0.015)
log(title_words)	-0.123	-0.215***	0.088*
	(0.114)	(0.046)	(0.038)
log(news)	0.040	0.283***	0.060
	(0.121)	(0.047)	(0.044)
log(price)	-0.757***		
	(0.204)		
log(share_out)	-0.255		
	(0.221)		
log(investor_past_tweets)		-0.464***	
		(0.037)	
log(firm_past_tweets)			-0.283***
			(0.042)
Observations	1,566	1,566	1,566
R-squared	0.057	0.354	0.169
Firm FE	Yes	Yes	Yes
Month FE	No	Yes	Yes
Week FE	No	Yes	Yes

Notes: *** *p*<0.001, ** *p*<0.01, * *p*<0.05, + *p*<0.1. The model is estimated using 3SLS, which allows the error terms in the three equations to be correlated.

ONLINE APPENDIX F. IMPACT OF SOCIAL MEDIA REACTIONS ON ABNORMAL RETURNS AFTER CONTROLLING FOR SOCIAL MEDIA ENGAGEMENTS

To address the concern that some tweets posted by investors or firms may go viral and hence have cascading effects, we conduct a robustness check in which we control for the total numbers of likes and retweets on all the tweets posted by investors and firms on the event day. Because the total numbers of likes and retweets include second-order and higher-order likes and retweets resulting from retweet cascades, they can capture the potential cascading effects of firms' and investors' tweets beyond the event day. Table F1 summarizes the regression results regarding how the numbers of abnormal investor and firm tweets affect the abnormal returns of large and small firms, with and without controlling for the total numbers of likes and retweets on the tweets. The effects of abnormal investor and firm tweets on abnormal returns of large and small firms are highly similar with and without controlling for the engagements on tweets, suggesting that our findings are robust to potential cascading effects.

Table F1. Results after Controlling for the Engagements on Tweets

	ihs(abnormal_return)			
VARIABLES	(1)	(2)	(3)	(4)
log(words)	0.018	0.020	0.040	0.041
	(0.031)	(0.031)	(0.037)	(0.037)
log(title_words)	0.133+	0.134+	0.027	0.030
	(0.078)	(0.078)	(0.094)	(0.094)
log(news)	0.014	0.012	0.133	0.136
	(0.096)	(0.096)	(0.115)	(0.115)
log(price)	-0.766***	-0.761***	-1.023***	-1.027***
	(0.173)	(0.181)	(0.181)	(0.181)
log(share_out)	-0.488+	-0.501+	-0.525	-0.530
	(0.292)	(0.296)	(0.337)	(0.338)
sentiment	0.204**	0.203**	0.246**	0.247**
	(0.067)	(0.067)	(0.082)	(0.083)
sentiment_var	-0.009	-0.010	0.148	0.148
	(0.114)	(0.114)	(0.163)	(0.163)
abnormal_investor_tweets	0.295***	0.276***		
	(0.050)	(0.051)		
abnormal_investor_tweets*large_cap	-0.294*	-0.293*		
	(0.123)	(0.132)		
log(investor_tweets_likes)		0.090		
		(0.060)		
log(investor_tweets_retweets)		-0.071		
		(0.055)		
abnormal_firm_tweets			0.007	0.022
			(0.068)	(0.075)
abnormal_firm_tweets*large_cap			0.246	0.288
			(0.394)	(0.431)
log(firm_tweets_likes)				0.008
				(0.067)
log(firm_tweets_retweets)				-0.026
				(0.059)
Observations	2,994	2,994	2,121	2,121
R-squared	0.111	0.111	0.104	0.104
Firm FE	Yes	Yes	Yes	Yes

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.1. Robust standard errors clustered by firms in parentheses.

