Social Media and Financial News Manipulation

Shimon Kogan*
The Wharton School, University of Pennsylvania
Interdisciplinary Center Herzliya

Tobias J. Moskowitz
Yale School of Management
NBER
AQR Capital Management

Marina Niessner
The Wharton School, University of Pennsylvania

March 2022

*We thank Tony Cookson, Kent Daniel, Anastassia Fedyk, Diego Garcia, Matt Gentzkow, Gary Gorton, Bryan Kelly, Elizabeth Kempf, Bonnie Moskowitz, James Pennebacker, Jesse Shapiro, Kelly Shue, Eric So, Denis Sosyura, Sam Hartzmark, as well as conference and seminar participants at UCLA (Anderson), Rice University (Jones), University of Miami Business School, ASU Sonoran Winter Finance Conference, 3rd Annual News & Finance Conference, University of Colorado at Boulder, Northwestern University (Kellogg), FSU SunTrust Beach Conference, MIT Sloan, Yale SOM, Catolica-Lisbon, University of Kentucky Finance Conference, FEB, 3rd Rome Junior Finance Conference, the 2018 WFA meetings, the U.S. Securities and Exchange Commission Division of Economic and Risk Analysis, the Red Rock Conference, Behavioral NBER, Jackson Hole Finance Conference, AQR Capital Management, the NBER Asset Pricing meetings, 2020 AFA meetings in San Diego, and the Global Virtual Seminar Series on Fintech for their helpful comments and suggestions. We also thank Elli Hoffmann and Keren Ben Zvi for providing and helping organize the data. AQR Capital Management is a global investment management firm, which may or may not apply similar investment techniques or methods of analysis as described herein. The views expressed here are those of the authors and not necessarily those of AQR. Contact emails: skogan@upenn.edu, tobias.moskowitz@yale.edu, and marina.niessner@gmail.com. A previous version of this paper was titled, "Fake News in Financial Markets."

Abstract

We examine an undercover SEC investigation into the manipulation of financial news on social media. While fraudulent news had a direct positive impact on retail trading and prices, revelation of the fraud by the SEC announcement resulted in significantly lower retail trading volume on all news, including legitimate news, on these platforms. For small firms, volume declined by 23.5% and price volatility dropped by 1.3%. We find evidence consistent with concerns of fraud causing the decline in trading activity and price volatility, which we interpret through the lens of social capital, and attempt to rule out alternative explanations. The results highlight the indirect consequences of fraud and its spillover effects that reduce the social network's impact on information dissemination, especially for small, opaque firms.

We study a unique setting involving fraudulent news in financial markets that culminates in a Securities and Exchange Commission (SEC) investigation of shared financial news networks. We use this setting to document the effects of market manipulation on consumers and producers of information. The SEC investigation revealed the presence of fraud to market participants, which we show led to a 4.1% decrease in retail trading volume in response to all news, including legitimate news, from the social media platforms. The effects are even stronger for small, opaque firms who experience a 23.5% decline in trading volume and a 1.3% reduction in price volatility, relative to a baseline volatility of 7.2%, following the SEC announcement. The results highlight the indirect consequences of fraud and its spillover effects that reduce the social network's impact on information dissemination. We interpret these results through the lens of models of trust and the importance of social capital for financial activity, and attempt to address alternative explanations.

We examine social media platforms for financial news, which are crowd-sourced providers of information that democratize news. These networks grew in popularity, influence, and impact over the past decade which, coupled with lax regulatory monitoring, created incentives for manipulation. In 2014, an industry whistle blower, Rick Pearson, who was an author on one of the most prominent crowd-sourced platforms, Seeking Alpha, was approached by a public relations firm to write paid-for false content to promote a stock. Instead of rejecting the offer, he went undercover and turned over to the SEC 171 articles written by 20 authors about 47 companies who knowingly wrote false information about these firms. The SEC launched an investigation that became public in March 2014 and ultimately led to legal action against these firms and the articles' authors. We use the SEC investigation as our empirical laboratory to assess the indirect consequences of fraud. In fact, the 1934 Act that formed the SEC was originally motivated by a desire to restore the public's faith in capital markets following the 1929 stock market crash, with one of its mandates to monitor and investigate price manipulation.

We first confirm that social media news platforms have impact on markets. Articles written on the social platforms are associated with higher subsequent retail trading activity and price volatility. We focus on retail trading volume, measured by Boehmer et al. (2020), since retail investors are the primary participants on these social networks. Later, we also examine institutional investor trading volume as a placebo test. Obtaining the 171 fraudulent articles identified by Rick Pearson and the SEC, we show that these false promotional articles increased abnormal retail trading volume by more than 55% over the three days following the article's publication relative to legitimate articles written by the same author. These results are concentrated among small firms. In addition, the fraudulent media campaign was effective at manipulating prices, causing an average 13% rise in affected firm share prices immediately following the articles, that eventually reverses in the long run and ultimately becomes

cumulatively negative at −5%. These results are consistent with investors being deceived by the fraud that eventually gets corrected in markets. The findings provide evidence for the premise behind a theory of trust: investors manipulated by false news requires a level of trust in the medium in the first place.

Although the social media platforms do not control the content of the articles or fact check them, the idea behind the platforms is to democratize news and allow for disagreement. The SEC investigation, however, revealed that some of these articles were company-sponsored propaganda masquerading as genuine author opinion. It is this fraud that we view as a shock to the intent and operation of these platforms that investors were previously unaware. Given that investors responded to these platforms (via trading), indicating some trust, we hypothesize that this shock may have weakened that trust.

We interpret the SEC investigation and revelation of the social media manipulation as a shock to trust in crowd-sourced platforms. We use the announced SEC investigation and media coverage of the scandal in February and March 2014 as an event revealing fraud on these platforms, and measure fraud's broader, indirect effects on markets. Specifically, we collect 203,545 articles from Seeking Alpha from 2005 to 2015 and a competitor platform, Motley Fool (147,916 articles from 2009 to 2014), that cover 7,700 publicly traded firms. Uniquely, we test an implication from theory that fraud imposes externalities on other, legitimate news (Aymanns et al. (2017) and Allcott and Gentzkow (2017)) and relate this test more generally to trust in markets (Guiso et al. (2004)). We find that abnormal retail trading volume and price volatility drop significantly for *any* news article written on these platforms after the announcement, particularly for small firms with high retail interest. Retail abnormal trading volume drops by 4.1% for an article published in the six months after the SEC announcement versus the six months before the announcement. For the smallest firms, the drop is 23.5%. Price volatility for small firms also drops by 1.3%, with no significant price response for larger firms, consistent with externalities of manipulation on information networks. The effect of the SEC announcement is immediate and transitory. We document a reduction in retail investors' response to all news as quickly as one week into the event, but the effect dissipates over time, becoming insignificant a year after the event.

Identifying the impact of trust on markets, versus other possible channels impacting markets, is empirically challenging. There are many factors driving news and trading activity simultaneously that may coincide with the SEC event. For example, trends in news, changes in the market's response to news, or unobservables affecting trading may coincide with the SEC event. However, this identification challenge is largely mitigated by the SEC investigation being an unanticipated event, having been spawned by Rick Pearson's self-initiated undercover investigation. This fact makes the event more likely to be exogenous to the outcome variables, similar in spirit to a random shock to the information environment.

To help rule out other factors affecting markets at this time, we also show that there are no pre-trends in news or in the market's response to news prior to the announcement, which is consistent with the unanticipated and exogenous nature of the SEC event. In addition, we also employ a number of controls for other sources of news beyond the social platforms (corporate filings, press releases, and Wall Street Journal and New York Times articles) to help capture the market's response to news more generally. Finally, we run several "placebo" or falsification tests designed to pick up general effects associated with news and the market's response to news, but that are otherwise unrelated to these platforms or the SEC's investigation. For example, we examine institutional trading activity in response to articles published pre- versus post-SEC event as a placebo test. While we find a significant decline in *retail* trading volume for articles published after the SEC event, there is no significant change to *institutional* trading volume to articles published after the event. Since retail and institutional investors should both be affected by general trends in news or responses to news, or other omitted variables driving news and trading, but retail investors should mainly be affected by shocks to the social network platforms, these findings suggest that the effect of the scandal, and not other omitted factors, drives the precipitous drop in retail trading activity.

As another placebo test, we also look at news media unlikely to be affected by the SEC investigation. For instance, newspaper articles from the Wall Street Journal (WSJ) and New York Times (NYT) are unlikely to be affected by the fraud on social media platforms – WSJ or NYT journalists were not being paid to write false promotional articles about firms. However, general trends in news and the market's response to news should be evident for these media as well, and so we should see a decline in retail trading to WSJ and NYT articles, too, if general market conditions are driving these effects. We do not find a commensurate decline in trading volume after the SEC event in response to NYT or WSJ articles. This result is consistent with the SEC event only affecting responses to the social media platforms and not being a symptom of wider trends in news. Conducting similar falsification exercises for corporate filings and press releases – sources of news that are also highly unlikely to be polluted by fraud or be interpreted differently after the SEC investigation – we find no significant changes in market response either. These results suggest that other news trends or unobservables are not likely confounding the event study. The result is consistent with investors' increased distrust of news from the social media platforms, but not other forms of media.

We also show that the reduction in response to news is tempered for authors with better reputations and larger followings, consistent with consumers' trust motives. In addition, articles that contain more numbers, such as articles written about earnings, receive a more muted trading response reduction from investors after the shock. Conversely, measures of disagreement, such as the number and length of

comments on an article, and mentions of the word "disagree" in the comments section have no *differential* impact on trading volume after the SEC announcement, despite having an unconditional positive impact on trading. We also find that when comments contain the words "fake" or "fraud," the impact on volume declines even further, but *only* after the SEC announcement. Before the event, claims of "fraud" or "fake" in the comments have no impact. Finally, author disclosure of a position they have in the stock that they are writing about has a negative effect on volume after the scandal, but no effect prior to the scandal. These results are consistent with increased mistrust in news after the scandal and not changes in disagreement or other drivers of trading activity.

We argue that the spillover effects of market manipulation result in a new market equilibrium, where investors discount all news, including legitimate news, from the social platforms after the scandal is public. Applying natural language processing (NLP) to the comments section of the articles, we find a significant increase in use of the words "fake" and "fraud" in the comments section after the SEC investigation, consistent with participants being more concerned or sensitive to fraud. Using a linguistic algorithm, we also find a much more negative tone in the comments after the SEC announcement, consistent with readers having less trust in the articles.

We interpret our results through the lens of several theories. While the evidence is broadly consistent with themes about social trust affecting economic activity, we also find that producers of news on the social platforms respond to the SEC event, creating a new equilibrium that also contributes to the results. Applying NLP to the articles themselves, we find no differences in tone, clout, or other linguistic cues after the SEC shock, with one notable exception: authenticity, a measure designed to detect deception in language. Articles score higher on authenticity *after* the SEC scandal, suggesting that producers of news on these platforms are optimally responding to the new equilibrium. This change by producers of news should *increase* trading activity in response to their articles under a model of trust. Hence, the equilibrium response of news producers would not explain our results under a model of trust and would underestimate the decline in trading response from consumers. However, under a model of uncertainty and disagreement, where fraud raises the level of uncertainty (Fisher and Verrechia (2000), Dye and Sridhar (2004), and Fang et. al (2021)), increased authenticity reduces disagreement, which reduces trading, and hence our result could be consistent with those models. We also find that articles written after the scandal focus more on earnings and other "hard" information that is more easily verifiable and hence more trustworthy, or more precise.

Our results highlight the spillover effects and indirect consequences of market manipulation. A vast literature focuses on the direct effects of stock market manipulation.[1] We study their broader implica-

[1] Aggarwal and Wu (2006) examine SEC litigation cases, Frieder and Zittrain (2007), Hanke and Hauser (2008), Hu

tions, such as deterioration of trust in markets, merging the literature on stock market manipulation with the literature on social capital (Knack and Keefer (1997), Guiso et al. (2004), Guiso et al. (2008), Guiso et al. (2010), Sapienza and Zingales (2012), Gomes et al. (2021)). Little evidence exists on how trust evolves in markets or the consequences of changes in trust. We interpret the SEC event as a shock to public trust in social news platforms generally that deteriorated public trust in this particular news source, but not necessarily other news media (e.g., newspapers). We provide some of the first direct evidence of how a new equilibrium evolves following a shock to public trust.

The results may also relate to the implications of fake news and media bias. Misleading information can impact social, political, and economic relationships. While analysis of these issues has primarily been theoretical,[2] our paper provides novel empirical evidence. Our findings may be consistent with news being tailored to readers' priors, such as in Vosoughi et al. (2018) and Edelson et al. (2021), and news-producers sacrificing longer-term reputational capital in lieu of short-term gains (Allcott and Gentzkow (2017)). Contrary to the social media literature on fake news, which is limited to providing evidence on the excess proliferation and readership of such news, we provide evidence that these news affect behavior. The decline in trading activity to all news, including legitimate news, following the revelation of the SEC investigation is also consistent with Aymanns et al. (2017) and Allcott and Gentzkow (2017), where fake news increases distrust of media in general.

Finally, our setting suggests reasons to be both cautious and optimistic in generalizing the findings. One of the benefits of financial markets is the ability to quantify outcomes. On the other hand, arbitrage forces and market efficiency may mute its effects. If market manipulation can impact U.S. equity markets, where competition for information is fierce, markets are liquid, and arbitrage activity exists, then it could have even greater influence in settings where information costs are higher and the ability to correct misinformation is more limited (e.g., online consumer or political markets).

The rest of the paper is organized as follows. Section 1 briefly motivates our analysis through the lens of several theories. Section 2 describes the shared financial networks and discusses the scandal of fraudulent paid-for articles that led to the SEC's investigation. Section 3 documents the direct impact

et al. (2009), Nelson et al. (2013), and Hu and McInish (2013) examine the efficacy of email spams touting stocks, Delort et al. (2009) and Sabherwal et al. (2011) examine the use of online message boards to manipulate prices, and Ullah et al. (2014) examine third-party false information releases.

[2] Allcott and Gentzkow (2017) model fake news as an extension of Gentzkow and Shapiro (2005) and Gentzkow et al. (2015) on media bias, where fake news occurs in equilibrium when agents cannot costlessly verify the truth and the news matches the agent's priors. Aymanns et al. (2017) provide an equilibrium model of an adversary using fake news to target agents with a biased private signal, where knowledge of the adversary causes agents to discount all news. Kshetri and Voas (2017) discuss the pervasiveness of fake news and its dissemination across news consumers. False content may impose private and public costs by making it more difficult for readers to infer the truth, reduce positive social externalities from shared-information platforms, increase skepticism and distrust of legitimate news, and potentially cause resource misallocation. Readers may also derive utility from fake news (as entertainment or if slanted toward their biases, as in Mullainathan and Shleifer (2005)).

on markets from articles on these platforms, including the fraudulent articles from the scandal. Section 4 uses the revelation of the SEC event to examine the indirect impact of manipulation on markets. Section 5 concludes.

1. A Brief Theoretical Motivation

We briefly discuss several theories that guide the empirical analysis and help interpret the results. Our sample of fraudulent articles from the SEC are authored by paid agents acting on behalf of a principal to manipulate the stock price.

Financial markets provide a useful setting to examine the impact of manipulation because they provide high frequency outcomes, such as trading volume and market prices. Trading volume provides a measure of whether investors pay attention to, and act upon, the news in the articles, which may or may not result in price impact. For example, news can cause trading without impacting prices if markets are informationally efficient (Fama (1970)), or prices can move without trading (Milgrom and Stokey (1982)). We therefore examine both volume and price movements in response to news.

