

Social media surges and stock manipulation

Douglas Cumming

School of Business,
Stevens Institute of Technology,
New Jersey, US

Monika Tarsalewska

University of Exeter Business School,
University of Exeter, Exeter, UK

Vu Tran

ICMA Centre,
Henley Business School,
University of Reading, Reading, UK

January 15, 2026

Social media surges and stock manipulation

Abstract

Social media posts are a cost-effective way for stock manipulation compared to engaging in real trades. We find that there is a 300% increase in the number of tweets with buy recommendations on a social media platform around a stock manipulation event. We show that changes in social media volume, disagreement, and more negative tone are significantly associated subsequent manipulation events. Our results suggest that the change in the number of tweets is not sufficient to drive stock prices. Only when combined with sentiment it affects returns. For example, tone might be used for certain manipulation strategies such as “spoofing” where benefits occur when the trading volume decreases, while disagreement might be used for “pump-and-dump” when manipulators benefit when volume increases. In the presence of institutional block holders, the increased attention irrespective of the tone is sufficient for stock price manipulations. Our results are corroborated with robustness and falsification tests.

Keywords: Market Manipulation, Social Media, StockTwits, Surveillance.

JEL Codes: G10, G11, G12, G40

1. Introduction

Stock manipulation can deteriorate financial markets development and undermine integrity and trust, reduce innovation, and distort reallocation of assets (Aggarwal and Wu, 2006; Guiso et al., 2008; Cumming et al. 2020a, 2020b). A rise in social media has transformed how information is spread and absorbed in financial markets. While it facilitates faster and broader access to market news and investor sentiment, it also introduces risks of misinformation, investor herding behavior, and potential for stock manipulation. For example, in 2022 eight social media influencers were charged by Security Exchange Commission (SEC) in \$100 million stock manipulation scheme.¹ The social media misinformation is typically coordinated (Allen et al., 2023). At the end of January 2021, a group of stocks listed on US stock exchanges experienced sudden price increases due to coordination of retail investors on social media platforms. Similarly, anecdotal evidence shows, a significant pump-and-dump scheme involving China Liberal Education, a Nasdaq-listed company, was uncovered by a U.S. Department of Justice investigation where individuals coordinated spread of information through social media. Following the disclosure of these manipulative activities, China Liberal Education's stock plummeted resulting in significant financial losses for many investors.² These types of manipulations on social platforms when happen are not limited to individual stocks but also have wider spillover effects. There is a up to 4.1% decrease in retail trading volume following break in trust due to manipulation (Kogan et al., 2023).

In this paper, we investigate the association between the social media surges and stock market manipulation. While previous literature shows that activity on a social media platform is linked to trading activity, its primary focus is on volume, disagreement, attention, stock prices

¹ https://www.upi.com/Top_News/US/2022/12/14/Eight-social-media-influencers-charged-in-100-million-stock-manipulation-scheme/9681671035732

² <https://www.ft.com/content/829e668b-d7d7-4ee6-a8fc-3d1be6593b0c>

and earnings surprises, and not directly linked to market manipulation (Antweiler and Frank 2004; Chen et al. 2014; Cookson and Niessner 2020; Allen et al. 2023; Cookson et al. 2024).

We study this novel research question by using social media activity data from StockTwist and data on stock manipulation used by surveillance provider for major stock exchanges Nasdaq Market Surveillance (SMARTS).³ Our data sample covers more than 8,800 stocks in NYSE and NASDAQ during the period from January 2010 to December 2018. It is important to acknowledge that in our data we have instances of detected market manipulation cases and that not all of those detected cases lead to formal prosecutions, as enforcement practices may vary. High-frequency data from financial markets is a useful setting to our study as it captures investor attention to the news content. Previous literature shows that news itself, even without affecting prices can trigger trading if markets are informationally efficient or prices can move without trading (Fama 1970; Milgrom and Stokey 1982).

Our initial evidence based on the event studies shows that there were significant social media surges before stock manipulation events. We observe over 300% increase in the number of tweets with stock recommendations within 10 days prior manipulation events. This number only slightly drops after excluding any confounding events such as earnings announcements.

Identifying the causal impact of social media versus other possible channels, is empirically challenging as there can be many factors driving trading activity that can coincide simultaneously with social media surges (omitted variable problem). We therefore use the change in the number of tweets from day -10 to day -1 before manipulation, control for lagged volatility, returns, and trading volume, and other firm characteristics and the firm and year, month fixed effects to eliminate any unobserved firm and year variation. Another possible concern is reverse causality that manipulation affects social media activity. Yet, it is unlikely in our research design. First, we use lagged values of social media surges that occur before the

³ <https://www.nasdaq.com/solutions/fintech/nasdaq-market-surveillance>

manipulation. Second the manipulation incidents are unknown to public on the day they occur. The manipulation alerts issued by surveillance software are only known to authorities and might serve as evidence in subsequent prosecution.

We find that the change in the number of tweets during 10-day period prior has a positive and statistically significant explaining power for stock manipulation. Next, controlling for the change in positiveness and disagreement measures in the previous literature, we find that the change of number of tweets and disagreement positively affects stock manipulation alerts while the change in positiveness has negative and statistically significant effect suggesting that more negative narratives in social media can explain subsequent manipulative trading. In addition, we validate if all those measures of social media have in fact any effect on stock returns and trading volume. All three measures increase the trading volume but only positiveness and disagreement affect stock returns.

This evidence is supported by placebo tests. First, we create placebo random manipulation events and find that the effect of social media surges becomes statistically insignificant. Second, we use SEC investigation based on Kogan et al. (2023) that deteriorated public trust in social media news sources. We find that post this event the effect of social media surges on stock manipulation becomes insignificant and the effect is also present in case of returns and volume. This results further corroborate our results.