ONLINE APPENDIX G. SIMULTANEOUS EQUATIONS MODEL ON ALL FIRMS

As mentioned earlier, one limitation with the three-equation simultaneous equations model is that the sample is restricted to the subset of firms who were active on Twitter as otherwise the variable *abnormal_firm_tweets* is undefined. Consequently, 33% of firms and 29% of events are removed. To increase the sample size, we remove the firm equation from the simultaneous equations model and estimate the remaining two equations simultaneously on all firms whose abnormal returns are available, including those who were inactive on Twitter. The results of the model are reported in columns (1a) and (1b) of Table G1. The substantive findings from this model are highly consistent with those in Table 5, which not only demonstrates the robustness of our findings but also suggests that the impact of investors' abnormal social media activity on stock returns is not limited to firms active on social media.

To investigate whether our earlier finding in Table 8 that abnormal investor tweets have a weaker effect on the abnormal stock returns of large-cap firms persists after accounting for the endogeneity of abnormal investor tweets, we also estimate a second model that considers the interactions with firm size. The results are reported in columns 2a and 2b. The significantly negative coefficient of *abnormal_investor_tweets *large_cap* in column 2a suggests that this finding is robust to the endogeneity of abnormal investor tweets. Note, the effect of abnormal investor tweets on the abnormal return of large firms remains insignificant (effect size = 0.242 − 0.581 = −0.339, p-value = 0.231).

Table G1. Simultaneous Equations Model on All Firms

VARIABLES	(1a) ihs(abnormal_return)	(1b) abnormal_investor_tweets	(2a) ihs(abnormal_return)	(2b) abnormal_investor_tweets
ihs(abnormal_return)		-0.027		-0.061
		(0.073)		(0.074)
abnormal_investor_tweets	0.227**		0.242**	
	(0.080)		(0.082)	
abnormal_investor_tweets *large_cap			-0.581*	
			(0.244)	
sentiment	0.206***	0.044	0.211***	0.051+
	(0.060)	(0.028)	(0.060)	(0.028)
sentiment_var	0.013	0.223***	0.003	0.225***
	(0.116)	(0.045)	(0.116)	(0.046)
log(words)	0.023	0.088***	0.023	0.089***
	(0.029)	(0.012)	(0.029)	(0.012)
log(title_words)	0.123+	-0.181***	0.121	-0.179***
	(0.074)	(0.029)	(0.074)	(0.030)
log(news)	0.032	0.291***	0.030	0.294***
	(0.081)	(0.031)	(0.081)	(0.032)
log(price)	-0.789***		-0.774***	
	(0.138)		(0.138)	
log(share_out)	-0.504*		-0.512**	
	(0.198)		(0.196)	
log(investor_past_tweets)		-0.398***		-0.403***
		(0.024)		(0.024)
Observations	2,998	2,998	2,998	2,998
R-squared	0.110	0.369	0.110	0.344
Firm FE	Yes	Yes	Yes	Yes
Month FE	No	Yes	No	Yes
Week FE	No	Yes	No	Yes

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.1. Columns (1a) and (1b) report the results of the basic simultaneous equations model. Columns (2a) and (2b) report the results of the simultaneous equations model that considers the interactions with firm size.

ONLINE APPENDIX H. RESPONSES IN A TWO-DAY TIME WINDOW

In the main analysis, we focus on the responses of investors, firms, and the stock market to the news on the event day, similar to Chen et al. [9]. For abnormal investor tweets and abnormal returns, this decision often makes sense as investors are known to react quickly to the news on social media and the stock market is also known to be rather efficient to absorb the signals in news. Firms, on the other hand, may be relatively slower in responding to news depending on the nature of the news. In occasional cases, the effect of the news on abnormal investor tweets and abnormal returns may last a bit longer. Therefore, as a robustness check, we also consider the responses of the three parties in a two-day window, namely, the event day and the day immediately after (a business day for abnormal return). For abnormal returns, we use the cumulative abnormal return over these two days. For abnormal investor (firm) tweets, we use the deviation of the average numbers of investor (firm) tweets from the average number of investor (firm) tweets in the past two weeks, again on the log scale. The results of the corresponding simultaneous equations model are summarized in Table H1.