We also use the SEC event and its public announcement as a shock to trust in news on the social media platforms and assess the implications of that trust. We think of the SEC event as a shock to public trust in the social news platforms rather than just "private" trust in the authors who wrote the articles. We test the idea that participants on these platforms placed less faith in their credibility following the revelation of fraud. Although the social platforms do not monitor the content of the articles or their veracity – and this is known to market participants – investors had no idea outright fraud was taking place and that articles masquerading as news or opinion were actually company-sponsored propaganda. Hence, the SEC event was a shock exposing fraud that likely deteriorated trust in these networks. However, this public distrust of the social news platforms likely did not extend to all news sources, such as print media (e.g., the Wall Street Journal or NY Times), the associated press, and corporate announcements that operate very differently than crowd-sourced news platforms, with different incentives and rigorous monitoring. These are empirical questions we aim to test by looking at the response to social news media sources and to non-crowd sourced news media as a placebo test for our event study.

The premise we have in mind is an investor trying to extract signals from a news source – in this case one provided through crowd sourcing – about a stock's future share price. Consequently, investors will consider how reliable or trustworthy this news source is. Revelation of fraud on the social platforms will erode that trust and make news from the social networks less credible, all else equal. In a perfectly

rational world with full information, a representative investor should place no weight on any manipulated or fraudulent signal and hence revelation of fraudulent news from the SEC investigation should have no impact (Stein (1989)). However, with uncertainty, fraud can raise the conditional variance of reported news (through a variety of channels: Fisher and Verrechia (2000), Dye and Sridhar (2004), Fang et. al (2021)), which a rational investor will take into account, downweighting the news source when it is deemed less reliable. The SEC revelation may be a shock to the uncertainty of news coming from the social platforms, where investors may rationally discount news after the shock.

Under a model of heterogeneous beliefs, the manipulated news may influence some agents more than others. For example, suppose there are naive investors who are persuaded by the news and rational arbitrageurs who are not, but may be limited in the size of their positions due to risk aversion or limited capital. In this case, the fraudulent news will increase trading and may or may not result in increased price impact and volatility, depending on the relative dominance of naive versus rational traders – another reason to examine volume separately from price changes. Any price movement should reverse if driven by naive traders being fooled by fraudulent news. A revelation of fraudulent news on these platforms may then alter the beliefs of naive investors who may discount or distrust the signals they receive from these platforms. Moreover, the announcement of fraud may cause investors to discount all news from these platforms, including legitimate news, causing a spillover effect from the announcement on all news from this source. In our empirical tests, we examine the indirect spillover effects of the revelation of fraud from the SEC announcement on other news and other firms, to test the broader implications of market manipulation.

2. Knowledge-Sharing Platforms and the Social Media Scandal

We describe the knowledge-sharing financial news platforms and the SEC-investigated scandal of manipulated articles.

2.1 Knowledge-sharing platforms

Our sample comes from a website called Seeking Alpha, supplemented with a competitor platform, Motley Fool. Seeking Alpha is an online news service provider for financial markets, whose content is provided by independent contributors. The company has distribution partnerships for its content with MSN Money, CNBC, Yahoo Finance, MarketWatch, NASDAQ, and TheStreet. The Motley Fool is a multimedia financial-services company that provides financial advice for investors through a shared-knowledge platform. They are the two largest financial crowd-sourced sites.

As the popularity of these platforms grew (Seeking Alpha grew from two million unique monthly

visitors in 2011 to over nine million in 2014, generating 40 million visits per month), concerns of their susceptibility to fraud surfaced due to these sites being virtually unregulated, frequented by retail investors, and because authors can use pseudonyms (though the platforms claim they know the true identity of each author, which was eventually subpoenaed by the SEC). Authors are allowed to talk up or down a stock that they are long or short, provided they disclose any positions they have in the stock in a disclaimer accompanying the article. Failure to disclose has legal ramifications. According to Section 17b of the securities code, it is illegal to fail to disclose any direct or indirect compensation that the author received from the company, its representative, a broker-dealer, or an underwriter. Appendix A details how authors on these sites contribute and are compensated.[3]

2.2 Rick Pearson and the SEC Investigation

In 2014, an industry insider, Rick Pearson, who was a regular contributor to Seeking Alpha, was approached by an investment-relations firm, DreamTeam, to promote certain stocks looking for "good news" by writing articles with false information for a fee without disclosing the payment. Specifically, Mr. Pearson was asked to write paid promotional articles on Galena Biopharma and CytRx Corporation. Instead, Mr. Pearson went undercover to investigate how rampant this practice was and uncovered more than one hundred fraudulent, paid-for articles by other authors who did not disclose their compensation. He turned the evidence over to the SEC, who investigated each of these cases starting in February 2014. The first lawsuit (against Galena Biopharma) was brought on October 31, 2014, in a claim against the authors, the promotion firms who paid them, and the companies and their executives who hired the promotion firms. Subsequent lawsuits were also filed on April 10, 2017 and September 26, 2018.

We use the announcement of the initial SEC investigation as a shock to the public revelation of fraud, and show that the timing of this shock was unlikely confounded by other effects, recent trends, or unobservables likely to affect our outcome variables. Since the investigation stemmed from Rick Pearson's undercover work, the timing of the announcement should be unrelated to any outcome variable we analyze, except from the event itself. Below, we briefly describe some of the details of the first case, against Galena Biopharma, which provides a micro study of the direct impact these fraudulent articles had on the stock's trading activity and prices, as well as the motivation behind them.

2.3 Galena Biopharma

On October 31, 2014 the SEC filed a lawsuit in the United States District Court on behalf of all persons who bought Galena's common stock between August 6, 2013 and May 14, 2014.[4] Figure 1

[3] In June 2012, Seeking Alpha announced it would no longer permit publication of articles for which outside compensation had been paid.

[4] (Case 3:14-cv-00558-SI): http://securities.stanford.edu/filings-documents/1051/GBI00_01/20141031_r01c_14CV00367.pdf.

depicts the stock price of Galena from April 2013 to May 2014, as well as the events that led to the lawsuit. According to the lawsuit, Galena worked with investment relations companies Lidingo and DreamTeam to publish a series of promotional articles on Seeking Alpha that Galena paid for, where the payments were not disclosed by the authors and in some cases, authors falsely claimed *not* to have received any payment. Appendix B contains an example of one of the fake articles written about Galena. Figure 1 shows that Galena's share price rose from about $2 to over $7 between the summer of 2013 and January of 2014. The publications of the promotional articles are highlighted on the graph by green boxes and often coincide with a bump in stock price on that day and a steady increase in price several days after. The motivation behind the campaign seems to be a pump-and-dump scheme, as Galena insiders took advantage of the price rise through corporate actions and their own personal trading. On September 18, 2013 Galena sold 17,500,000 units of stock in a seasoned equity offering for net proceeds of $32.6 million. On November 22, 2013, Galena held a board meeting and granted stock options to executives and directors with a strike price of $3.88. In January 2014, after the stock price reached its highest level since 2010, seven Galena insiders sold most of their shares in less than a month, for more than $16 million. These events are highlighted in Figure 1.

In February and early March 2014, several investigative journalists published exposé articles documenting the fraud, including in Barron's and Fortune. On March 17, 2014 Galena revealed in a 10-K filing that it was the target of an SEC investigation over the promotion. The SEC brought charges against Galena and its former CEO Mark Ahn "regarding the commissioning of internet publications by outside fake firms." Mr. Ahn was fired in August 2014 over the controversy, and in December 2016, the SEC, Galena, and Mr. Ahn reached a settlement. Appendix B reports the 8-K form documenting the settlement. By that point, Galena's stock price had dropped to $2 a share.[5] We exploit the timing of the SEC announcement and subsequent media attention as a shock to investors' information of fraudulent news on these platforms, and examine the impact on markets before versus after the event.

2.4 Fraudulent Articles

Mr. Pearson kindly provided the articles to us that he determined to be fraud: 111 articles by 12 authors covering 46 publicly traded companies. We also obtained a second set of known fraudulent articles that the SEC identified during their investigation.[6] Seeking Alpha kindly shared 147 of those articles with us, as they had been removed from the platform. Among those articles, we match 60 to

[5]Interestingly, while Galena is a relatively small firm, it was not an obscure one. For example, in July 2013, before the promotion started, it had a market cap of approximately $350 million, and it was followed by analysts at Cantor Fitzgerald, JMP Securities, Oppenheimer & Co., and others. Furthermore, according to the SEC lawsuit, more than a hundred market makers facilitated trading in the company's stock.

[6]The full list can be found here:
https://ftalphaville-cdn.ft.com/wp-content/uploads/2017/04/10231526/Stock-promoters.pdf.

firms publicly traded on U.S. exchanges to obtain price and volume information from the Center for Research in Security Prices (CRSP). The rest of the articles pertain to firms traded over the counter. Combining all of the data sources, our final sample consists of 171 fraudulent articles written by 20 different authors about 47 publicly traded firms.[7]

It is useful to define what we mean by *fraudulent* articles. In the sample from Rick Pearson and the SEC, the articles were paid for by a promotional firm to deceive the market and manipulate the stock price. Consequently, these articles contained information of some kind that authors knew to be incorrect at the time. How false or wrong that information turned out to be is difficult to assess. Our sample of fraudulent articles is about *intent* to deceive, where the articles contain information that the authors knew at the time to be false.

To provide some insight into the content of these promotional articles, we highlight a recent example from our sample that was part of the SEC's lawsuit filed in September 2018. One of the fraudulent, paid-for articles in this case was a publication that appeared on Seeking Alpha on September 26, 2013 about the company Biozone. The article stated,

> *Biozone has developed a new method of drug delivery, QuSomes that provides improved efficacy, reduced side effects, and lower costs. This technology will allow Biozone to reformulate and sell certain FDA approved drugs at a reduced cost, which should help Biozone capture a large percentage of these drug markets.*

From the SEC lawsuit filed in September 2018 in the District Court of New York City:

> *Keller misleadingly stated that Company A had a formulation ready for testing to be brought to the billion-dollar injectable drug market. Yet, as Keller knew, as of summer 2012, all R&D efforts had been shut down without the successful formulation of an injectable drug and Company A had ceased all efforts to develop this technology in mid-2012.*

"Keller" refers to Brian Keller, the co-founder and Chief Scientific Officer of Biozone, who had paid for the promotional article. Many of the paid-for articles involve similar issues.

Our focus, however, is on investor *trust* or credibility in these platforms, and not necessarily the actual content of the articles. The scandal of fraudulent articles and the SEC investigation provide a shock to that trust in this medium, allowing us to assess the indirect impact of market manipulation. The content of the articles themselves embody direct manipulation, which requires knowing the content of the articles, how that deviates from the truth, and the market's expectations. There is little evidence on the indirect effects of manipulation and the equilibrium response.

[7] While we gain 60 additional articles from the SEC, we only gain one additional firm. Most of the additional articles pertain to firms already covered by Rick Pearson, and hence simply give us more articles about the same firms.

3. Direct Impact of Social Media Articles

We first establish that articles on these platforms matter and that investors pay attention to them, and then show that fraudulent articles had a disproportionate impact on markets.

3.1 Do articles on the platforms have impact?

We briefly address whether articles posted on these social platforms impact markets at all, or whether these social platforms are just a side show. Prior research (Antweiler and Frank (2005), Chen et al. (2014), Cookson and Niessner (2020)) finds that social media news platforms predict trading volume, stock prices, and earnings surprises. We supplement those findings with our unique sample to further support that these social information platforms have impact.

We manually download all articles: 203,545 articles on Seeking Alpha over the period 2005 to 2015 and 147,916 articles on Motley Fool from 2009 to 2014, to obtain the content of the articles, authorship, and in the case of Seeking Alpha, comments from other users. This broader sample serves two purposes. It allows us to gauge how markets respond to articles on these platforms in general, and it allows us to assess the response of consumers and producers of information on the social news platforms to the revelation of fraud from the SEC investigation in the next section.[8]

We examine *abnormal* trading volume around the publication of articles on these platforms to capture whether investors "react" to the articles. It is likely the reverse is true as well, that articles react to trading activity. To try to establish some causal interpretation, we examine *abnormal* or unexpected changes in *future* trading volume from when the article is written. A reverse causality story would imply that authors are writing articles in anticipation of *future unexpected* trading activity, which we would call "news." We also control for lagged abnormal volume from the previous trading day to capture reaction to all events up to day $t-1$. Establishing a direct causal link between articles and trading activity is challenging due to omitted variables likely driving both. However, establishing causality here is not the goal. Rather, our aim is to examine the equilibrium response of market participants, consumers, and producers of news to the shock from the SEC event.

We specifically focus on retail trading activity (versus institutional). We conjecture that retail traders, in particular, are participating and responding to these platforms and hence expect a stronger response from retail trading activity. A conjecture we confirm in the data. We identify retail trades using TAQ data and the algorithm proposed by Boehmer et al. (2020). Their method uses the fact that

[8]In a previous version of the paper, we also use these articles for a third purpose to apply a linguistic algorithm to detect false content, that we validate and calibrate using Rick Pearson's known fraud and legitimate articles as a training sample. The algorithm had some success, but also a lot of noise. We relegate those results to an internet appendix, as they are tangential to our focus here.

most retail trades do not take place on registered exchanges. Instead, they are often filled internally by the broker or are sent to a wholesaler (e.g., Citadel). Those trades must be reported to a FINRA Trade Reporting Facility (TRF). These TRF executions are reported in TAQ with exchange code "D." Many orders that are executed off-exchange are given a small price improvement relative to the National Best Bid of Offer. Following Boehmer et al. (2020), we classify trades with TAQ exchange code of "D" and prices with just above or below a round penny as retail trades.

Table 1 examines the relation between future three-day abnormal retail trading volume and article publication on these sites. We define abnormal retail trading volume for stock i on day t as $RetVol(i,t)/\frac{1}{T}\sum_{k=1}^{140} RetVol(i,t-k)$, which is the retail trading volume for stock i on day t relative to the average daily retail trading volume in stock i over the last 6 months.[9] We sum abnormal retail volume over days $t=0$, $t+1$, and $t+2$, where $t=0$ is the date the article appears on the website. We then regress the natural logarithm of abnormal retail volume on an indicator variable for whether an article on these sites appeared about firm i on date $t=0$. We include year-month fixed effects in the regression. We examine only firms that had at least one article published on Seeking Alpha or Motley Fool over the sample period (about 7,700 unique firms and 9.8 million firm-day observations).

As the first column of Table 1 shows, an article published on these platforms is associated with a 36.5% increase in abnormal retail trading volume over the three days following publication. This result implies that investors are either trading in direct response to the articles or, more generally, in response to whatever news is coming out that day that these articles are discussing. While the increase seems large, this first regression controls for no other variables, except time fixed effects. As we show below, the bulk of the effect is also concentrated in very small and illiquid firms, where percentage trading volume changes can be large.

Articles are often written following press releases or corporate filings with the SEC, so in the second column we control for whether there is an SEC filing (10-K, 10-Q, or 8-K), a company-issued press release, or a media article (WSJ and NYT) in the three days leading up to the article. We also include lagged abnormal retail trading volume as a regressor to help capture other events affecting trading activity. With these controls, the effect on abnormal retail trading volume over the next three days is 19.5%. To make sure that these results are not all coming from the day the news is released, Table B1 in Appendix B reports the effect on retail trading volume separately for the same day and for one and two days after the article's publication. Of the 19.5% rise in abnormal retail trading volume, 8.8% occurs on the day the article is published, 6.2% the following day, and 4.5% two days later. The abnormal retail trading volume following an article increases for about two weeks before returning back to its previous

[9] Results are nearly identical defining abnormal volume relative to the last 30, 60, or 180 days.

6-month daily average.

The third column of Table 1 reports results for small firms separately, where we interact the article dummy with a small firm dummy (defined as firms smaller than the bottom 10^{th} percentile of NYSE firms based on beginning of month market cap). The effect on abnormal retail trading volume rises strongly for small firms, with the effect being nearly three times larger for small firms. This result is consistent with small firms having less volume and liquidity, less active large investors, and a more opaque information environment.