Further, we check what kind of techniques can manipulators use through social media surges. We interact stock manipulation with social media activity. We find that change in number of tweets is a way for manipulators to affect the trading volume but not returns suggesting on its own might not be sufficient strategy to benefit. Only change in positiveness and disagreement is a possible channel to affect returns as it has positive effect on returns yet different effect on the trading volume. It is possible that that positiveness might be used for

certain manipulation strategies such as “spoofing” where benefits occur when the trading volume decreases while disagreement might be used for “pump and dump”.

Social media can impact political, economic, and firm dynamics. While most of the previous literature focuses on the direct effects social media on stock markets, we analyse its implication for stock manipulations and market efficiency. Our study is close to Kogan et al. (2023) who examine how fraudulent news disseminated via social media contributes to pump-and-dump schemes and affects trading behavior, market prices, and investor trust. We contribute to this literature by empirically analyzing large-scale data to assess whether surges in social media activity are associated with increased instances of stock price manipulation. Kogan et al. (2023) highlight the mechanisms and consequences of manipulated financial news, we extend this analysis by focusing on the relationship between social media surges and detected manipulation events. Social media networks can lead to coordinated trading, overreaction, price dislocations especially if they are coordinated (Cookson and Niessner 2020; Cookson et al. 2023; Allen et al. 2023). Bushee et al. (2020) show that around IPO quiet periods media can help or mislead retail investors.

The novelty of our study is access to detected market manipulation from surveillance company, that helps us to overcome data limitations on discovered cases of fraud. Proving market manipulation is challenging, as trading activity preceding information releases may be attributed to market anticipation, unusual volatility or end-of-day dynamics. Even in case of enforcement actions detailed data on such cases is often limited or not publicly disclosed (Cumming et al., 2020a). We therefore provide novel evidence on the role of social media activity as a potential signal or catalyst for market abuse.

2. Literature Review

2.1. Social media and financial markets

Traditionally, financial information was often provided by accredited experts such as accountants, analysts, or financial journalists. It was disseminated to investors through a limited number of specialized outlets and personal networks. The emergence of financial social media such as StockTwits, FinTwit, SeekingAlpha, among others, combined with digitalization of financial decisions has reshaped information production and consumption. It allowed non-experts to share market insights widely and retail investors to rely on freely available social media content. It introduced new dynamics how decentralized knowledge across users affects asset prices (Cookson et al., 2024).

Investors on social media express their views on individual stocks on a regular basis enabling the construction of belief measures. Yet, Da et al. (2021) show that beliefs in social media forecasts are systematically biased. Compared to traditional media, social media differ in terms of sentiment. Social media platforms have strongly positive sentiment while traditional media more negative tone (Tetlock et al., 2008; van Binsbergen et al., 2024). Previous literature also focuses on disagreement by measuring the dispersion between bullish and bearish sentiment about securities. Early research by Antweiler and Frank (2004) shows that differences in opinions on message boards translated into trading. Cookson and Niessner (2020) show that disagreement among people with the same investment philosophy is positively associated with trading volume. Hirshleifer et al. (2024) show that while greater social connectedness helps to incorporate news into prices it also increases disagreement. Cookson et al. (2022) show that informed buyers and sellers such as activists and short sellers are attracted to stocks with high unsophisticated disagreement on StockTwits as they have greater liquidity. Yet, when disagreement occurs with shorting constraints informed buyers trade less since it can lead to overvaluation. Ben-Rephael et al. (2024) show that the relationship between disagreement and trading is more pronounced when there is less uncertainty about a firm's trading environment.

Social media news has significant impact on financial markets. Cookson et al. (2023) show that tweets by “tech Twitter” users had immediate market impacts during the 2023 run on Silicon Valley Bank. Separate strand in the literature analyses the implications of attention to stock on social media that is proxied by the volume of messages or the relative volume of messages posted about some stocks versus others (Irvine et al., 2021; Cookson et al., 2024). Yet, there is a difference between the effects of social media sentiment and attention. Cookson et al. (2024) show that the return dynamics following high attention versus high sentiment are quite distinct. While after the run-ups in stock prices, returns revert for up to 10-15 after high-attention days, but not following high-sentiment days.

Given this abundance of information some papers ask what at its implications for markets in general. Kogan et al. (2023) show that post-manipulation trust break on social media has wide economic implications. Dessaint et al. (2024) claim that while social media had a positive effect only on the quality of short-term analyst forecasts but not long-term. In case of mergers social media seems to be a useful signal of corporate decisions (Ang et al., 2021; Cookson et al., 2024)

2.2. Financial misconduct

Financial market misconduct can have various forms. Two of the common forms include: insider trading (Allen and Gale, 1992; Allen and Gorton, 1992; Meulbroek, 1992; Bebchuk and Fershtman, 1994; Agrawal and Cooper, 2015; Bernile et al., 2015; Aitken et al., 2015b) and end-of-day manipulation (Atanasov et al., 2015; Aitken et al., 2015a; Lin et al., 2019). Stock price manipulations benefit the manipulator due to price change. However, there are also significant incentives to manipulate closing stock prices by ramping up end-of-day trading and pushing closing prices to artificially high levels as those prices serve to determine the expiration value of derivative instruments, directors’ options, they are used to estimate the

seasoned equity prices, assess broker performance, calculate the net asset values of mutual funds and stock indices (Aitken et al., 2015a; Cumming et al., 2020a).