The results using a two-day window are highly consistent with those using a one-day window, which demonstrates the robustness of our findings. The only noticeable difference is that the number of abnormal firm tweets has no effect on abnormal returns anymore. The effect of abnormal firm tweets on stock returns is short-lived (i.e., significant in a one-day window, but not in a two-day window) likely due to the high visibility of firm tweets to the public, as compared to tweets from investors.

Table H1. Simultaneous Equation Model Using a Two-Day Time Window

VARIABLES	(1) ihs(abnormal_return)	(2) abnormal_investor_tweets	(3) abnormal_firm_tweets
ihs(abnormal_return)		0.011	0.006
		(0.033)	(0.025)
abnormal_investor_tweets	0.240+		0.043
	(0.140)		(0.045)
abnormal_firm_tweets	-0.163	0.075	
	(0.255)	(0.104)	
sentiment	0.462***	-0.010	0.055*
	(0.095)	(0.030)	(0.022)
sentiment_var	0.198	0.209***	-0.063
	(0.183)	(0.051)	(0.040)
log(words)	-0.026	0.100***	0.025*
	(0.044)	(0.012)	(0.010)
log(title_words)	0.176	-0.214***	0.107***
	(0.118)	(0.033)	(0.026)
log(news)	0.023	0.214***	0.019
	(0.122)	(0.033)	(0.027)
log(price)	-1.875***		
	(0.213)		
log(share_out)	-0.839**		
	(0.260)		
log(investor_past_tweets)		-0.455***	
		(0.027)	
log(firm_past_tweets)			-0.385***
			(0.031)
Observations	2,110	2,110	2,110
R-squared	0.139	0.388	0.251
Firm FE	Yes	Yes	Yes
Month FE	No	Yes	Yes
Week FE	No	Yes	Yes

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.1. The model is estimated using 3SLS, which allows the error terms in the three equations to be correlated.

ONLINE APPENDIX I. CONTINUOUS MEASURE OF FIRM SIZE

In addition to the binary measure of firm size, we have also considered using the market capitalization of a firm on the prior trading day as a continuous measure of its size. Tables I1-I3 reports the results of the heterogeneity analyses based on this continuous measure. The results are highly consistent with those reported in Tables 6-8.

Table I1. Investor Response to News

	(1)	(2)	(3)	(4)
VARIABLES	abnormal_investor_tweets			
log(words)	0.080***	0.080***	0.079***	0.079***
	(0.013)	(0.013)	(0.013)	(0.013)
log(title_words)	-0.211***	-0.214***	-0.207***	-0.211***
	(0.032)	(0.031)	(0.031)	(0.031)
log(news)	0.289***	0.289***	0.290***	0.290***
	(0.034)	(0.034)	(0.034)	(0.034)
log(market_cap)	-0.081	-0.082	-0.077	-0.078
	(0.061)	(0.061)	(0.060)	(0.060)
sentiment	0.051*	0.054*	0.048+	0.052*
	(0.025)	(0.026)	(0.025)	(0.026)
sentiment_var	0.206***	0.204***	0.210***	0.207***
	(0.049)	(0.048)	(0.048)	(0.048)
sentiment*log(market_cap)		-0.009		-0.015
		(0.010)		(0.010)
sentiment_var*log(market_cap)			-0.042*	-0.050**
			(0.018)	(0.018)
Observations	3,368	3,368	3,368	3,368
R-squared	0.296	0.296	0.296	0.297
Firm FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.1. Robust standard errors clustered by firms in parentheses. The moderator ***log(market_cap)*** is demeaned.