3.2 How impactful are the fraudulent articles?

The last column of Table 1 adds a dummy variable for the 171 fraudulent articles identified by Rick Pearson and the SEC. The estimated coefficient on these fraudulent articles from the scandal is 0.554 with a t-stat of 2.22, indicating that the fraudulent articles increased retail abnormal trading activity by 55.4%, which is 2.5 times larger than the effect from a typical article published on the social news platforms. This result indicates that the fraudulent promotional articles were successful in having impact on investor trading. The effect is large, but reasonable considering that the SEC selects cases ex-post that had the largest impact, and is in line with excerpts from recent SEC lawsuits.[10] The stronger impact on trading activity may also be driven by fraudulent articles being more sensational and diffusing more quickly across readers (Vosoughi et al. (2018)). The larger impact may also indicate that promotional articles are different than other articles along other dimensions that might also affect trading volume, which we investigate below.

3.3 Return reaction event study

If the fraudulent articles were successful in pumping up the stock price, then we should see an impact on share prices as a direct result of these promotional articles. To test this conjecture, we focus on the sample of firms from Rick Pearson and the SEC and compare the fraudulent articles to a set of legitimate articles written by the same authors in order to difference out unobservable heterogeneity in author style and reputation to better identify the impact of the promotional articles controlling for other effects that may also affect investor attention. We obtain 334 non-fake or legitimate articles published on the same social news platform, covering 171 companies, from the same set of authors that were investigated by the SEC.

Figure 2 plots the difference between cumulative abnormal returns (CARs) for days with fraudulent

[10] From Case 1:18-cv-08175 filed on September 7, 2018 in the U.S. District Court, Southern District of New York: *"The market reacted strongly to the Company A promotion: the trading volume of Company A stock rose from approximately 1,100 shares on September 25, 2013 to over 4.5 million shares on September 27, 2013 and to more than 6 million shares on October 2, 2013."* And, in the same case about another firm, *"The article did not disclose that the author had been paid by Company B – at Honig's direction – to write the article. After the article was published on February 3, 2016, there was a 7000% increase from the previous day's trading volume, and an intraday price increase of over 60%."*

articles, relative to days with legitimate articles, from 20 days before the article's publication to 250 days after, for small and non-small firms separately. CARs are defined as the return on the stock minus a portfolio of stocks with similar size, BE/ME, and momentum characteristics following Daniel, et al. (1997).

Abnormal returns for small firms increase after a fraudulent article is published (relative to legitimate articles), reaching as much as 13% cumulatively after about 60 days, before giving up all the gains and ending with a cumulative −5% return after a year. This finding supports another implication that the initial price impact from fraudulent news should eventually reverse. This pattern matches that of Galena Biopharma, the first SEC-prosecuted case, in Figure 1. The permanent price impact of −5% for small firms indicates either that once the market figured out there was fraud, investors viewed this as a bad signal for the firm, or that the true price should have dropped by 5% initially, but the promotional campaign temporarily propped up the price. For non-small firms, the price starts dropping immediately after the fraudulent article comes out and continues to decrease over the year. This result is consistent with the market figuring out the news is false immediately for larger firms, where the cost of information is lower. The results are consistent with the SEC's claim that the fraudulent articles were part of pump-and-dump schemes designed to temporarily prop up share prices.

For robustness, Table B2 in Appendix B reruns the regressions in Table 1 using idiosyncratic price volatility of the stock as the dependent variable. We examine price volatility as opposed to signed returns because it is difficult to sign the direction of the content of the articles (see for example Loughran and McDonald (2016)). We find that daily price volatility also rises following articles published on these platforms. The effect is also strongest for the smallest firms. The effect of a published article on idiosyncratic stock volatility is about 0.07% over the three days, with the effect an order of magnitude larger for the smallest firms, which is consistent with larger price impact for the smallest stocks (Frazzini et al. (2018)). Fraudulent articles have an additional 3% significant impact on price volatility over the next three days relative to legitimate articles. We also obtain a proprietary supplemental dataset from Seeking Alpha on the readership of articles. Table B3 in Appendix B shows that future abnormal retail trading volume is positively related to the number of clicks and number of times the article is read by investors, suggesting articles that influence circulation and readership are also associated with more retail trading activity in the stock. This collection of results indicates that the SEC announcement of the scandal could have a significant impact on markets if it changes the way participants consume news from these platforms and/or how news is subsequently produced on these platforms.

4. Indirect Impact of Fraud

Our main and novel test is on the indirect impact of fraudulent news. The public revelation of the SEC's investigation and subsequent media attention around it may provide a shock to investor awareness of fraud and hence trust in social media news sources. We use this event to test the spillover effects of fraud and its impact on financial activity. The indirect effects of fraud help capture the equilibrium response to the shock and are novel to the manipulation literature, which focuses primarily on the direct effects of misinformation, and distinguishes among theories that do not imply equilibrium spillover effects (Allcott and Gentzkow (2017), Kshetri and Voas (2017), and Aymanns et al. (2017)).

4.1 Spillover effects from unexpected revelation of fraud

We use the period from February to March 2014 as the event that reveals the fraud on knowledge-sharing platforms. We examine the abnormal retail trading activity associated with published articles over the six months prior versus the six months after the event (i.e., August 2013 to January 2014 versus April 2014 to September 2014).

Panel A of Table 2 examines the impact of all published articles on future abnormal retail trading volume before versus after the scandal. The first column reports results from a regression of future abnormal retail volume on an article indicator, plus controls for SEC filings, press releases, print media, and lagged volume. Confirming the results from Table 1 over this one-year snapshot of our sample (six months before to six months after the SEC announcement), published articles on the social media platforms are associated with significantly higher future abnormal retail trading activity. The second column adds the 6-month post event indicator, and its interaction with the article dummy. The negative interaction term with the post-event dummy shows that the effect of articles on future retail trading volume decreases significantly after the scandal by about 3.8% relative to before the scandal. This result is consistent with investors becoming aware of fraudulent content and muting their response to news in general on these platforms.

The third column adds the small firm dummy and interacts it with the article indicator and the post-event dummy. First, future retail trading volume associated with a published article declines by 4.1% after the scandal's announcement (t-stat of -2.03). Second, the retail volume response to articles is still much stronger for small firms (coefficient of 0.643 with a t-stat of 10.00). But, third, the triple interaction of article $\times$ small $\times$ post-event indicates that retail trading volume declines even more so for articles about small firms *after* the scandal (coefficient of -0.235 with a t-stat of -2.94). This last result indicates that retail trading volume for small stocks is the most sensitive to news from these platforms and hence is affected most by the announcement of fraud on these platforms. Retail trading activity in

the firms where it matters most – small firms – are the most responsive to the shock of the scandal.

Panel B of Table 2 reports results from regressions using idiosyncratic price volatility as the dependent variable. There is also significantly reduced impact on price volatility from articles after the scandal for small firms. In particular, after the revelation of the scandal, the effect of a published article on small firms is a 1.3% decline in idiosyncratic volatility over the next three days, relative to baseline volatility of 7.2%. Conversely, we find an *increase* in price impact from press releases, SEC filings, and other news media after the scandal, which indicates that other trends in news or omitted variables in the market's response to news are, if anything, biasing our results in the opposite direction.

4.2 Identification and addressing alternative explanations

While we interpret our findings as the shock of fraudulent news on these platforms subsequently affecting the impact of future articles on financial markets, we consider here alternative explanations. For instance, the market environment surrounding the SEC announcement may have changed, which could also have caused variation in the relationship between news and market activity. In addition, the news generation process itself could have changed. In short, there is an identification challenge in understanding the relationship between news and trading activity, since there are many potential omitted variables affecting both.

To address alternative explanations for our results, we run a number of additional tests. First, we remind the reader that the timing of the SEC event should be exogenous to the outcome variables we study, since the source of the scandal was Rick Pearson's undercover work, which was initiated by him being idiosyncratically approached by a PR firm. In addition, we control for other news sources (SEC filings, press releases, and print media) happening at the same time in order to account for variation in news in general. However, to more directly and convincingly address alternative hypotheses and potential omitted variables, we examine pre-trends in trading activity in response to news and test whether reaction to news is simply lower in the post-scandal period for other, unrelated reasons. We also conduct several placebo, or falsification, tests to rule out alternative explanations by looking at institutional trading activity (institutions are largely not attending to these social media cites) and examining the market's response to news sources unrelated to and unlikely affected by the scandal (such as the WSJ and NYT or press releases).

4.2.1 Pre-trends

We first examine pre-trends in the market's reaction to news leading up to the event. We regress daily abnormal retail trading volume on dummy variables for an article being published, the post-scandal

period, and their interaction:

$$Log(RetAbVol)_t = \alpha + \beta_1 Article \times PostEvent + \beta_2 Article + \beta_3 PostEvent + Controls + \epsilon.$$

The controls include other sources of news about the firms at the same time from the WSJ, NYT, press releases, and corporate filings, as well as lagged volume and firm fixed effects. Figure 3 plots the coefficient β_1 at the daily level (with 95% confidence error bars drawn on the graph), which represents the difference-in-difference reaction to published articles after versus before the scandal, from seven days before to four trading weeks after each article's publication. A week prior to an article's publication, there is no difference in retail trading volume response after versus before the scandal, indicating no pre-trend in response to news before the scandal. Hence, the decline in retail trading volume we find after the SEC announcement does not appear to be driven by trends in the market's response to news that were happening at the same time. After the scandal, however, there is a significant drop in retail volume response to articles on the publication date and for the next two trading weeks. These results suggest that investors' reaction to articles on these platforms decreases significantly after the scandal, lasting for two weeks before returning to normal trading levels, and appears to be a direct result of the announcement rather than any trends in market activity.

Another possibility is that articles became less timely after the scandal, and thus investors react less to them, as that information has already been incorporated into the market. Table B4 in Appendix B looks at the timeliness of articles before versus after the scandal by regressing the incidence of an article on the lagged abnormal return (with respect to size, value, and momentum benchmarks) of the firm discussed in the article, using lagged returns over the previous day, week, and month. We find no evidence that articles are responding to returns of firms any faster or slower before versus after the scandal. Hence, the timeliness of article publication, at least with respect to abnormal price movement, appears to be the same after the scandal and hence is unlikely driving the lower trading volume we find after the SEC announcement.

4.2.2 Falsification tests

We conduct falsification exercises to rule out many alternative explanations for our results.

Falsification test #1: Institutional trading.

We first examine institutional trading activity, which should not be affected by the social networks, and hence the SEC announcement, since institutional traders do not actively participate on the social networks. However, if the SEC announcement is confounded by omitted factors that impact trading and

news in the economy generally, then we expect institutional trading to show (just as large) an effect. The idea is that institutional trading should also be affected by omitted variables driving news and general trading activity, but because institutional traders are not primary participants on these social networks, the effect of the scandal should not otherwise impact institutional trading.

To measure institutional investor trading, we do not simply subtract retail volume from total trading volume, recognizing that the Boehmer et al. (2020) method for identifying retail trades will not capture all retail trading. Instead, we follow Bushee et al. (2020) and use large trades (trades greater than or equal to $50,000) as a measure of institutional volume. While some institutional investors split their trades into small orders (Frazzini et al. (2018)), implying we might miss those, we are confident that large trades should reflect only institutional investor activity. The point here is to capture trades we are confident are institutional trades and not to necessarily capture all institutional trading.

Panel C of Table 2 examines the impact of all published articles on abnormal *institutional* trading volume before versus after the scandal. As the first column shows, future abnormal institutional trading volume also rises following the publication of an article on these platforms. Since institutions do not participate on these platforms, we conclude that this effect is likely driven by omitted variables driving trading and news generally. However, the second column indicates that the response of institutional trading was *unaffected* by the SEC's announcement of the scandal. The interaction between the Post-event dummy and an article is -0.013 with an insignificant t-stat of -0.59. Looking at the third column of Panel B, we also find no evidence that institutional trading dropped in response to the SEC announcement for the smallest firms either. Hence, unlike for retail trading volume, we find no evidence that institutional trading volume in response to articles on the social news platforms was influenced by the SEC announcement. These results are consistent with the scandal having an impact on those market participants who pay attention to the social networks and helps rule out concerns that omitted variables associated with trading and news generally, which should affect both retail *and* institutional traders, is driving these results.

Falsification test #2: News sources unrelated to the scandal.

Second, we examine the retail trading response to other news sources unlikely to be impacted by the SEC announcement, as they were not part of the scandal, such as articles published in the NYT and WSJ, press releases, and corporate filings. We expect no differential effect after the scandal's announcement for these other media sources, unless the announcement happened to coincide with other news or changes in the market's response to news generally. Panels A and B of Table 2 show that these other news sources all have strongly significant and positive impact on abnormal retail and institutional trading volume unconditionally. However, we do not expect the market's reaction to these news sources to change after

the SEC announcement, unless the SEC announcement is confounded by omitted unobservables.

We interact the Post-event dummy with SEC filings, firm press releases, and other news media (WSJ and NYT articles). The interaction terms serve as falsification tests or "placebo" tests of whether the market responds to other news differently after the scandal. If the post-event period happened to coincide with less information content, less news, less firm activity, or less trading to news, for instance, then the Post-event reaction to *all* news sources should be significant. If, however, investors responded to the scandal itself, which just pertained to the social knowledge-sharing platforms, then we should only see a decline in trading response to news on the social platforms and not to other media sources.

Panel C of Table 2 reports the results of these falsification tests. As the first column shows, the interaction terms for retail trading volume are statistically and economically zero, with two of the three coefficients having the wrong sign to be consistent with the alternative story of omitted variables. We find no discernible difference in abnormal retail trading volume response to other news sources before versus after the scandal. This evidence is inconsistent with trends in news, trends in trading activity, or changes in response to news generally explaining our findings. The scandal only seems to affect the market's response to articles contained on the social networks.

The second column of Panel C of Table 2 repeats these tests for abnormal institutional volume, which also shows no effects, economically or statistically, after the scandal. Institutional trading in response to other news sources – NYT, WSJ, press releases, and corporate filings – exhibit no change after the SEC announcement of fraud on the social networks.

The evidence from these falsification tests indicates that the market's response to news in general – both from retail and institutional investors – exhibits no changes after the SEC announcement, rejecting the alternative hypothesis that omitted variables are driving the decline in response to social media articles after the announcement. Rather, our results appear to be most consistent with retail investors, who primarily participate on the social platforms, discounting all news on the social media platforms, including legitimate news, due to distrust after the revelation of fraud on those networks. The magnitude of the drop in abnormal retail volume is even larger and more significant (-5.6% drop per article with a t-stat of -2.97) after accounting for interactions of the other news sources with the Post-event dummy. In other words, controlling for potential changes after the SEC announcement in the market's response to other news sources that are unlikely affected by the scandal, our results become stronger, suggesting that omitted variables may be understating our findings. The results for institutional trading volume remain statistically no different from zero, consistent with institutional traders being largely inattentive to, and unaffected by, the social media news source.

4.2.3 Alternative explanations

While our tests and results indicate that the announcement of the scandal itself caused the precipitous decline in retail trading activity, the channel through which that decline occurs is open for debate. For instance, we interpret the reaction to the announcement through the lens of trust in the social media platforms and interpret the event as a shock to trust. However, other mechanisms, such as a change in the precision of news or investors' perception of the uncertainty in news (Fang et. al (2021)) could also be consistent with the results. The SEC event had a direct impact on the market's response to news on the social platforms, but the exact channel driving that response is not clear. Below, we provide some additional evidence suggestive of the trust channel, but cannot completely rule out other channels leading to the change in market response after the SEC event.

One test, however, that can help distinguish between uncertainty and trust is to look at disagreement among investors. If there is either a change in the precision of news or investors' perception of the uncertainty in news, we expect that investors would disagree less about the news after the scandal. Since we focus on trading of retail investors, we use a measure of disagreement among mostly retail investors proposed by Cookson and Niessner (2020) to study this question. The disagreement measure is constructed at the daily level from opinions expressed by mostly retail investors on a popular social network for investors called StockTwits. We examine whether disagreement among investors increases on the days with Seeking Alpha articles, and whether that reaction is different after the disclosure of the SEC investigation, which we expect it to be if news on these platforms becomes less uncertain. The results are presented in Table 3. The first two columns look at all firms, and the last two look at small firms and non-small firms, respectively. We also control for the number of messages about the firm on StockTwits on a given day, to capture disagreement controlling for attention. The disagreement measure is standardized to be zero mean and unit standard deviation. We find that days with articles on Seeking Alpha are associated with about 1/3 standard deviation more disagreement than days without articles. However, that difference doesn't change after the revelation of the SEC investigation, suggesting that our main results are unlikely to be driven by changes in uncertainty of news published on Seeking Alpha.