2.3. Social media and stock manipulation

Social media reduced the ability of traditional information producers to maintain exclusive control over news dissemination and verification. It also facilitated spread of misinformation, rumours, and unverified content (Cookson et al., 2024). While traditional media stories were driven by editor's choice, social media attention is driven by popularity and algorithms. Social media can also amplify the effects of news. Users on social media platforms are encouraged to post *viral* content i.e. information that is reshared widely (Van Bavel et al., 2021). Often unverified and controversial content is more likely to spread further and attract more attention (e.g. Allcott and Gentzkow, 2017). Chen et al. (2014) show that opinions shared on platforms like Twitter and StockTwits can significantly influence stock valuations, especially if they are amplified by many users. This creates the new scope for coordinated influence on investors (Bikhchandani et al., 2024). This influence created new opportunities for stock manipulation that is mostly costless as the manipulator does not have to incur any bid-ask spread payments as in case of traditional manipulation strategies. We, therefore, predict that social media surges can be associated with stock manipulation.

Cookson et al. (2024) show that while attention is highly correlated across platforms, but sentiment is not. Disagreement and tone of the posts might have a significant effect on trading volumes (Antweiler and Frank 2004; Cookson and Niessner, 2020). Negative tone has implications for corporate decisions (Ang et al., 2021). Cookson et al. (2024) show stock prices return following high-sentiment days suggesting that sentiment and tone can be used for manipulation. We therefore predict that disagreement and more negative tone of posts are positively associated with stock manipulation.

3. Data

Data includes 8,816 companies listed in NYSE and NASDAQ during the period from 1st January 2010 to 31st December 2018. We collected 96,519,390 tweets on the StockTwist social media platform during the sample period. StockTwits is the finance-specific platform. It has a similar format to Twitter where users post short texts about stocks and use cashtags (that were introduced in 2012). It ranks as the 505th most popular website in the US as of June 2021 (website analytics tool Alexa) with monthly 5 million messages and 300,000 users (Cookson et al., 2024).

For each tweet, we collect the tweet content, the time stamp, the user ID number of the poster. We use cashtag tokens ('\$') to identify the mentioned stock(s). Standard cleaning procedure of these tweets is applied following prior papers (e.g. Fan et al., 2020). First, we drop all non-English tweets. Next, special characters are deleted from tweets, such as link tokens (e.g., 'http', 'https', and 'www'), hashtag tokens ('#'), cashtag tokens ('\$'), and user identifier tokens ('@'). In addition, we clean emojis and non-alphabet characters and single-character words from the tweets. All tweets with only links (URLs) and/or hashtag tokens and/or cashtag tokens and/or user identifier tokens are dropped. All tweets posted over weekends are pulled to the following Mondays. After the cleaning procedure, we have over 68,634,972 cleaned tweets.

We focus on tweets that explicitly recommend transactions. Following Loughran and McDonald (2023) methodology, we develop a lexicon of bigrams for stock recommendations. The lexicon contains 190 bigrams such as 'bull flag', 'buy buy', 'buy rating', 'buy signal', 'buy today', 'easy money', 'go higher', 'go up', 'going higher', 'going up', 'good buy', 'hot list', 'long long', 'make money', 'move higher', 'moving up', 'much higher', 'power hour', 'price target', 'ratio short', 'short covering', 'short interest', 'short ratio', 'short sell', 'short squeeze', 'shorts cover',

'stocks buy', 'strong buy', 'strong growth', 'target price', 'today buy', 'today up', 'top stocks', 'up today', 'up tomorrow'. Applying this lexicon filtering, we end up with 11,499,386 tweets.

Our measure of market manipulation comes from Nasdaq Market Surveillance (SMARTS) database. We utilise suspicious transaction alerts Nasdaq Market Surveillance. These alerts are based on intraday trading indicators including trading volume, trading value, return, effective spread, quoted spread. If a given stock exhibits three or more abnormal trading metrics during a given 30-minute interval, a “Continuous Trading Manipulation” alert for the stock during the interval. An “abnormal” trading metric is recorded if the idiosyncratic element of changes in trading indicators compared to the average values in the last 30 days jump up 3 standard deviations away from the average daily difference. This flagging algorithm captures abnormal trading compared to its average and contemporaneous market-wide changes. It also focuses on the spikes in trading indicators over a given 30-min interval but reverting to normal values in the next 30-min intervals throughout the trading day.

We control for lagged accounting-based firm characteristics to account for heterogeneity in firm fundamentals that may affect both social media activity and the market manipulation. These include *Log Market Cap* (logarithm of total assets), *Market-to-Book* ratio, *Quick Ratio* (ratio of current assets to current liabilities), and *Log Revenues* (logarithm of normalized revenues). We control for prior stock performance with the lagged return of the stock (*Return*). It helps to isolate the effects of social media surges and manipulation events from momentum or reversal patterns in stock prices. It helps to ensure the effect of our social media surges is not confounded by recent trends in the firm's stock price. We also control for lagged volatility (*Volatility*), estimated using the Parkinson (1980) method. This method provides a reasonable measure of the latent volatility without requirements of intraday data for realized volatility. By including lagged volatility, we proxy for prior uncertainty in stock prices, which may affect trading behavior. We also include as a control *Log Trading Volume* that is the

logarithm of abnormal trading volume, calculated as the log-difference between the current trading volume and its six-month average, to proxy for events affecting trading activity that could coincide with increased social media activity or manipulation alerts. Finally, we control for lagged S&P 500 return (*S&P 500 Return*) to capture general market movements and macroeconomic sentiment.