Table I2. Firm Response to News

	(1)	(2)	(3)	(4)
VARIABLES		abnormal_firm_tweets		
log(words)	0.030**	0.031**	0.031**	0.031**
	(0.011)	(0.011)	(0.011)	(0.011)
log(title_words)	0.057+	0.056+	0.056+	0.056+
	(0.032)	(0.032)	(0.032)	(0.032)
log(news)	0.039	0.039	0.039	0.039
	(0.030)	(0.030)	(0.030)	(0.030)
log(market_cap)	0.032	0.033	0.031	0.031
	(0.063)	(0.063)	(0.063)	(0.063)
sentiment	0.093***	0.094***	0.094***	0.094***
	(0.026)	(0.026)	(0.026)	(0.026)
sentiment_var	-0.082+	-0.082+	-0.083+	-0.083+
	(0.044)	(0.044)	(0.044)	(0.044)
sentiment*log(market_cap)		-0.005		-0.003
		(0.015)		(0.015)
sentiment_var*log(market_cap)			0.014	0.013
			(0.024)	(0.023)
Observations	2,438	2,438	2,438	2,438
R-squared	0.122	0.122	0.122	0.122
Firm FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.1. Robust standard errors clustered by firms in parentheses. The moderator *log(market_cap)* is demeaned.

Table I3. Impact of News and Social Media Reactions on Financial Performance

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES			ihs(abnormal_return)			
log(words)	0.038	0.038	0.038	0.015	0.041	0.023
	(0.032)	(0.032)	(0.032)	(0.030)	(0.037)	(0.036)
log(title_words)	0.081	0.079	0.081	0.134+	0.029	0.084
	(0.080)	(0.081)	(0.080)	(0.078)	(0.094)	(0.092)
log(news)	0.098	0.098	0.098	0.019	0.133	0.071
	(0.095)	(0.095)	(0.095)	(0.096)	(0.115)	(0.117)
log(price)	-0.876***	-0.877***	-0.876***	-0.752***	-1.018***	-0.911***
	(0.177)	(0.176)	(0.177)	(0.175)	(0.180)	(0.175)
log(share_out)	-0.478	-0.478	-0.478	-0.485+	-0.524	-0.522+
	(0.312)	(0.313)	(0.313)	(0.285)	(0.336)	(0.312)
sentiment	0.227***	0.229**	0.227***	0.200**	0.247**	0.225**
	(0.067)	(0.072)	(0.067)	(0.067)	(0.082)	(0.081)
sentiment_var	0.089	0.087	0.089	-0.011	0.147	0.057
	(0.116)	(0.117)	(0.117)	(0.114)	(0.163)	(0.161)
sentiment*log(market_cap)		-0.007				0.031
		(0.029)				(0.052)
sentiment_var*log(market_cap)			0.000			0.065
			(0.053)			(0.089)
abnormal_investor_tweets				0.274***		0.234***
				(0.047)		(0.058)
abnormal_investor_tweets*log(market_cap)				-0.051*		-0.046
				(0.026)		(0.034)
abnormal_firm_tweets					0.010	0.012
					(0.067)	(0.067)
abnormal_firm_tweets*log(market_cap)					-0.016	-0.014
					(0.040)	(0.040)
Observations	3,007	3,007	3,007	2,994	2,121	2,110
R-squared	0.096	0.096	0.096	0.112	0.104	0.116
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: *** p<0.001, ** p<0.01, * p<0.05, + p<0.1. Robust standard errors clustered by firms in parentheses. The moderator *log(market_cap)* is demeaned. Its main effect is not included in this table as we have already controlled for *log(price)* and *log(share_out)*.

REFERENCES

1. Chen, Y., Ganesan, S., and Liu, Y. Does a firm's product-recall strategy affect its financial value? An examination of strategic alternatives during product-harm crises. *Journal of Marketing*, *73*, 6 (2009), 214–226.

2. Stock, J.H. and Yogo, M. Testing for weak instruments in Linear IV regression. In *Identification and Inference for Econometric Models*. Cambridge University Press, 2005, pp. 80–108.