4.3 Further corroborating evidence of trust

We provide some corroborating evidence to suggest that change in trust in social media due to the scandal may be driving the market's changing response to articles on these platforms. Specifically, we examine cross-sectional characteristics that are consistent with investor's awareness of fraud causing the decline in retail trading activity.

4.3.1 Article and author characteristics

Panel A of Table 4 reports results from regressing abnormal retail trading volume associated with an article on various article and author characteristics (controlling for other news from corporate filings, press releases, WSJ and NYT articles, lagged abnormal trading volume, and year-month and firm fixed effects, and in some specifications, author fixed effects). We first look at the number of followers of the author, a proxy for author popularity. Conditional on an article being published, authors with greater past followings have a significant positive impact on retail trading volume (an additional 1.3% with a t-stat of 8.96). We also show (second column) that authors who have written more articles in the past have a bigger impact on retail volume (increase of 0.5% with a t-stat of 3.29). These results suggest that articles written by authors with better reputations and more experience are associated with more subsequent retail trading volume. Reputation and experience seem more consistent with trust than other channels (e.g., precision).

Columns 3 and 4 analyze whether the retail trading volume response to an article is higher when the article is more quantitative in nature and/or references accounting data, where presumably the information is easier to verify from other sources. We look at the fraction of the article text comprised of numbers, and fraction of text containing the bigram "earn." These articles have a bigger impact on retail trading volume of 2.6% and 4.5% (t-stats of 9.07 and 12.65), respectively. The results are consistent with investors putting more trust in articles with hard numbers or that such articles contain more precise news or are associated with bigger news. To distinguish between these hypotheses, we look at how they relate to retail volume before versus after the scandal.

We also examine the comments section to each article. The number of comments is strongly positively related to retail trading volume (increase of 5.4% for a one standard deviation increase in number of comments, with a t-stat of 10.84), as is the average length of comments (increase of 2.3% with a t-stat of 6.68). The number and length of comments may proxy for disagreement about or interest in the articles, or may be a symptom of more activity associated with articles containing more news. To more directly look at disagreement, we use natural language processing on the comments and create a dummy variable for whether disagreement words ("disagree," "differ," "counter," "oppose," or "argue") or "wrong" words ("wrong" or "not right") show up in the comments section. In both cases, associated abnormal retail trading volume accompanying these articles rises. Viewed as proxies for disagreement, these results are consistent with theories arguing that disagreement is associated with more trading.

We also add up mentions of the words "fake" or "fraud" in the comments section and compute a dummy variable equal to one if readers use these words. We find an additional positive impact on

retail trading volume when comments contain these words. Thus, comments of "fake" and "fraud" are associated with *more* trading activity. However, we show that this result changes after the SEC announcement. Finally, we also include a dummy variable for authors who disclose they have a position in the stock they are writing about. The impact on volume is negative but insignificant.

Panel B of Table 4 examines whether the author and article characteristics have any differential impact after the SEC announcement of the scandal. We rerun the regressions from Panel A, examining the time period six months before to six months after the SEC scandal, and interact each characteristic with the Post-event dummy. The interaction terms suggest that after revelation of fraud, the reduction in retail trading activity is driven by investor concerns over the credibility of news on the social media platforms. For example, as the first two interactions show, the drop in trading volume is not as large if the author has more followers and has written more articles in the past. Reputational cues of the author become more important after the scandal, and do not appear to be important before the scandal. The results are consistent with market participants maintaining more trust in articles written by authors with better reputations and that reputation mattering more after the revelation of fraud. The alternative hypothesis that certain authors endogenously match with certain types of news articles should be unaffected by the event, unless the event caused a shift in the matching of authors to news. We look at the responses of consumers and producers of articles in the next subsection. We also find positive interactions with the Post-event dummy for articles that discuss earnings or offer hard numbers, suggesting these articles are not discounted as much after the scandal.

Looking at measures of disagreement, such as the number and length of comments and comments containing "disagree" or "wrong" words, we find no significant interactions with the Post-event dummy. Although these measures of disagreement are shown to have a strong positive overall association with retail trading volume (Panel A), their impact on trading is no different after the scandal. These results indicate that disagreement is no more influential on trading activity before versus after the scandal, which makes sense since the SEC announcement should have no effect on disagreement per se. In essence, the results for disagreement provide another placebo test that shows that investor disagreement and its impact on trading had no coincidental shift with the SEC announcement. This result suggests that other influences on trading besides revelation of fraud did not change after the scandal. Only measures of reputation and credibility seem to have more prominence after the scandal.

Finally, we also interact the Post-event dummy with the dummy for "fake" or "fraud" appearing in the comments section. While such comments have a positive impact on retail trading volume before the scandal, they are associated with even larger declines in retail trading volume after the scandal. Hence, concerns of fraudulent content echoing in the comments to an article significantly reduces the article's

impact on retail trading activity, but only *after* the scandal came to light. This result suggests that investors became aware of the fraud and increased their concern of fraud after the SEC announcement, and that shift in awareness resulted in a more muted response to news on the social media platforms. Finally, disclosure of an author's position in the stock also has a significant negative impact on retail trading, but only after the scandal (-10.3% with a t-stat of -3.10). Overall, these results further support that the decline in retail trading activity for articles published after the SEC announcement is a response to the shock of distrust in news from these platforms due to the scandal.

4.4 Equilibrium response of authors and readers

Another explanation for a muted volume response after the scandal is that readers are responding to changes in author and article characteristics, rather than trading less due to erosion in trust. Here, we examine the equilibrium response of readers and authors to the SEC announcement. Table 5 reports results from regressions of article and comment characteristics – fraction of numbers used in the text, mentions of earnings, number and length of comments and use of disagreement, "wrong," and "fake" words in comments – on the Post-event dummy, looking only at articles published six months before to six months after the SEC announcement period. We also control for other news about the firms through SEC filings, press releases, NYT and WSJ articles, and lagged abnormal volume, and include firm fixed effects. We run the regression separately for small firms and non-small firms, reported in Panels A and B, respectively.

For small firms, we find that after the scandal, articles are more likely to be written about earnings or accounting information, which makes sense if authors recognize that the scandal may have eroded trust from their readers. Note, however, that our previous results show that these changes should *increase* rather than decrease retail trading volume. Hence, authors' equilibrium response to focus on more hard accounting information could understate our estimated decline in retail trading volume after the SEC event. If, however, hard information is more precise and reduces disagreement among traders, then that could be consistent with a decrease in trading volume.

We also examine the comments to articles to gauge the equilibrium response of readers after the scandal. Consistent with readers distrusting social platforms and questioning their credibility more after the scandal, incidences of the words "fake" and "fraud" in comments increased significantly after the SEC event. Conversely, measures of disagreement, such as the number and length of comments, use of "disagree" or "wrong" words, is no more prevalent after versus before the scandal. These results further support our interpretation of the decline in retail trading volume from Table 2.

We can also see if author characteristics changed after the scandal, either due to self-selection or

perhaps the platforms sought to improve their reputation. In Panels C and D of Table 4, we regress the number of followers an author has, as well as the number of articles the author has written in the past, on the Post-event dummy. We find that both measures are higher in the six months after the scandal relative to the six months prior to the scandal for small firms. Hence, producers of news on these platforms do seem to respond to a new equilibrium created by the SEC shock. We interpret that response as being consistent with a shock to trust and a response to try to improve that trust. Note, however, that this response by news producers may increase retail trading (evidenced in Table 4, Panel A, which then underestimates our finding that retail trading activity decreases after the scandal.

4.5 Article Characteristics and Authenticity

We further explore the equilibrium response of consumers and producers by applying linguistic tools to the actual text of the articles and comments. We use the Linguistic Inquiry Word Count model (LIWC2015) from Pennebaker et al. (2015), which focuses on individuals' writing or speech style to measure individuals' cognitive and emotional states across various domains. The LIWC model outputs the percentage of words that fall into one of more than 80 linguistic, psychological, and topical categories. We use the broad categories of writing style measured by LIWC: "clout" (a measure of confidence or expertise in expression), "analytical" (formal, logical, and hierarchical as opposed to informal, personal, and narrative), and "emotional tone" (positive or upbeat),[11] to see if articles and comments change after the scandal along linguistic dimensions.

Table 6 reports the results from regressing the linguistic scores of the comments and articles on the Post-event dummy, including all the controls for other news and firm fixed effects. Panel A reports results for small firms and Panel B for non-small firms. The comments show a much more negative emotional tone after the scandal. The more negative tone could be consistent with readers being more skeptical of these platforms after the scandal.

Turning to the articles themselves, the linguistic algorithm detects no difference in tone, clout, or analytic characteristics of the articles before versus after the scandal. Hence, the change in tone from the comments does not appear to be driven by a change in writing style of the articles after the scandal. The linguistic characteristics of the articles appear to remain the same after the event, with the exception of one category we discuss now.

One of the LIWC categories is "authenticity," designed to detect deception in expression. Pennebaker (2011) describes which linguistic traits are associated with authenticity. In particular, truth-tellers tend to use more self-referencing words and communicate through longer sentences compared to liars. When

[11] Tone here should not be confused with "sentiment" as used in the finance literature (Tetlock (2007)). The former relates to emotional expression, while the latter relates to whether the news for prices is expected to be positive or negative.

people lie, they tend to distance themselves from the story by using fewer self-referencing words ("I" or "me"). Furthermore, liars use fewer insight words such as *realize*, *understand*, and *think*, and include less specific information about time and space. Liars also tend to use more discrepancy verbs, like *could*, that assert that an event might have occurred, but possibly did not. Newman et al. (2003) use an experimental setting to develop an authenticity score based on expression style components using similar techniques, and the Central Intelligence Agency and Federal Bureau of Investigation use similar methods to assess authenticity.

As an example, consider the two statements by former U.S. congressman Anthony Weiner before and after his admission in the "sexting" scandal.

Before admission:

> ***We*** *know for sure I didn't send this photograph. [...]* ***We don't know*** *where the photograph came from.* ***We don't know*** *for sure what's on it. [...]* ***If*** *it turns out there's something larger going on here,* ***we****'ll take the requisite steps.*

After admission:

> ***I*** *would like to make it* ***clear*** *that* ***I*** *have made terrible* **mistakes**, *that* ***I*** *have* ***hurt*** *the people* ***I*** *care about the most, and* ***I*** *am deeply sorry.* ***I*** *have not been honest with* ***myself***, ***my*** *family,* ***my*** *constituents,* ***my*** *friends,* ***my*** *supporters and the media.*

The use of "we" versus "I" and "my", the discrepancy words "don't know" and "if," and the lack of insight words like "mistakes," "clear," and "hurt" are all more prevalent in his statements when he was lying.

The algorithm uses a combination of these linguistic traits to generate the authenticity measure. We apply this measure to the articles' text and examine whether the authenticity of articles changed before versus after the SEC scandal. As the last column of Table 6 shows, the authenticity of articles about small firms increases significantly after the scandal. This finding suggests that authors responded to the scandal by providing more authentic content or that the bad actors were removed from these platforms. The results are also consistent with small companies, who engage or were willing to engage in promotional articles before the scandal, ceasing or decreasing this activity after the scandal.[12]

While this equilibrium response by authors could also affect retail trading volume after the scandal, the response from authors should increase trading volume according to a trust mechanism, though might decrease it if authenticity lowers disagreement. This result also indicates that one of the fallouts from the scandal was a shift to try and increase authenticity on these platforms, either by writing more truthful content or removing bad actors who wrote fraudulent articles.

Finally, another test to rule out that the equilibrium response of news producers is causing all of the volume decline post-event, is to shorten the event window to a period where it is less plausible that the

[12]In a previous version of the paper, we also tried to identify fake content among all of the articles in our sample, using the known fraudulent articles identified by Rick Pearson and the SEC to calibrate the model. Our ability to detect fake content was somewhat successful, the details of which are provided in an online Internet Appendix.

platforms could alter their behavior. For example, if looking immediately after the SEC announcement, it is plausible consumers of news could respond immediately by discounting news from these platforms, but less plausible that producers of news would be able to respond that quickly to alter their articles. Reducing the time-band around the scandal period reduces statistical power, but also makes it less likely that our results are driven by platform changes. Figure 4 examines the three-day retail volume response for articles before versus after the scandal over different horizons, ranging from articles published one week after the announcement up to 18 months after the scandal (with 95% confidence error bars drawn on the graph). As the figure shows, even defining the event to be very narrow – articles published a week after the scandal – the three-day abnormal retail trading volume response to an article is lower after the scandal. The decline is sharpest when looking one month after, but is still significant even if we define the event from a year before to a year after the scandal. While the decline in trading activity a year after the scandal is likely due to both demand and supply effects from news on these platforms, the results only a week or two after the scandal seem much more likely to come from the demand side.[13]

5. Conclusion

We investigate a novel setting of fraud and market manipulation on social financial news networks. Using the announced SEC investigation as a shock to the revelation of fraud on these platforms, we analyze the broader effects of stock market manipulation on subsequent investor behavior. After the shock, we find significant spillover effects from investors discounting their response to all news from these platforms, including legitimate news, where trading activity declines significantly. We attempt to rule out alternative explanations for these results, including omitted variables that may drive changes in news, trading activity, and investor response to news. The evidence points to the SEC event itself as having a significant impact on the market's subsequent response and not other explanations or trends that happen to coincide with the event. These findings provide new evidence on the indirect effects of market manipulation and we interpret these results through the lens of models of trust and social capital.

Our findings may provide some of the first evidence of indirect spillover effects of fraud that match theory (Allcott and Gentzkow (2017), Aymanns et al. (2017), and Kshetri and Voas (2017)) and provide novel evidence for the role of social capital and trust (Guiso et al. (2004)). Our setting of financial markets may underestimate the broader impact of disinformation for settings where information costs are higher and the ability to take corrective action (i.e., through trading) is more limited.

[13]Panel B of Figure 4 repeats the plot for abnormal institutional trading volume that exhibits no reliable changes after the SEC event for any window length.

References

Aggarwal, R. and G. Wu (2006). Stock market manipulations. *The Journal of Business 79*(4), 1915–1953.

Allcott, H. and M. Gentzkow (2017). Social media and fake news in the 2016 election. *The Journal of Economic Perspectives 31*(2), 211–235.

Aymanns, C., J. Foerster, and C. Georg (2017). Fake news in social networks. *CoRR abs/1708.06233*.

Boehmer, E., C. M. Jones, X. Zhang, and X. Zhang (2020). Tracking retail investor activity. *Journal of Finance, Forthcoming*.

Bushee, B., M. Cedergren, and J. Michels (2020). Does the media help or hurt retail investors during the ipo quiet period? *Journal of Accounting and Economics 69*(1), 101261.

Chen, H., P. De, Y. J. Hu, and B.-H. Hwang (2014). Wisdom of crowds: The value of stock opinions transmitted through social media. *The Review of Financial Studies 27*(5), 1367–1403.

Cookson, J. A. and M. Niessner (2020). Why don't we agree? evidence from a social network of investors. *The Journal of Finance 75*(1), 173–228.

Delort, J.-Y., B. Arunasalam, M. Milosavljevic, and H. Leung (2009, 11). The impact of manipulation in internet stock message boards. *International Journal of Banking and Finance 8*.

Edelson, L., M.-K. Nguyen, I. Goldstein, O. Goga, D. McCoy, and T. Lauinger (2021). Understanding engagement with us (mis) information news sources on facebook. pp. 444–463.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance 25*(2), 383–417.

Frazzini, A., R. Israel, and T. J. Moskowitz (2018). Trading costs.