In Table 1 we present the summary statistics. Manipulation dummy indicates that stock manipulation happened with mean probability of about 2% of firm-day observations. The median and the 95th percentile are zero, implying that stock manipulations tend to be clustered among certain firms and within certain days. Average (median) number of tweets on a given company in a day is 6.16 (0), implying heavy tails distribution. In most day-firm observations, number of tweets usually very low. We also estimate the 75th percentile which is 2. The mean, median of log-change in social media volume over 2-week periods is anchored at zero. The 5th and 95th percentile of these changes is -1.39 and 1.39, respectively.

The mean value of the logarithm of market capitalization is 6.79 with a standard deviation of 2.04. It corresponds to an average market capitalization of approximately \$889 million (i.e. $\exp(6.79) \approx 889$). The logarithm of total assets has a mean value of 7.15 and a standard deviation of 2.22. When transformed back to levels, the average total assets correspond to approximately \$1,278 million (i.e., $\exp(7.15) \approx 1,274$) that proxies for medium-sized firms. The market-to-book ratio mean is 4.37, indicating that, on average, firms are valued by the market at more than four times their book value. The logarithm of revenues has a mean of 7.98 and a standard deviation of 0.94, corresponding to average revenues of approximately \$2,951 million ($\exp(7.98) \approx 2,922$). The average stock return is 0.04% with a standard deviation of 3.29%, reflecting generally low daily returns but moderate volatility at the firm level. The average S&P 500 return is approximately 0.04% with a standard deviation of 0.94% and the abnormal return has a mean of approximately 0.04% and a standard deviation of 3.29%.

The near-zero average is expected due to the de-meaning inherent in abnormal return calculations. The logarithm of abnormal trading volume has a mean of 11.96 and a standard deviation of 2.35, suggesting substantial heterogeneity in trading activity across firms. The average volatility is 0.03 with a standard deviation of 0.04.

[Insert Table 1]

4. Empirical results

4.1 Event study

In our first test we examine the changes in numbers of tweets around manipulation events. We present the results in Table 2. We focus on windows within 2-months around manipulation events i.e. $[-22, -10]$ in column (1), $[-10, -5]$ in column (2), $[-5, -1]$ in column (3), $[-1, 0]$ in column (4), $[0, 5]$ in column (5), $[5, 10]$ in column (6), and $[10, 22]$ in column (7) with day 0 is the manipulation event date. For example, in column (1), we examine average cumulative changes in numbers of tweets on the sampled companies during the period from one month prior to manipulation events to two weeks prior to manipulation events i.e. $[-22, -10]$.

Event study results in Tables 3 Panel A show that there are significant surges in social media information flows within 2-week prior to manipulation events (columns 2 – 6). Post-event windows, social media volume drop quickly i.e. mostly within 1-week. Magnitude of social media changes around manipulation events are significant. Recall that the mean (median) value of tweet with buy recommendation numbers is 0.99 (0) and the average change in tweet number around manipulation is about 3.3 ($=0.15 + 1.26 + 1.89$) which is an over 330% increase in the tweet numbers scaled by the mean value.

It is possible that earnings announcements can affect our results as they might cause significant social media surges, we therefore remove them from our sample. In panel B of Table 2, we drop observations that there was any earning announcement on day 0, manipulation day. In panel C, we drop observations that there were any earnings announcements within the last

two weeks i.e. during the time window $[-10, 0]$. Consistent with Panel A, social media surges prior to manipulations are still strongly significant and the magnitude of such surges is over 300%. While the increase is abnormally large, the event study does not control for other variables that are correlated with manipulations.

[Insert Table 2]

4.2 Baseline regression

In our baseline model we run a regression of our manipulation measure on the changes in numbers of tweets from day -10 to day -1 i.e. 10-day prior to manipulation events.

$$Manipulation_{i,t}^* = \alpha + \beta \Delta NTweets_{i,[t-10,t-1]} + \gamma Controls + FE + \epsilon_{i,t} \quad (1)$$

Where $Manipulation_{i,t}^*$ is a proxy for stock manipulation on day t . $\Delta NTweets_{i,[t-10,t-1]}$ is the cumulative change in numbers of tweets on firm i for the prior 10 day period. $Controls$ is a vector of control variables. We also include firm and time fixed effects to account for firm- and time invariant heterogeneity.

We try to mitigate omitted variable problem as it is possible that there are other factors driving trading activity that can coincide simultaneously with social media surges (omitted variable problem) by using as our main the change in the number of tweets from day -10 to day -1 before manipulation, control for lagged firm accounting indicators, lagged stock trading return, lagged volatility, lagged trading volume, lagged S&P 500 return, lagged VIX index, and the firm, year, month fixed effects to eliminate any unobserved firm and year variation. Another possible concern in our research design is the reverse causality that manipulation affects social media activity instead. This is quite unlikely in our setting. First, we use lagged values of social media surges that precede manipulation events. Second the manipulation incidents are not publicly known on the day they occur. The alerts generated by surveillance software are accessible only to authorities and may be used as evidence in subsequent prosecutions

[Insert Table 3]

Table 3 presents the baseline results. The key independent variable is change in log number of tweets during 10-day period prior to day t i.e. $[t-10, t-1]$. By including month, year, individual fixed effects we only focus on the within a firm and a month variation. We estimate equation (1) where the dependent variables are: the dummy variable *Manipul.Dummy* in column (1), the logarithm of the number of manipulative transactions plus one (*Log.Manipul.Count*) in column (2), manipulative trading value on day t (*Manipul.Value*) in column (3), and value ratio on day t (*Manipul.Val.Ratio*) in column (4). We estimate equation (1) using OLS linear regression for the sake of interpretation in this baseline. Consistently, the log-change in tweet numbers is highly significant and positive. For example, in column (1) of Table 3, the *coeff.* is 0.0633, implying that 1-unit increase in the log-change in tweet numbers during prior two-week window is associated with 6.33% increases in the manipulation probability. The magnitude is the largest compared to other variables, and about 150% of the impact from lagged trading volume which is 0.0409. This is economically meaningful. Numbers of tweets are much easier and cheaper to be manipulated compared to the trading volume. Even for highly liquid stocks, one should pay 20-30 bpts for transaction fees if one wants to inflate trading volume (e.g. wash trading, two-way buy and sell trading). This could translate to thousands of dollars in costs for few percentage point increases in trading volume. On the other hand, numbers of tweets can be programmatically inflated with much lower costs. Recall the average increase of tweet numbers around manipulation events is about 300% (see Table 2) which is equivalent to the impact of over 450% increases in trading volume.