Frieder, L. and J. Zittrain (2007, 03). Spam works: Evidence from stock touts and corresponding market activity. *Social Science Research Network Working Paper Series*.

Gentzkow, M. and J. Shapiro (2005, October). Media bias and reputation. Working Paper 11664, National Bureau of Economic Research.

Gentzkow, M., J. Shapiro, and D. Stone (2015). Media bias in the marketplace: Theory. handbook of media economics, simon anderson, david stroemberg and joel waldfogel, eds.

Gomes, F., M. Haliassos, and T. Ramadorai (2021). Household finance. *Journal of Economic Literature Forthcoming*.

Guiso, L., P. Sapienza, and L. Zingales (2004, June). The role of social capital in financial development. *American Economic Review 94*(3), 526–556.

Guiso, L., P. Sapienza, and L. Zingales (2008). Trusting the stock market. *The Journal of Finance 63*(6), 2557–2600.

Guiso, L., P. Sapienza, and L. Zingales (2010, March). Civic capital as the missing link. Working Paper 15845, National Bureau of Economic Research.

Hanke, M. and F. Hauser (2008, 02). On the effects of stock spam e-mails. *Journal of Financial Markets 11*, 57–83.

Hu, B. and T. McInish (2013). Greed and fear in financial markets: The case of stock spam e-mails. *Journal of Behavioral Finance 14*(2), 83–93.

Hu, B., T. H. Mcinish, and L. Zeng (2009). The can-spam act of 2003 and stock spam emails.

Knack, S. and P. Keefer (1997). Does social capital have an economic payoff? a cross-country investigation. *The Quarterly Journal of Economics 112*(4), 1251–1288.

Kshetri, N. and J. Voas (2017, November). The economics of fake news. *IT Professional 19*(6), 8–12.

Loughran, T. and B. McDonald (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research 54*(4), 1187–1230.

Milgrom, P. and N. Stokey (1982). Information, trade and common knowledge. *Journal of Economic Theory 26*(1), 17–27.

Mullainathan, S. and A. Shleifer (2005, September). The market for news. *American Economic Review 95*(4), 1031–1053.

Nelson, K. K., R. A. Price, and B. R. Rountree (2013). Are individual investors influenced by the optimism and credibility of stock spam recommendations? *Journal of Business Finance & Accounting 40*(9-10), 1155–1183.

Newman, M. L., J. W. Pennebaker, D. S. Berry, and J. M. Richards (2003). Lying words: Predicting deception from linguistic styles. *Personality and social psychology bulletin 29*(5), 665–675.

Pennebaker, J., R. Booth, R. Boyd, and M. Francis (2015). Linguistic inquiry and word count: Liwc 2015 operator's manual. retrieved april 28, 2016.

Pennebaker, J. W. (2011). The secret life of pronouns: what our words say about us. Chapter 6. New York: Bloomsbury Press.

Sabherwal, S., S. K. Sarkar, and Y. Zhang (2011). Do internet stock message boards influence trading? evidence from heavily discussed stocks with no fundamental news. *Journal of Business Finance & Accounting 38*(910), 1209–1237.

Sapienza, P. and L. Zingales (2012). A trust crisis. *International Review of Finance 12*(2), 123–131.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance 62*(3), 1139–1168.

Ullah, S., N. Massoud, and B. Scholnick (2014, March). The Impact of Fraudulent False Information on Equity Values. *Journal of Business Ethics 120*(2), 219–235.

Vosoughi, S., D. Roy, and S. Aral (2018). The spread of true and false news online. *Science 359*(6380), 1146–1151.

Figure 1. **The Case that Launched the SEC Investigation: The Pump-and-Dump Scheme of Galena Biopharma Inc.**

This figure depicts the stock price of Galena Biopharma Inc. from April 2013 to May 2014, as well as occurrences of fake articles being published on Seeking Alpha, instances of SEO and stock options being granted to senior executives, as well as instances of insider trading and exposé articles about the promotional articles. This information was obtained from the SEC Lawsuit filed against Galena on 31 October, 2014 in the United States District Court (Case 3:14-cv-00558-SI). According to the lawsuit, the fake articles were published on August 6 and 22, 2013, September 26 and 30, 2013, November 12, 13, and 22, 2013, December 4, 10, 16, 2013, January 15, 2015, and February 5, 2014. While this was happening, Galena sold on September 18, 2013 in an SEO 17,500,000 units of stock for net proceeds to Galena of $32.6 million. On November 22, 2013, Galena held a board meeting and granted stock options to executives and directors with a strike price of $3.88. The CEO received 600,000 options, the CMO and COO 300,000 options, the CAO 150,000 options and each of the six directors received 200,000 options. Galena has historically awarded options either at the end of December or in early January. During the board meeting on January 16, 2014, where the board reviewed the preliminary 2013 earnings which had not been made public yet, the CEO declared that insiders could trade the company's stock immediately. Between January 17 and February 12, 2014 insiders sold over $16 million of their stock. On January 24 and 27, 2014 attention was drawn to the large insider trades. Then on February 1, 13, 14 and on March 13, 2014 articles started to appear on Seeking Alpha and TheStreet, documenting the promotional scheme. Finally on March 17, 2014, Galena disclosed in it's 10-K form the SEC investigation.

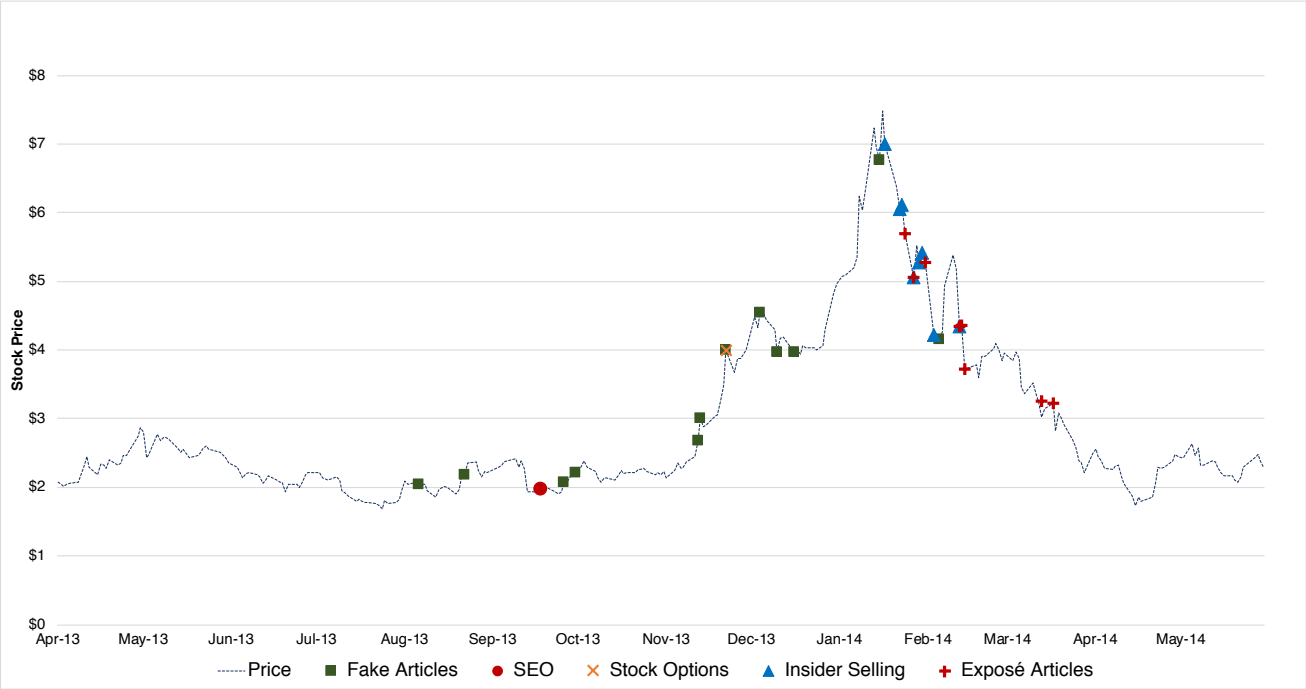


Figure 2. **Abnormal Returns for Fake Articles**

The figure depicts the difference in cumulative abnormal returns (residuals from a matched portfolio of stocks on size, BE/ME, and momentum) between days with fake articles and days with non-fake articles separately for small and non-small firms in our sample. The figure plots the returns for the for-sure fake articles from Rick Pearson and the SEC. Cumulative returns are measured starting with the day after the article's publication until 251 trading days after the article's publication. Before the article's publication, we measure cumulative returns starting with day -20 and ending on the day before publication ($t = -1$). Small firms are defined as firms in the bottom 10th percentile of NYSE firms.

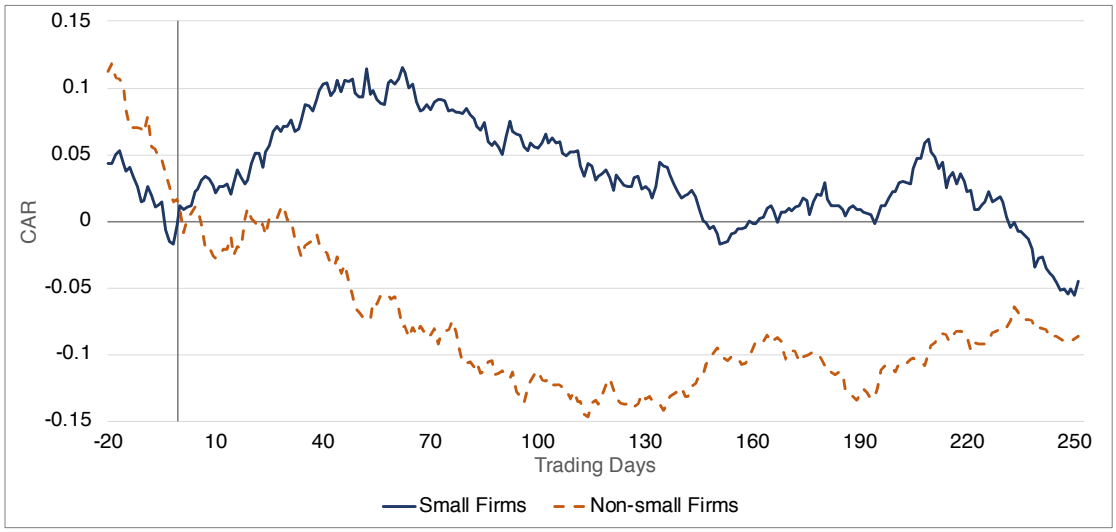


Figure 3. **Retail Trading Volume Response to Articles during the Six months *After* the 2014 SEC Investigation**

This figure plots the retail investors' reaction to articles published on the social media platforms during the six months *after* the SEC Lawsuit disclosure that publicly announced the presence of fake news on these platforms, and in the days around the article. We proxy for the market's reaction to the articles using daily abnormal retail trading volume. We plot the difference in reaction of abnormal retail volume around articles in the 6 months before versus 6 months after the disclosure period of the SEC announcement and press coverage (February-March 2014). We obtain retail trading volume from TAQ using Boehmer et al. (2020) method.
We estimate the following model for every day an article is published on Seeking Alpha and Motley Fool, and for days around the article publication:

$$Log(AbRetVol)_t = \alpha + \beta_1 Article \times \text{Post-Event} + \beta_2 Article + \beta_3 \text{Post-Event} + Controls + \epsilon$$

where $Log(AbRetVol)_t$ is defined as $RetVol(t)/AvgRetVol(t-146, t-20)$. *Post-Event* is a dummy variable equal to 1 if the article was published during the six month period after news of the SEC investigation broke, April 1 to September 30, 2014. The sample compares the period August 1, 2013 to January 31, 2014 to the period April 1 to 30 September, 2014. The figure plots the daily estimates of β_1 from seven days before the article's publication to four weeks after (trading days $t = -7, ..., 19$). Hence, $\beta_{1,t}$ represents the differential impact of the average retail trading volume response of an article after the SEC investigation relative to before the investigation on day t following the publication of an article on these social media platforms. The bars represent 90% confidence bands around the point estimates.
We also control for the presence of SEC filings, press releases, and print media coverage over the prior three days. SEC filing is a dummy variable if there was at least one SEC filling (10K, 10Q, or 8K) over the past three trading days, and press releases is a dummy variable if there was at least one press release issued by the firm over the past three trading days, and print media is a dummy if there was at least one WSJ or NYT article about the firm in the past three trading days.

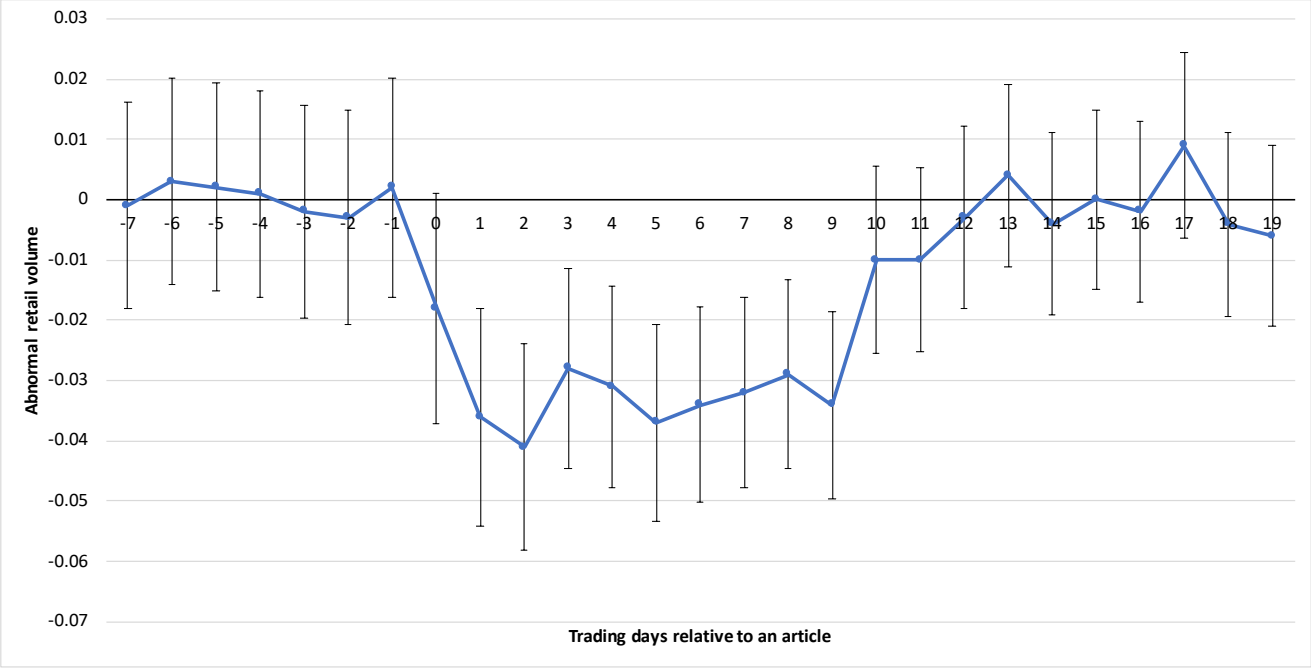


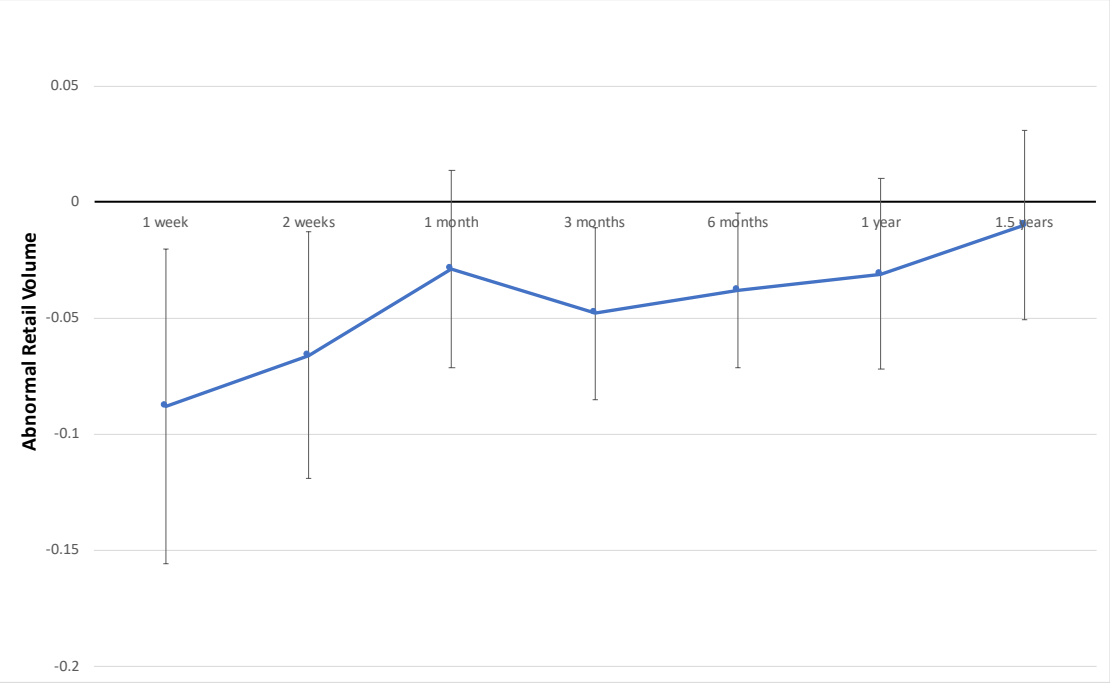
Figure 4. **Trading Volume Response to Articles over different Horizons *After* the 2014 SEC Investigation**

This figure plots the market's reaction to articles published on the social media platforms at various time bands *after* the SEC Lawsuit disclosure that publicly announced the presence of fake news on these platforms, and how that reaction changes over time after the scandal. We proxy for the market's reaction to the articles using abnormal daily trading volume. We plot the differences in reaction of abnormal volume to articles during 1 week, 2 weeks, 1 month, 3 months, 6 months, 1 year, and 1.5 years around the disclosure period of the SEC announcement and press coverage (February-March 2014). In Panel A we examine retail trading volume obtained from TAQ using Boehmer et al. (2020) method, and in Panel B we focus on institutional trading, proxied for by trades greater than $50,000.
We estimate the following model for the day of, the day after, and two days after an article is published on Seeking Alpha and Motley Fool:

$$Log(AbVol)_t = \alpha + \beta_1 Article \times \text{Post-Event} + \beta_2 Article + \beta_3 \text{Post-Event} + Controls + \epsilon$$

where $Log(AbVol)_t$ is defined as $Vol(t)/AvgVol(t-146, t-20)$. *Post-Event* is a dummy variable equal to 1 if the article was published during a given time band (1 week, 2 weeks, 1 month, 3 months, 6 months, 1 year, and 1.5 years) after news of the SEC investigation broke, April 1 to September 30, 2014. The figure plots the estimates of β_1 summed over days $t = 0$, $t+1$, and $t+2$, which represent the cumulative effect of the scandal on the abnormal trading volume reaction over the given time period. We estimate the regression for 1 week, 2 weeks, 1 month, 3 months, 6 months, 1 year, and 1.5 years after the scandal (i.e., after April 1, 2014). The bars represent 90% confidence bands around the point estimates.
We also control for the presence of SEC filings, press releases, and print media coverage over the prior three days. SEC filing is a dummy variable if there was at least one SEC filling (10K, 10Q, or 8K) over the past three trading days, and press releases is a dummy variable if there was at least one press release issued by the firm over the past three trading days, and print media is a dummy if there was at least one WSJ or NYT article about the firm in the past three trading days.

Panel A: Retail Trading Volume



Panel B: Institutional Trading Volume

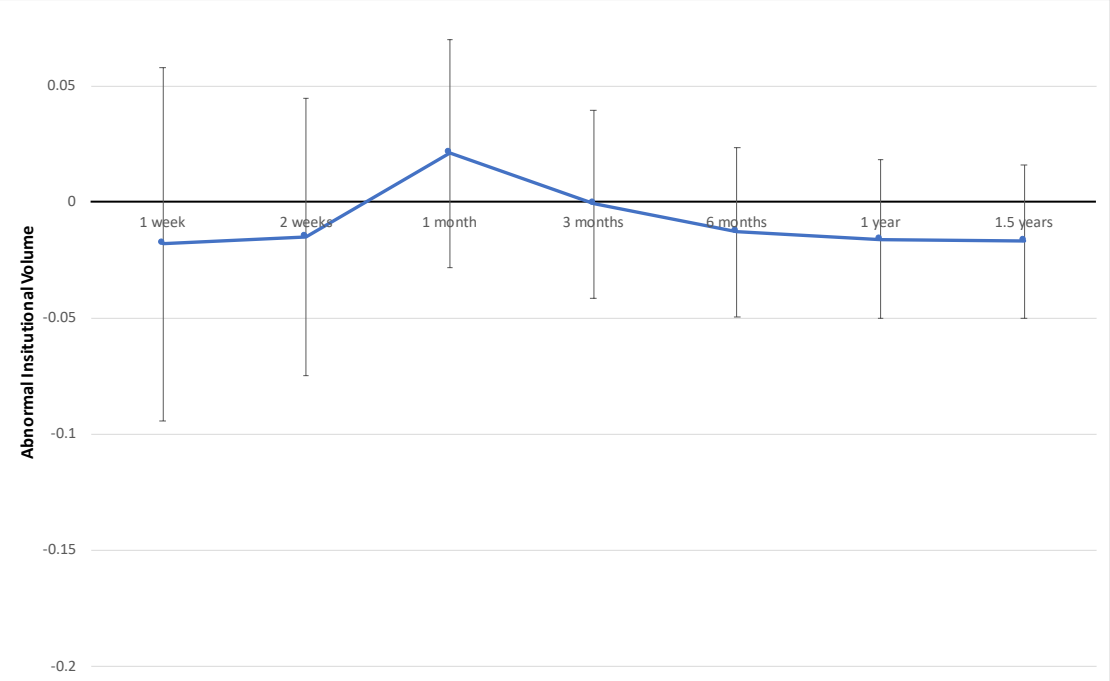


Table 1. **Relationship Between Articles and Subsequent Retail Trading Volume**

Reported are retail trading volume response to articles posted on Seeking Alpha and Motley Fool. We report results from regressions of the log of abnormal retail trading volume on the three days following the publication of an article on these platforms. We obtain retail trading volume from TAQ using Boehmer et al. (2020) method. Abnormal retail volume is defined as the log of $RetVol(t)/AvgRetVol(t-146, t-20)$, summed over days $t=0$, $t+1$, and $t+2$. We examine all firms that have ever had an article written about them on Seeking Alpha or Motley Fool. The main independent variable is *Article*, a dummy variable equal to 1 if there was at least one article published about the firm on day $t=0$. *Small firm* equals 1 if the firm is in the bottom 10th percentile of NYSE firms, and 0 otherwise. Firm size is measured in the prior trading month. In addition, we include as controls *SEC Filing*$_{t-3,t}$, which is a dummy variable if there was at least one SEC filling (10K, 10Q, or 8K) over the past three trading days and *Press release*$_{t-3,t}$, which is a dummy variable if there was at least one press release issued by the firm over the past three trading days. *Print Media*$_{t-3,t}$ is a dummy variable if there was at least one WSJ or NYT article about the firm in the past three trading days. We also control for lagged abnormal retail trading volume on day $t-1$. We include year-month fixed effects, and indicate statistical significance at the ten, five, and one percent levels with *, **, and ***, respectively. Standard errors are clustered at the firm level, with t-statistics in parentheses.

Dependent variable =	Log(Abnormal daily retail volume$_{t,t+2}$)			
Article	0.365***	0.195***	0.161***	0.194***
	(42.96)	(34.11)	(29.78)	(34.07)
Small firm			-0.028***	
			(-11.48)	
Article × Small firm			0.313***	
			(19.73)	
Fraudulent article				0.554**
				(2.22)
SEC filing$_{t-3,t}$		0.048***	0.047***	0.048***
		(26.60)	(26.13)	(26.60)
Press release$_{t-3,t}$		0.114***	0.111***	0.114***
		(51.70)	(50.37)	(51.72)
Print media$_{t-3,t}$		0.023***	0.020***	0.023***
		(4.81)	(4.39)	(4.82)
Abnormal retail volume$_{t-1}$		1.220***	1.219***	1.220***
		(174.78)	(174.76)	(174.66)
Observations	9,796,231	9,789,124	9,789,124	9,789,124
R-squared	0.023	0.284	0.284	0.284
Year-month F.E.	Y	Y	Y	Y

Table 2. **The Effect of 2014 SEC Lawsuit on Trading Volume and Return Volatility**

The table examines whether the salience of the presence of fake news on the platforms, stemming from the public announcement of the SEC investigation and lawsuit, impacted investors' reaction to articles on these platforms. In Panel A we examine retail trading volume obtained from TAQ using Boehmer et al. (2020) method. In Panel B, we examine the effect on return volatility, measured over the three days following the publication of an article on these platforms, multiplied by 100. In Panel C we focus on institutional trading, proxied for by trades greater than $50,000, over the three days following the publication of an article on these platforms. We compare the volume response to articles in the six months before the SEC investigation and six months after the investigation, where we identify the February-March 2014 period as the period when the SEC investigation was announced and covered in the press. We include all firms that have ever had at least one article written about them on Seeking Alpha or Motley Fool during that time period. Abnormal volume are defined as in Table 1. The regressors include the dummy variable *Article*, which equals 1 if there was at least one article published about the firm on day t, and 0 otherwise, and we include the dummy variable *Post-event*, which equals 1 if the time period is April 1 to September 30, 2014 and is zero if the article was published from August 1, 2013 to January 31, 2014. We exclude all observations prior to August 2013 and after September 2014. We then interact the *Post-event* dummy with the *Article* dummy in the regressions to test for the differential response to articles before versus after the SEC announced investigation. *Small firm* equals 1 if the firm is in the bottom 10th percentile of NYSE firms, and 0 otherwise. Firm size is measured in the prior trading month. We also include controls for SEC filings, press releases, and abnormal volume over the previous three days before the article, plus the day of the article. *SEC filing*$_{t-3,t}$ is a dummy variable if there was at least one SEC filling (10K, 10Q, or 8K) over the past three trading days, and *Press release*$_{t-3,t}$ is a dummy variable if there was at least one press release issued by the firm over the past three trading days. *Print media*$_{t-3,t}$ is a dummy variable if there was at least one WSJ or NYT article about the firm in the past three trading days. In Panel D we control for any changes in the reaction of investors to news sources unrelated to the scandal on the social media platforms (SEC filings, Press releases, and articles in the WSJ and NYT) after the investigation announcement. We indicate statistical significance at the ten, five, and one percent levels with *, **, and ***, respectively (t-statistics in parentheses). Standard errors are clustered at the firm level.

Panel A: Effect on Abnormal Retail Volume

Dependent variable =	Log(Abnormal daily retail volume$_{t,t+2}$)		
Article	0.201***	0.239***	0.188***
	(19.12)	(15.03)	(13.50)
Post		-0.095***	-0.029***
		(-17.10)	(-5.56)
Article × Post		-0.038*	-0.041**
		(-1.88)	(-2.03)
Small firm			0.114***
			(12.79)
Article × Small firm			0.643***
			(10.00)
Post × Small firm			-0.174***
			(-13.91)
Article × Post × Small firm			-0.235***
			(-2.94)
SEC filing$_{t-3,t}$	0.052***	0.057***	0.058***
	(12.89)	(13.99)	(14.30)
Press release$_{t-3,t}$	0.124***	0.122***	0.128***
	(25.51)	(25.15)	(26.74)
Print media$_{t-3,t}$	-0.009	-0.011	0.009
	(-0.83)	(-1.05)	(0.81)
Abnormal retail volume$_{t-1}$	1.239***	1.232***	1.224***
	(88.63)	(88.81)	(89.47)
Observations	1,314,455	1,314,455	1,314,455
R-squared	0.275	0.277	0.280

Panel B: Effect on Return Volatility

Dependent variable =	Return volatility$_{t,t+2}$		
Article	0.138***	0.195***	0.073***
	(4.02)	(4.32)	(4.44)
Post		-0.055***	-0.014
		(-4.90)	(-1.41)
Article × Post		-0.076	0.056
		(-1.33)	(1.29)
Small Firm			0.362***
			(13.81)
Article × Small Firm			1.809***
			(4.59)
Post × Small Firm			-0.113***
			(-3.87)
Article× Post×Small Firm			-1.282***
			(-3.03)
SEC filing$_{t-3,t}$	0.032***	0.035***	0.038***
	(2.59)	(2.83)	(3.13)
Press release$_{t-3,t}$	0.043***	0.042***	0.077***
	(3.02)	(2.94)	(5.52)
Print media$_{t-3,t}$	-0.172***	-0.174***	-0.074***
	(-11.72)	(-11.87)	(-5.71)
Return Volatility$_{t-1}$	4.664***	4.654***	4.379***
	(3.24)	(3.24)	(3.20)
Observations	1,005,279	1,005,279	1,005,279
R-squared	0.002	0.002	0.004

Panel C: Effect on Abnormal Institutional Volume

Dependent variable =	Log(Abnormal daily institutional volume$_{t,t+2}$)		
Article	0.369***	0.399***	0.242***
	(30.03)	(20.69)	(14.95)
Post		-0.153***	-0.129***
		(-23.61)	(-18.48)
Article × Post		-0.013	0.012
		(-0.59)	(0.60)
Small firm			-0.422***
			(-31.61)
Article × Small firm			0.837***
			(8.47)
Post × Small firm			-0.072***
			(-4.90)
Article × Post × Small firm			-0.005
			(-0.93)
SEC filing$_{t-3,t}$	0.062***	0.069***	0.050***
	(9.83)	(10.96)	(8.57)
Press release$_{t-3,t}$	0.228***	0.226***	0.174***
	(32.92)	(32.51)	(26.51)
Print media$_{t-3,t}$	0.227***	0.224***	0.127***
	(14.43)	(13.85)	(8.08)
Abnormal institutional volume$_{t-1}$	0.862***	0.853***	0.789***
	(82.42)	(81.79)	(69.41)
Observations	1,283,716	1,283,716	1,283,716
R-squared	0.163	0.166	0.194

Panel D: Effect from Unrelated News Sources

Dependent Variable =	Log(Ab Retail Vol)$_{t,t+2}$	Log(Ab Inst Vol)$_{t,t+2}$
Article	0.248***	0.408***
	(16.06)	(21.48)
Post	-0.097***	-0.147***
	(-17.12)	(-22.01)
Article × Post	-0.056***	-0.029
	(-2.97)	(-1.35)
SEC filing$_{t-3,t}$	0.062***	0.088***
	(9.34)	(9.71)
SEC filing$_{t-3,t}$ × Post	-0.010	-0.035
	(-1.20)	(-0.07)
Press release$_{t-3,t}$	0.111***	0.244***
	(15.72)	(25.88)
Press release$_{t-3,t}$ × Post	0.023	-0.05
	(0.48)	(-0.92)
Print media$_{t-3,t}$	-0.052***	0.159***
	(-4.18)	(8.91)
Print media$_{t-3,t}$ × Post	0.08	0.026
	(0.70)	(1.03)
Abnormal retail volume$_{t-1}$	1.232***	
	(88.83)	
Abnormal institutional volume$_{t-1}$		0.853***
		(81.79)
Observations	1,314,455	1,283,716
R-squared	0.277	0.167

Table 3. **Disagreement around Articles**

The table examines whether disagreement among retail investors as a response to news released has changed after the the SEC announced investigation. We measure disagreement among mostly retail investors using data from a social media platform called StockTwits. The details of the disagreement measure construction are described in Cookson and Niessner (2020)). The disagreement measure is standardized by subtracting the mean and dividing by the standard deviation. We compare disagreement among investors in response to Seeking Alpha articles six months before versus six months after the SEC investigation by regressing a measure of disagreement among mostly retail investors on *Article*, which equals 1 if there was at least one article published about the firm on day t, and 0 otherwise, and we include the dummy variable *Post-event*, which equals 1 if the time period is April 1 to September 30, 2014 and is zero if the article was published from August 1, 2013 to January 31, 2014. We exclude all observations prior to August 2013 and after September 2014. We then interact the with the *Article* dummy in the regressions to test for the differential response to articles before versus after the SEC announced investigation. We also control for the number of messages posted about the firm on StockTwits, and interact that number of messages with the *Post-event* dummy. In columns (1) and (2) we examine all firms, and in columns (3) and (4) we look at small and non-small firms, respectively. *Small firm* equals 1 if the firm is in the bottom 10th percentile of NYSE firms, and 0 otherwise. Firm size is measured in the prior trading month.We also include controls for SEC filings, press releases, and abnormal volume over the previous three days before the article, plus the day of the article. *SEC filing*$_{t-3,t}$ is a dummy variable if there was at least one SEC filling (10K, 10Q, or 8K) over the past three trading days, and *Press release*$_{t-3,t}$ is a dummy variable if there was at least one press release issued by the firm over the past three trading days. *Print media*$_{t-3,t}$ is a dummy variable if there was at least one WSJ or NYT article about the firm in the past three trading days. All regressions include firm fixed effects and statistical significance is denoted at the ten, five, and one percent levels by *, **, and ***, respectively (t-statistics in parentheses).