We also employ different estimation techniques for columns (1) and (2): Logit and Poisson regression, respectively. We show them in Appendix A.2. Across all specifications, the association between the change in the number of tweets and stock manipulation is strongly significant and positive.

Overall, Table 3 confirms that the surges in social media can explain subsequent manipulative trading events. The most likely interpretation is that social media is used prior to manipulation events which is consistent with Allen et al. (2023) and Kogan et al. (2023).

4.3 Regression with Positiveness and Disagreement

Following prior papers (e.g., Cookson & Niessner, 2020) we define additional variables that are *Positiveness* and *Disagreement*. The *Positiveness* measures the net positive tweets on stock i in day t and is defined as follows:

$$Positiveness_{i,t} = \frac{N_{i,t,positive} - N_{i,t,negative}}{N_{i,t,positive} + N_{i,t,negative}} \quad (2)$$

where $N_{i,t,positive}$ and $N_{i,t,negative}$ denotes number of positive and negative tweets on stock i in day t , respectively. The *Disagreement* measures degree social media users disagree with each other (the uncertainty of news) i.e. positive versus negative and is defined as follows:

$$Disagreement_{i,t} = \sqrt{1 - \left(\frac{N_{i,t,positive} - N_{i,t,negative}}{N_{i,t,positive} + N_{i,t,negative}} \right)^2} \quad (3)$$

where $N_{i,t,positive}$ and $N_{i,t,negative}$ denotes number of positive and negative tweets on stock i in day t , respectively. We then add additional two variables to equation (1) that are defined as change in $\Delta Positiveness$ i.e. $\Delta Positiveness_{i,[t-10,t-1]}$ is the cumulative change in net positive narratives of tweets on firm i during the prior 2-week period, and *Disagreement* i.e. $\Delta Disagreement_{i,[t-10,t-1]}$ is the cumulative change in how tweet users disagree with each other on firm i during the prior 2-week period. C is a vector of control variables. The remaining controls in equation (1) remain the same.

We present the results in Table 4. The results consistently show that the log-change in tweet numbers is highly significant and positive. $\Delta Positiveness$ is negative and statistically significant suggesting that more negative narratives in social media are significantly explanatory of subsequent manipulative trading. The coefficient on $\Delta Disagreement$ is not

consistently significant across all columns it only affects *Manipulation* measure positively and in a statistically significant way when it is proxied by manipulation value (manipulation trading value scaled by total trading value) on day t and ratio. The magnitude of both coefficients is smaller than the number of tweets.⁴

[Insert Table 4]

4.4. Social media, stock returns, and trading volume

To make sure that social media surges are used for manipulation we try to establish if they are related changes in future returns and trading volume. We run regression of stock returns and trading volume on the changes in the social media surges.

$$\begin{aligned} Reaction_{i,t} = & \alpha + \beta_1 \Delta NTweets_{i,[t-10,t-1]} + \beta_2 \Delta Positiveness_{i,[t-10,t-1]} + \\ & + \beta_3 \Delta Disagreement_{i,[t-10,t-1]} + \gamma Controls + u_i + \epsilon_{i,t} \end{aligned} \quad (4)$$

where $Reaction_{i,t}$ is return (abnormal) and trading volume (abnormal) of stock i on day t . The remaining variables are defined as before. The change in the number of tweets consistently affect the trading volume but has no effect on abnormal stock returns. Change in positiveness and disagreement have relatively consistent effect on trading volume and abnormal stock returns.

Tables 6 shows estimation results of stock indicators on day t by changes social media aggregate measurements during the prior 2-week period i.e. $[t-10,t-1]$. The dependent variables are raw return, market-model abnormal return, log-trading volume, and abnormal trading volume in columns (1) - (4), respectively. Abnormal return is estimated using the market model with S&P500 proxied for market portfolio. Abnormal trading volume is log-difference between trading volume and the average trading volume in the last 6-month period. We control for

⁴ We present OLS estimation results in this Table 5 for the sake of interpretation and presentation. The corresponding results based on Logit and Poisson techniques are quantitatively similar (reported in Appendix A.3).

(lagged) firm characteristics and trading indicators, and month, year, individual fixed effects. Standard errors are clustered within firms and years. In other words, we only focus on the within a firm and a year variation.

Among the social media measurements, only changes in *Positiveness* are consistently and strongly significant in explaining returns and abnormal returns. The positive coefficient of change in positiveness during last 10-day period means that more positive narratives in social media lead to increase in returns. Coefficient of *Disagreement* is negative implying as people increasingly disagree on a given stock, its prices tend to reduce mildly (about 43 basis points). Effects among social media users during $[t-1, t-10]$ are negatively and significantly linked to today abnormal returns. These results are similar to Cookson et al. (2024) that analyse StockTwist platform.