Dependent variable =	Disagreement$_t$			
	All firms		Small firms	Non-small firms
Article	0.292***	0.301***	0.292***	0.297***
	(16.02)	(11.33)	(6.14)	(10.48)
Post-event		0.014**	0.013**	0.009
		(2.31)	(2.01)	(1.16)
Article × Post-event		-0.033	-0.000	-0.026
		(-0.89)	(-0.01)	(-0.67)
Num messages	0.004***	0.004***	0.006***	0.003***
	(3.37)	(3.46)	(2.70)	(2.89)
Num messages × Post-event		0.003***	0.008***	0.002***
		(3.16)	(2.71)	(2.70)
SEC filing$_{t-3,t}$	0.138***	0.136***	0.079***	0.168***
	(12.94)	(12.95)	(9.83)	(12.71)
Press release$_{t-3,t}$	0.228***	0.227***	0.182***	0.246***
	(20.41)	(20.57)	(15.04)	(19.95)
Print media$_{t-3,t}$	0.105***	0.104***	0.092	0.106***
	(7.75)	(7.81)	(1.37)	(7.89)
Disagreement$_{t-1}$	0.154***	0.153***	0.191***	0.133***
	(40.37)	(40.27)	(26.71)	(36.32)
Observations	1,406,229	1,406,229	553,677	852,552
R-squared	0.323	0.324	0.304	0.330
Firm FE	X	X	X	X

Table 4. **Retail Volume Impact Across Article and Author Characteristics**

The table reports the relationship between author and article characteristics and the abnormal retail trading volume response to articles published on the social media platforms. Abnormal retail trading volume is defined in Table 1. Panel A reports results for the entire sample period and Panel B focuses on the time period six months before and six months after the SEC announced investigation (February/March 2014). The independent variables include the log of the number of followers the author of the article has, the log of one plus the number of articles that the author has written prior to the article, the log of one plus the number of comments that the article received, the standardized fraction or % of how often the author uses numbers in the article, and the standardized fraction or % of how often the word "earning" or the stem "earn" are used in the article. In addition, we include the number of comments, the average length of comments (word count) on the article, and dummy variables for whether disagreement words ("disagree", "differ", "counter", "oppose", or "argue"), "wrong" words ("wrong" or "not right"), or "fake" words ("fake" or "fraud") are used in the comments section, and Disclosed position is 1 if the author disclosed an existing position, and zero otherwise. Panel B interacts the *Post-Event* dummy, which equals 1 if the time period is April 1 to September 30, 2014, the six month period after the SEC announced investigation, with each of the independent variables. We also include controls for SEC filings, press releases, and print media coverage over the prior three days, and abnormal trading volume on day $t-1$. SEC filing is a dummy variable if there was at least one SEC filling (10K, 10Q, or 8K) over the past three trading days, and press releases is a dummy variable if there was at least one press release issued by the firm over the past three trading days, and print media is a dummy if there was at least one WSJ or NYT article about the firm in the past three trading days. We do not report the coefficient estimates on these controls for brevity and because their estimates are very similar to those in Table 2. We include fixed effects at the year-month, author, and firm level (where appropriate), and statistical significance is denoted at the ten, five, and one percent levels by *, **, and ***, respectively (t-statistics in parentheses). Standard errors are clustered at the firm level.

Panel A: Full Sample Period

Dependent variable =	Log(Abnormal daily retail volume$_{t,t+2}$)									
Log(Num followers)	0.013***									
	(8.96)									
Log(Num past articles)		0.005***								
		(3.29)								
% Numbers in text			0.026***							
			(9.07)							
% "Earn" mentions				0.045***						
				(12.65)						
log(Num comments)					0.054***					
					(10.84)					
Log(avg comment length)						0.023***				
						(6.68)				
Disagree comments							0.024***			
							(3.25)			
Wrong comments								0.047***		
								(5.16)		
Fake/Fraud comments									0.034**	
									(2.06)	
Discloses position										-0.006
										(-0.69)
Observations	174,377	320,228	317,443	317,443	172,002	123,005	172,002	172,002	172,002	171,525
R-squared	0.494	0.444	0.484	0.485	0.536	0.562	0.534	0.535	0.534	0.534
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year-Month F.E.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Author F.E.			Y	Y	Y	Y	Y	Y	Y	Y
Firm F.E.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Panel B: Comparing 6-months before versus 6-months after the SEC investigation

Dependent variable =	Log(Abnormal daily retail volume$_{t,t+2}$)									
Log(Num followers)	-0.006									
	(-1.05)									
Log(Num followers) x Post-event	0.016**									
	(2.20)									
Log(Num past articles)		-0.007								
		(-1.32)								
Log(Num past articles) x Post-event		0.051***								
		(7.41)								
% Numbers in text			-0.002							
			(-0.26)							
% Numbers in text x Post-event			0.003							
			(0.17)							
%“Earn" mentions				0.002						
				(0.15)						
% “Earn" mentions x Post-event				0.050**						
				(2.25)						
log(Num comments)					0.037***					
					(2.60)					
log(Num comments) x Post-event					0.004					
					(0.13)					
Log(avg comment length)						0.034***				
						(2.73)				
Log(avg comment length) x Post-event						-0.009				
						(-0.33)				
Disagree comments							0.048			
							(1.61)			
Disagree comments x Post-event							0.005			
							(0.07)			
Wrong comments								0.078**		
								(2.22)		
Wrong comments x Post-event								-0.010		
								(-0.14)		
Fake/Fraud comments									0.206**	
									(2.46)	
Fake/Fraud comments x Post-event									-0.125	
									(-1.39)	
Discloses position										0.146***
										(5.92)
Discloses position x Post-event										-0.103***
										(-3.10)
Post-event	-0.193***	-0.281***	-0.048	-0.042	-0.092**	-0.018	-0.085***	-0.081***	-0.077*	-0.034
	(-2.65)	(-3.42)	(-0.69)	(-0.61)	(-2.00)	(-0.13)	(-3.70)	(-2.78)	(-1.83)	(-0.85)
Observations	28,392	55,258	55,257	55,257	28,392	25,595	28,392	28,392	28,392	28,229
R-squared	0.395	0.394	0.390	0.391	0.396	0.410	0.395	0.395	0.395	0.392
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Table 5. **Response of Authors and Readers to the 2014 SEC Investigation**

The table reports regression results of author and reader responses to the SEC announced investigation by looking at the characteristics of articles and comments on those articles in the six month period after the SEC investigation. Specifically, we regress a host of article and comment characteristics on the dummy variable *Post-event*, which equals 1 if an article was published from April 1 to September 30, 2014 and zero if the article was published from August 1, 2013 to January 31, 2014. These periods correspond to six months after and six months before the February-March 2014 SEC investigation period. The dependent variables are the log of one plus the number of comments that the article received, the standardized fraction or % of how often the author uses numbers in the article, the standardized fraction or % of how often the word "earning" or the stem "earn" are used in the article, the average length of comments (word count) on the article and dummy variables for whether the disagreement words ("disagree", "differ", "counter", "oppose", or "argue"), "wrong" words ("wrong" or "not right"), or "fake" words ("fake" or "fraud") are used in the comments section. Each of these dependent variables are regressed on the *Post-event* dummy plus controls for SEC filings, press releases, and print media coverage over the prior three days. SEC filing is a dummy variable if there was at least one SEC filling (10K, 10Q, or 8K) over the past three trading days, and press releases is a dummy variable if there was at least one press release issued by the firm over the past three trading days, and print media is a dummy if there was at least one WSJ or NYT article about the firm in the past three trading days. We also controls for lagged abnormal volume from the previous day, defined in Table 1. Panel A reports results for articles written about small firms only (smallest 10% of stocks based on NYSE breakpoints) and Panel B reports results for all other firms (non-small firms). Panels A and B include firm fixed effects. In Panels C and D, we examine whether author characteristics have changed after the SEC investigation. The dependent variables are the log number of followers that the authors have and the log number of articles that the author has written in the past. Statistical significance is denoted at the ten, five, and one percent levels by *, **, and ***, respectively (t-statistics in parentheses).

Dependent variable =	# Comments	% Numbers in text	% "Earn" in text	Length of comments	"Disagree" words in comments	"Wrong" words in comments	"Fake" words in comments
Panel A: Articles about small firms							
Post-event	0.083	-0.051	0.123***	-0.101	-0.003	0.020	0.006**
	(1.46)	(-1.13)	(4.71)	(-1.43)	(-0.11)	(0.84)	(2.10)
SEC filing$_{t-3,t}$	-0.097	-0.018	0.037	-0.047	-0.041	-0.026	-0.027
	(-1.48)	(-0.35)	(1.26)	(-0.57)	(-1.42)	(-0.95)	(-1.60)
Press release$_{t-3,t}$	-0.098	-0.015	0.136***	-0.147	-0.019	-0.014	0.002
	(-1.29)	(-0.28)	(4.26)	(-1.58)	(-0.58)	(-0.46)	(0.09)
Print media$_{t-3,t}$	-0.266	0.063	0.009	-0.387*	0.029	-0.057	-0.029
	(-1.47)	(0.44)	(0.11)	(-1.79)	(0.37)	(-0.77)	(-0.63)
Abnormal volume$_{t-1}$	0.097***	-0.007	0.006	0.076***	0.019*	0.031***	0.016***
	(4.29)	(-0.37)	(0.53)	(2.74)	(1.90)	(3.32)	(2.71)
Observations	2,257	2,845	2,845	2,257	2,257	2,257	2,257
R^2	0.606	0.400	0.445	0.592	0.422	0.432	0.517
Firm F.E.	Y	Y	Y	Y	Y	Y	Y
Panel B: Articles about non-small firms							
Post-event	-0.060***	-0.178***	-0.057***	-0.069***	-0.017***	-0.006	0.004
	(-4.86)	(-1.22)	(-6.26)	(-4.21)	(-3.00)	(-1.16)	(1.45)
SEC filing$_{t-3,t}$	-0.034**	0.040***	0.169***	-0.054***	-0.022***	-0.007	-0.006
	(-2.31)	(3.91)	(17.80)	(-2.74)	(-3.14)	(-1.12)	(-1.49)
Press release$_{t-3,t}$	-0.077***	0.017	0.142***	-0.097***	-0.018**	-0.021***	-0.004
	(-4.91)	(1.62)	(14.25)	(-4.63)	(-2.45)	(-2.97)	(-1.10)
Print media$_{t-3,t}$	0.022	-0.007	0.029**	0.023	0.015*	-0.000	0.003
	(1.24)	(-0.60)	(2.57)	(0.97)	(1.77)	(-0.06)	(0.57)
Abnormal volume$_{t-1}$	0.029***	0.003	0.067***	0.032**	0.006	0.009**	0.004
	(2.84)	(0.40)	(9.77)	(2.40)	(1.20)	(1.99)	(1.46)
Observations	25,635	52,273	52,273	25,635	25,635	25,635	25,635
R^2	0.596	0.148	0.185	0.515	0.312	0.288	0.253
Firm F.E.	Y	Y	Y	Y	Y	Y	Y

Dependent variable =	log(Num followers)	log(Num past articles)
Panel C: Articles about small firms		
Post-event	0.380***	0.908***
	(4.79)	(12.51)
Observations	2,866	3,427
R-squared	0.008	0.044
Panel D: Articles about non-small firms		
Post-event	-0.168***	0.335***
	(-7.58)	(17.80)
Observations	26,413	52,891
R-squared	0.002	0.006

Table 6. **Linguistic Analysis of Comments and Articles**

The table reports regression results of reader and author responses to the SEC announced investigation by looking at the linguistic characteristics of comments on the articles and the articles themselves using the linguistic algorithm LIWC. We compare linguistic characteristics six months before versus six months after the SEC investigation by regressing a host of article and comment characteristics on the dummy variable *Post-event*, which equals 1 if an article was published from April 1 to September 30, 2014 and zero if the article was published from August 1, 2013 to January 31, 2014. The dependent variables are LIWC measures of Analytic, Clout, and Tone of the comments on the articles (first three columns) and the measures of Analytic, Clout, and Tone on the articles themselves (next three columns) plus the LIWC measure of authenticity of the articles (last column). Analytic is how analytical the language sounds, Clout is how authoritative the language sounds, and Tone is the sentiment of the article, where higher scores mean more positive tone. Authenticity measures how the article is using authentic linguistic cues, where a positive number indicates more authenticity. Each of these dependent variables are regressed on the *Post-event* dummy plus controls for SEC filings, press releases, and print media coverage over the prior three days. SEC filing is a dummy variable if there was at least one SEC filling (10K, 10Q, or 8K) over the past three trading days, and press releases is a dummy variable if there was at least one press release issued by the firm over the past three trading days, and print media is a dummy if there was at least one WSJ or NYT article about the firm in the past three trading days. We also controls for lagged abnormal volume from the previous day, defined in Table 1. Panel A reports results for articles written about small firms only (smallest 10% of stocks based on NYSE breakpoints) and Panel B reports results for all other firms (non-small firms). All regressions include firm fixed effects and statistical significance is denoted at the ten, five, and one percent levels by *, **, and ***, respectively (t-statistics in parentheses).