[Insert Table 5]

4.5. Placebo tests

In this section we run additional tests to corroborate our baseline results by running placebo tests. First, we run a placebo test where we generate 10,000 random fake manipulation events (when there was no actual manipulation that took place on the day). We present the results in Table 6 Panels A and B. The coefficients on social media variables: change in number of tweets, positiveness and disagreement become statistically insignificant.

Second, following Kogan et al. (2023) we use as an exogenous shock, the SEC investigation, that went public in March 2014 and revealed the presence of fraud to market participants in social media. It revealed that while undercover and cooperating with SEC Rick Pearson provided evidence of paid articles promoting certain stocks by disseminating false and promotional information on *Seeking Alpha*. The announcement deteriorated public trust in social media news sources. Kogan et al. (2023) show that it reduced investors' response to

within a week into the event, but the effect disappeared after a year after the event. In a way returns and trading volume became insensitive to social media. We therefore expect that post this event the relationship between social media surges and manipulation should become ineffective.

For our analysis we include 6 months before and 6 months after March 2024 and estimate the following equation:

$$Manipulation_{i,t}^* = \alpha + \beta Post + \beta Post \times \Delta NTweets_{i,[t-10,t-1]} + \beta NTweets_{i,[t-10,t-1]} + \gamma Controls + FE + \epsilon_{i,t} \quad (5)$$

where *Post* is equal to 1 after March 2014, and zero otherwise. Where $Manipulation_{i,t}^*$ is a proxy for stock manipulation on day t . $\Delta NTweets_{i,[t-10,t-1]}$ is the cumulative change in numbers of tweets on firm i for the prior 10 day period. *Controls* is a vector of control variables. We also include firm and time fixed effects to account for firm- and time invariant heterogeneity as in Eq. (1).

We present the results in Table 6 Panel C. Consistently with our predictions we find that post this event the effect of social media surges on stock manipulation becomes insignificant. We also show in Table 6 Panel C that the effect of social media surges on returns and volume is also statistically insignificant.⁵

[Insert Table 6]

4.6. Social media and manipulation: possible mechanisms

In this section we aim to determine what mechanism and schemes might be used for stock manipulations. In general, social media surges might be a cost-efficient way for schemes like “pump and dump” or “spoofing” compared to traditional ways.

⁵ We present OLS estimation results in this Table 7 Panel C for the sake of interpretation and presentation. The corresponding results based on Logit and Poisson techniques are quantitatively similar (reported in Appendix A.4).

In order to test for the mechanisms, we interact stock manipulation with social media activity and analyse its effects on returns and trading volume. We present the results in Table 7. The interaction term between manipulation and change in tweet numbers is insignificant in explaining (abnormal) returns while it is strongly significant and positive in explaining (abnormal) trading volume. The interaction between manipulation dummy and positive narratives in tweets is strongly significant and positive in explaining (abnormal) returns while it is also strongly significant but negative in explaining (abnormal) trading volume. The interaction between manipulation dummy and disagreement in tweets is strongly significant in explaining both (abnormal) returns and (abnormal) trading volume.

These results suggest that the number of tweets is a way for manipulators to affect the trading volume but not returns. Therefore, an increase in the number of tweets on its own might not be sufficient strategy to benefit for manipulators. Only change in *Positiveness* and *Disagreement* is a possible channel to affect returns as it has positive effect on returns yet different effect on the trading volume. It is possible that that positiveness might be used for certain manipulation strategies such as “spoofing” where benefits occur when the trading volume decreases while disagreement might be used for “pump and dump”.

[Insert Table 7]

4.8 *The role of institutional investors*

Finally, we investigate the role of institutional ownership on manipulation and social media. Allen et al. (2006) show that large investors and corporate insiders hold significant market power, enabling them to impact or manipulate prices. Therefore, one can expect that in the presence of institutional block holders can facilitate stock price manipulation. In Table 8 we

investigate the role of institutional investors on manipulation and social media.⁶ In Table 8 Panel A we add the dummy variable for high institutional ownership and its interaction with social media indicators in models explaining manipulative variables (in the left-hand-side of equations). The *High-Institutional* ownership dummy equals 1 if the long-term holding proportion is larger than 15% which is the empirical mean for our sample. We observe that the coefficient of *High-Institutional* ownership dummy is consistently negative and significant across all models and specifications. This implies that institutional investors help to deter manipulative behaviours and transactions in general. However, the coefficients of the interaction terms between High-Institutional ownership dummy and change in tweet numbers is significantly positive.

The presence of institutional investors mostly mitigates the effects of *Positiveness* and *Disagreement* on stock manipulation (Table 8 Panel B). Yet, the relation between change in the number of tweets and trading volume is more pronounced in the presence of institutional investors. This is quite surprising as the literature suggests that the retail investors are mostly susceptible to such manipulation. Yet, this evidence suggests that the positive relation between number of tweets and stock manipulation is more pronounced in the presence of institutional investors. In the presence of institutional investors, the attention to the stock itself by increasing the number of tweets is sufficient to affect positively manipulation and trading volume.

This evidence would suggest that in case of stock with institutional block holders the increased attention irrespective of the tone is sufficient for stock manipulations. It is consistent with Cookson et al. (2022) that informed traders respond meaningfully but differently to valuation changes induced by unsophisticated disagreement.

[Insert Table 8]

⁶ We present OLS estimation results in Tables 9 and 10 for the sake of interpretation and presentation. The corresponding results based on Logit and Poisson techniques are quantitatively similar (reported in Appendix A.5).

5. Conclusion

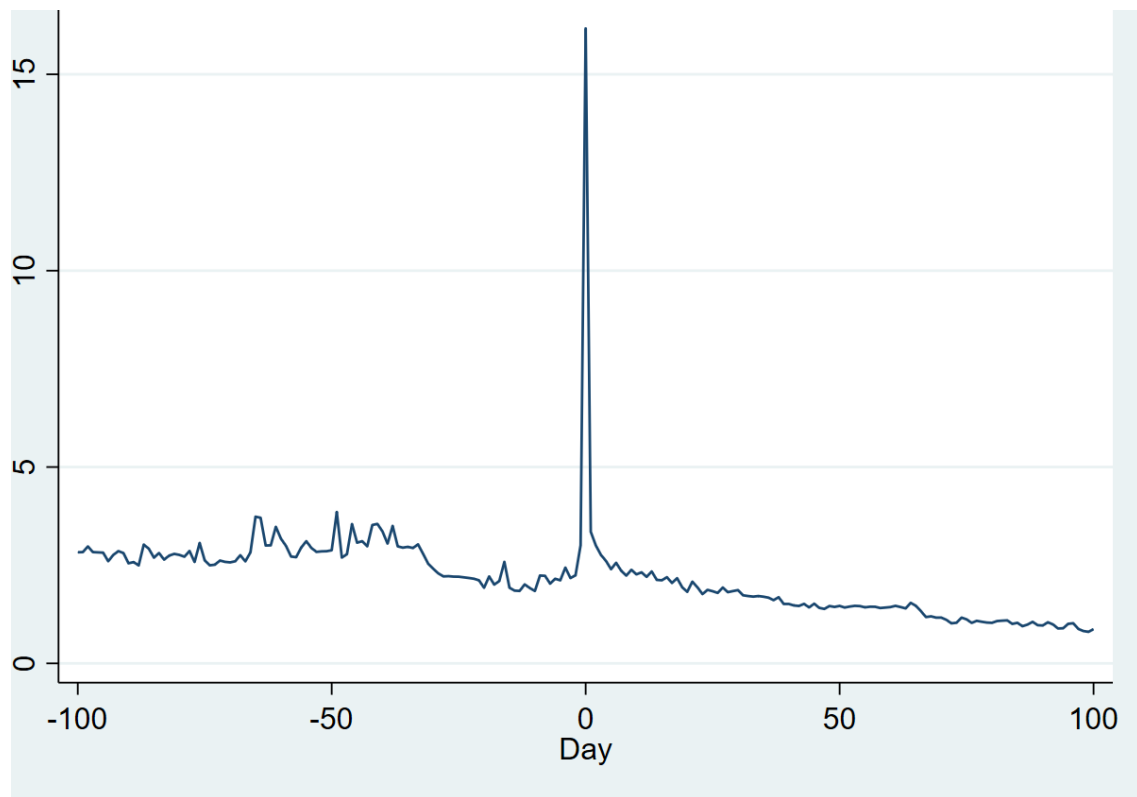
We show that stock manipulation follows social media surges. By studying social media activity data from StockTwist and stock manipulation from NASDAQ market surveillance (SMARTS) database, we find an approximately 300% increase in the number of tweets within 10 days prior to manipulation events. Using regression analysis, we find that the change in the number of tweets during 10-day period prior can significantly explain following stock manipulation. Our results are robust to a number of tests and supported by placebo tests. We further find that change in number of tweets alone can influence trading volume but not necessarily returns. In contrast, changes in sentiment, specifically positiveness and disagreement, can affect returns, though they impact trading volume differently. Positiveness may be used in manipulation strategies like “spoofing,” where the advantage comes from reduced trading volume. Disagreement, on the other hand, may facilitate “pump and dump” schemes by driving up returns alongside increased market activity. Overall, we show novel evidence on the role of social media activity as a precursor to market manipulation.

References

- Aggarwal, R. and Wu, G. (2006): Stock market manipulations, *Journal of Business* 79, 1915–1953.
- Agrawal, A., Cooper, T. (2015). Insider trading before accounting scandals. *Journal of Corporate Finance* 34, 169–190.
- Aitken, M., Cumming, D.J., Zhan, F. (2015a). High frequency trading and end-of-day price dislocation. *Journal of Banking and Finance* 59, 330–349.
- Aitken, M., Cumming, D.J., Zhan, F. (2015b). Exchange trading rules, surveillance, and suspected insider trading. *Journal of Corporate Finance* 34, 311–330.
- Allen, F., Gale, D. (1992). Stock-price manipulation. *Review of Financial Studies* 5, 503–529.
- Allen, F., Gorton, G. (1992). Stock price manipulation, market microstructure and asymmetric information. *European Economic Review* 36, 624–630.
- Allen, F., Litov, L., Mei, J. (2006). Large Investors, Price Manipulation, and Limits to Arbitrage: An Anatomy of Market Corners. *Review of Finance* 10 (4), 645–693.
- Allen, F., Haas, M.D., Nowak, E., and Tengulov, A. (2023). Squeezing Shorts through social media platforms. Swiss Finance Institute Working Paper.
- Allcott, H., Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives* 31, 211–236.
- Ang, J. S., Hsu, C., Tang, D., Wu, C. (2021). The role of social media in corporate governance. *Accounting Review* 96 (2), 1–32.
- Antweiler, W., Frank, M.Z. (2004). Is all that talk just noise? The information content of internet stock message boards. *Journal of Finance* 59 (3), 1259–1294.
- Atanasov, V., Davies, R.J., Merrick, J.J., Jr. (2015). Financial Intermediaries in the Midst of Market Manipulation: Did They Protect the Fool or Help the Knave? *Journal of Corporate Finance* 34, 210–234.
- Ben-Rephael, A., Cookson, J. A., Izhakian, Y. Y. (2024). Do I really want to hear the news? Public information arrival, ambiguity and investor beliefs. Available at SSRN 3631760.
- Bernile, G., Sulaeman, J., Wang, Q. (2015). Institutional trading during a wave of corporate scandals: ‘Perfect Payday’? *Journal of Corporate Finance* 34, 191–209.
- Bikhchandani, S., Hirshleifer, D., Tamuz, O., Welch, I. (2024). Information cascades and social learning. *Journal of Economic Literature* 62(3), 1040–1093.
- Bushee, B., Cedergrén, M., Michels, J. (2020). Does the media help or hurt retail investors during the IPO quiet period? *Journal of Accounting and Economics* 69, 101261.
- Chen, H., De, P., Hu, Y.J., Hwang, B.H. (2014). Wisdom of crowds: The value of stock opinions transmitted through social media. *Review of Financial Studies* 27, 1367–1403.
- Comerton-Forde, C., Putniņš, T.J. (2011). Measuring closing price manipulation. *Journal of Financial Intermediation* 20(2), 135–158.
- Comerton-Forde, C., Putniņš, T.J. (2014). Stock Price Manipulation: Prevalence and Determinants. *Review of Finance* 18 (1), 23–66.
- Cookson, J.A., Engelberg, J.E., Mullins, W. (2023). Echo Chambers. *Review of Financial Studies* 36, 450–500.
- Cookson, J.A., Mullins, W., Niessner, M. (2024). Social media and finance. Available at SSRN 4806692.
- Cookson, J.A., Niessner, M. (2020). Why don’t we agree? Evidence from a social network of investors. *Journal of Finance*.
- Cumming, D., Ji, S., Johan, S., Tarsalewska, M. (2020b). End-of-day price manipulation and M&As. *British Journal of Management* 31(1), 184–205.
- Cumming, D., Ji, S., Peter, R., Tarsalewska, M. (2020a). Market manipulation and innovation. *Journal of Banking and Finance* 120, 105957.