Panel A: Articles about small firms

	Comments			Articles			
Dependent variable =	Analytic	Clout	Tone	Analytical	Clout	Tone	Authenticity
Post-event	0.053	-0.300	-4.345***	0.333	0.569	0.550	3.079***
	(0.07)	(-0.44)	(-3.70)	(0.25)	(1.43)	(0.64)	(5.50)
SEC filing$_{t-3,t}$	0.015	-0.716	1.490	-0.600**	0.265	1.676*	-0.181
	(0.02)	(-0.91)	(1.10)	(-2.02)	(0.52)	(1.74)	(-0.26)
Press release$_{t-3,t}$	0.680	0.391	1.746	-0.581*	1.414**	-0.093	2.520***
	(0.73)	(0.44)	(1.13)	(-1.79)	(2.53)	(-0.09)	(3.35)
Print media$_{t-3,t}$	-4.007*	1.472	-5.695	-0.551	-1.366	1.524	-3.823**
	(-1.85)	(0.71)	(-1.58)	(-0.65)	(-0.94)	(0.55)	(-2.31)
Abnormal volume$_{t-1}$	-0.506*	0.114	-0.478	0.089	0.182	-0.137	0.175
	(-1.80)	(0.42)	(-1.03)	(0.84)	(1.00)	(-0.40)	(0.79)
Observations	2,257	2,257	2,257	2,845	2,845	2,845	3,421
R^2	0.312	0.333	0.384	0.347	0.399	0.424	0.014
Firm F.E.	Y	Y	Y	Y	Y	Y	Y

Panel B: Articles about non-small firms

	Comments			Articles			
Dependent variable =	Analytic	Clout	Tone	Analytical	Clout	Tone	Authenticity
Post-event	0.202	-0.406**	-1.232***	-0.019	0.007	0.652	-1.683***
	(1.24)	(-2.45)	(-4.29)	(-0.34)	-0.23	(1.43)	(-4.02)
SEC filing$_{t-3,t}$	0.303	-0.105	0.119	0.236***	0.396***	0.686***	1.615***
	(1.55)	(-0.53)	(0.35)	(3.95)	(4.33)	(3.46)	(10.26)
Press release$_{t-3,t}$	-0.098	0.087	0.434	0.162***	0.587***	0.566***	1.095***
	(-0.47)	(0.41)	(1.19)	(2.59)	(6.14)	(2.73)	(7.08)
Print media$_{t-3,t}$	-0.245	0.340	0.096	-0.190***	-0.299***	-0.159	-1.546***
	(-1.06)	(1.45)	(0.24)	(-2.71)	(-2.79)	(-0.69)	(-10.03)
Abnormal volume$_{t-1}$	-0.163	-0.217	-0.460*	0.020	0.076	-0.372***	0.980***
	(-1.21)	(-1.59)	(-1.94)	(0.47)	(1.17)	(-2.62)	(9.06)
Observations	25,635	25,635	25,635	52,273	52,273	52,273	52,871
R^2	0.151	0.169	0.191	0.122	0.180	0.157	0.010
Firm F.E.	Y	Y	Y	Y	Y	Y	Y

APPENDIX

Appendix A: Contributors and compensation for authorship on shared-knowledge platforms

For authors on Seeking Alpha, base payment is $35 plus $10 per 1,000 page-views. For analysis of stocks that have a large number of followers, Seeking Alpha has three additional payment tiers, from $150 to $500 per article. Finally, two articles are selected each week for a $2,500 "outstanding performance" prize on the basis of how well the stock idea played out. The articles are published as Premium articles, Standard articles, and Instablogs. Standard articles are allowed to be published elsewhere, and are unpaid, but also undergo a selection process. Instablogs are published instantly and with no pay.

The Motley Fool offers a wide range of stock news and analysis at its free website, www.fool.com, as well as through a variety of paid investment advice services, which provide online stock analysis and research with interactive discussion boards. The discussion boards are used heavily to recruit future Motley Fool staffers, where frequent posters are first awarded free subscriptions and then can receive a small stipend. The Motley Fool Blog Network was a stock analysis and news site that provided a platform for non-Motley Fool staff writers to submit articles. They received compensation ranging from $50 to $100 for each article submitted and additional compensation for how many recommendations or “editors picks” they received. Eventually the company merged the Blog Network with its primary site in 2014.

Appendix B: Supplemental Tables and Figures for "Social Media and Financial News Manipulation"

Figure B1. **Time Trend in Trading Volume**

This figure plots the trading volume in the year before and the year after the SEC Lawsuit disclosure that publicly announced the presence of fake news on these platforms (February and March 2014). In Panel A we examine retail trading volume obtained from TAQ using Boehmer et al. (2020) method, and in Panel B we focus on institutional trading, proxied for by trades greater than $50,000, over the three days following the publication of an article on these platforms. In grey we highlight the months in which news about the presence of fake articles on these platforms was disclosed.

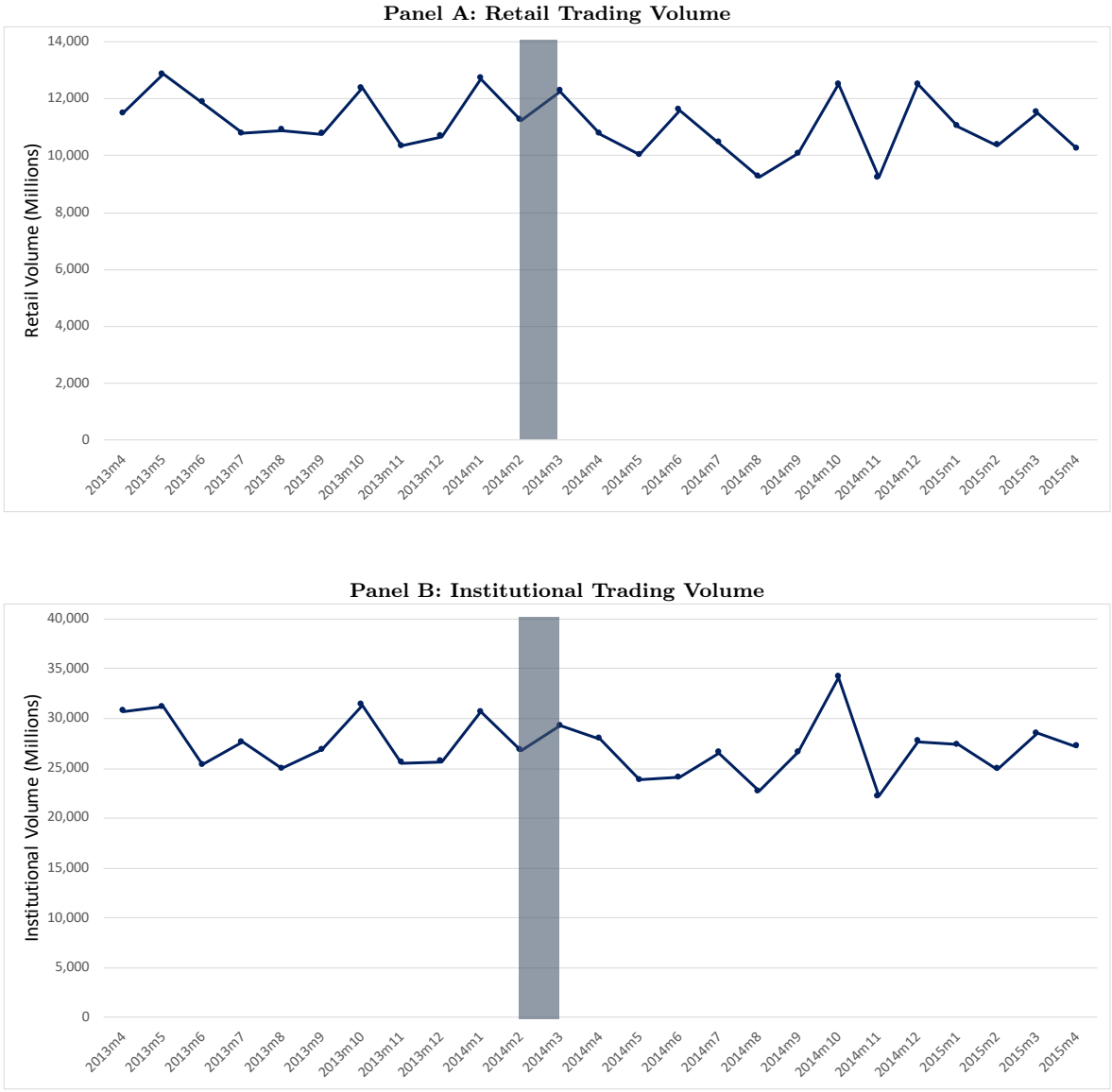


Table B1. **Effect of articles on retail trading volume - Daily Reaction**

Reported are daily retail trading volume responses to articles posted on Seeking Alpha and Motley Fool over the first three days after the article is published. The table reports results from regressions of the log of abnormal trading volume on the three days following the publication of an article on these platforms on the dummy variable *Article*, which equals 1 if there was at least one article published about the firm on day $t = 0$. In addition, we include as controls $SEC\ Filing_{t-3,t}$, which is a dummy variable if there was at least one SEC filling (10K, 10Q, or 8K) over the past three trading days. $Press\ release_{t-3,t}$, which is a dummy variable if there was at least one press release issued by the firm over the past three trading days. $Print\ media_{t-3,t}$ is a dummy variable if there was at least one WSJ or NYT article about the firm in the past three trading days. We obtain retail trading volume from TAQ using Boehmer et al. (2020) method. Abnormal retail volume is defined as the log of $RetVol(t)/AvgRetVol(t-146, t-20)$, summed over days $t = 0$, $t+1$, and $t+2$. We also control for lagged abnormal trading volume on day $t-1$. We include year-month fixed effects, and indicate statistical significance at the ten, five, and one percent levels with *, **, and ***, respectively (t-statistics in parentheses). Standard errors are clustered at the firm level.

Dependent variable =	Log(Abnormal retail trading volume)		
	Day 0 ($t = 0$)	Day 1 ($t + 1$)	Day 2 ($t + 2$)
Article	0.088***	0.062***	0.045***
	(32.26)	(32.89)	(31.30)
SEC filing$_{t-3,t}$	0.028***	0.017***	0.003***
	(47.49)	(25.75)	(4.25)
Press release$_{t-3,t}$	0.050***	0.043***	0.021***
	(63.94)	(50.37)	(28.19)
Print media$_{t-3,t}$	0.009***	0.006***	0.008***
	(4.80)	(3.40)	(5.47)
Abnormal retail volume$_{t-1}$	0.468***	0.394***	0.358***
	(193.75)	(169.08)	(156.71)
Observations	9,789,254	9,783,738	9,778,195
R^2	0.237	0.171	0.141
Year-month F.E.	Y	Y	Y

Table B2. **Relationship Between Articles and Subsequent Price Volatility**

Reported are results from the same regressions in Table 1, but using three-day squared price movements (volatility), multiplied by 100, as the dependent variable. We examine all firms that have ever had an article written about them on Seeking Alpha or Motley Fool. We examine all firms that have ever had an article written about them on Seeking Alpha or Motley Fool. The main independent variable is *Article*, a dummy variable equal to 1 if there was at least one article published about the firm on day $t = 0$. *Small firm* equals 1 if the firm is in the bottom 10th percentile of NYSE firms, and 0 otherwise. Firm size is measured in the prior trading month. In addition, we include as controls $SEC\ Filing_{t-3,t}$, which is a dummy variable if there was at least one SEC filling (10K, 10Q, or 8K) over the past three trading days and $Press\ release_{t-3,t}$, which is a dummy variable if there was at least one press release issued by the firm over the past three trading days. $Print\ media_{t-3,t}$ is a dummy variable if there was at least one WSJ or NYT article about the firm in the past three trading days. We also control for lagged abnormal retail trading volume on day $t-1$. We include year-month fixed effects, and indicate statistical significance at the ten, five, and one percent levels with *, **, and ***, respectively (t-statistics in parentheses). Standard errors are clustered at the firm level.

Effect on price volatility from articles				
Dependent variable =	Return Volatility$_{t,t+2}$			
Article	0.076***	0.104***	0.123***	0.066***
	(2.90)	(4.21)	(10.72)	(3.40)
Small Firm			0.484***	
			(29.10)	
Article × Small Firm			0.855***	
			(2.98)	
Fraudulent article				2.968***
				(3.97)
Observations	10,617,750	10,617,750	10,617,750	10,617,750
R-squared	0.001	0.001	0.002	0.001
Year-month F.E.	Y	Y	Y	Y
Controls	Y	Y	Y	Y

Table B3. **Article Readership Analysis**

The table reports regression results on the relation between Seeking Alpha readership and abnormal daily trading volume. The readership variables include the log of the number of clicks and number of reads (measured as those that scroll through the entire article) on the day the article is published plus the following two days. All regressions include date and firm fixed effects and indicate statistical significance at the ten, five, and one percent levels, with *, **, and ***, respectively (t-statistics in parentheses).

Dependent variable =	Abnormal daily volume			
log(# Clicks)	0.053***		-0.137***	
	(10.68)		(-6.24)	
log(# Reads)		0.060***	0.191***	
		(12.43)	(8.89)	
Fraction of reads				0.460***
				(8.51)
Observations	14,567	14,567	14,567	14,567
R^2	0.89	0.89	0.89	0.89
Daily F.E.	Y	Y	Y	Y
Firm F.E.	Y	Y	Y	Y

Table B4. **Timeliness of stories**

The table reports regression results on the relation between an article being posted on Seeking Alpha or Motley Fool and the returns in the prior periods. We examine whether that effect differs before and after the SEC investigation became public. AbRet prior day is abnormal return (residuals from a matched portfolio of stocks on size, BE/ME, and momentum) the day $t-1$ before the article is published. AbRet prior week and month are cumulative abnormal returns over $t-7$ or $t-20$ to $t-1$, respectively. All regressions indicate statistical significance at the ten, five, and one percent levels, with *, **, and ***, respectively (t-statistics in parentheses).

Dependent variable=	Article
AbRet prior day	0.009
	(1.15)
AbRet prior day × Post-event	-0.004
	(-0.35)
AbRet prior week	0.005
	(1.00)
AbRet prior week × Post-event	0.001
	(0.08)
AbRet prior month	0.009***
	(4.37)
AbRet prior month × Post-event	-0.005
	(-1.47)
Post-event	0.016***
	(46.93)
Observations	981,489
R-squared	0.002

Table B5. **The 2014 SEC Lawsuit on Trading Volume**

The table examines whether the salience of the presence of fake news on the platforms, stemming from the public announcement of the SEC investigation and lawsuit, impacted investors' reaction to articles on these platforms, as measured by abnormal trading volume. We compare the market's response to articles (in terms of trading volume and price volatility) in the six months before the SEC investigation and six months after the investigation, where we identify the February-March 2014 period as the period when the SEC investigation was announced and covered in the press. We include all firms that have ever had at least one article written about them on Seeking Alpha or Motley Fool during that time period. The regressors include the dummy variable *Article*, which equals 1 if there was at least one article published about the firm on day t, and 0 otherwise, and we include the dummy variable *Post-event*, which equals 1 if the time period is April 1 to September 30, 2014 and is zero if the article was published from August 1, 2013 to January 31, 2014. We exclude all observations prior to August 2013 and after September 2014. We then interact the *Post-event* dummy with the *Article* dummy in the regressions to test for the differential response to articles before versus after the SEC announced investigation. *Small firm* equals 1 if the firm is in the bottom 10th percentile of NYSE firms, and 0 otherwise. Firm size is measured in the prior trading month.We also include controls for SEC filings, press releases, and abnormal volume over the previous three days before the article, plus the day of the article. *SEC filing*$_{t-3,t}$ is a dummy variable if there was at least one SEC filling (10K, 10Q, or 8K) over the past three trading days, and *Press release*$_{t-3,t}$ is a dummy variable if there was at least one press release issued by the firm over the past three trading days. *Print media*$_{t-3,t}$ is a dummy variable if there was at least one WSJ or NYT article about the firm in the past three trading days. We indicate statistical significance at the ten, five, and one percent levels with *, **, and ***, respectively (t-statistics in parentheses). Standard errors are clustered at the firm level.

Dependent variable =	Log(Abnormal daily volume$_{t,t+2}$)		
Article	0.338***	0.389***	0.235***
	(23.50)	(18.24)	(12.61)
Post		-0.179***	-0.079***
		(-17.75)	(-8.80)
Article × Post		-0.041	-0.068**
		(-1.52)	(-2.57)
Small Firm			-0.217***
			(-13.25)
Article × Small Firm			1.000***
			(11.66)
Post × Small Firm			-0.265***
			(-11.56)
Article× Post×Small Firm			-0.277***
			(-2.62)
SEC filing$_{t-3,t}$	0.122***	0.131***	0.118***
	(16.74)	(17.86)	(16.52)
Press release$_{t-3,t}$	0.287***	0.285***	0.246***
	(34.80)	(34.53)	(30.57)
Print media$_{t-3,t}$	0.074***	0.070***	-0.006
	(4.58)	(4.25)	(-0.37)
Abnormal Retail volume$_{t-1}$	1.486***	1.479***	1.450***
	(136.69)	(136.94)	(131.01)
Observations	1,401,509	1,401,509	1,401,509
R-squared	0.366	0.367	0.375