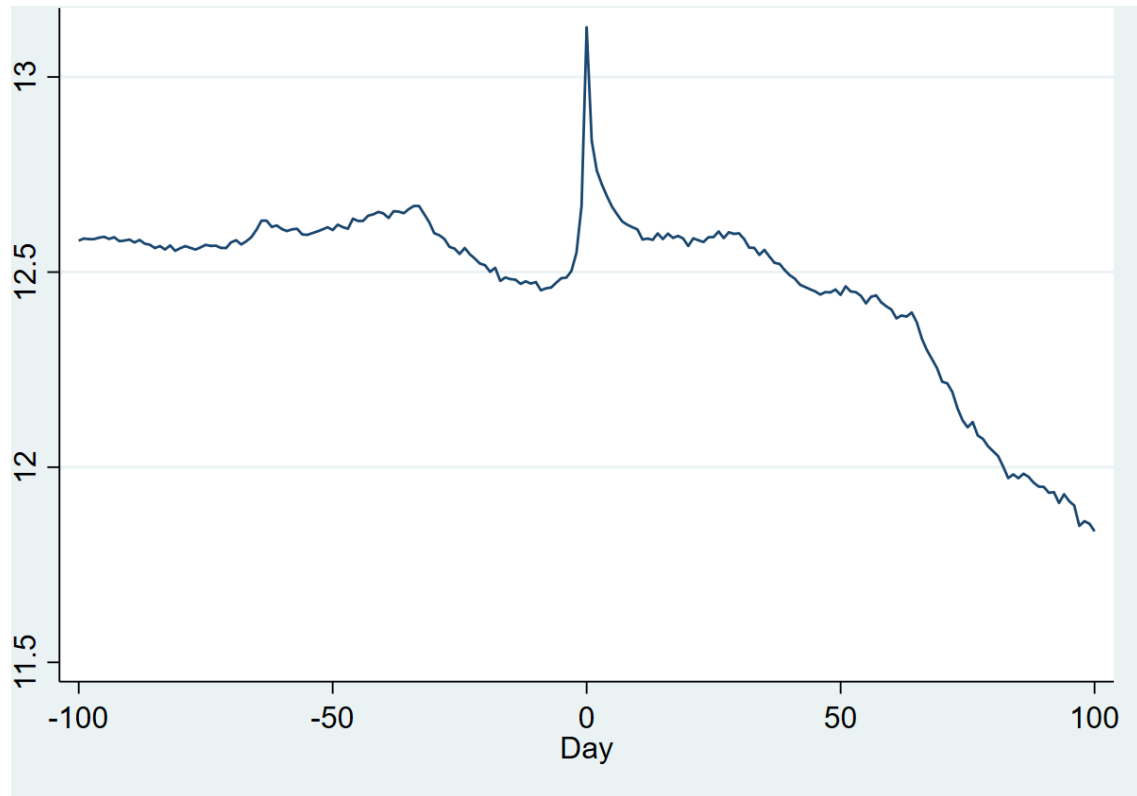
- Da, Z., Huang, X., Jin, L.J. (2021). Extrapolative beliefs in the cross-section: What can we learn from the crowds? *Journal of Financial Economics* 140(1), 175–196.
- Dessaint, O., Foucault, T., Frésard, L. (2024). Does alternative data improve financial forecasting? The horizon effect. *Journal of Finance* 79 (3), 2237–2287.
- Fama, E.F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance*.
- Fan, R., Talavera, O., Tran, V. (2020) Social media bots and stock markets. *European Financial Management* 26, 753-777.
- Guiso, L., Sapienza, P., Zingales, L. (2008). Trusting the stock market. *Journal of Finance* 63, 2557–2600.
- Hirshleifer, D., Peng, L., Wang, Q. (2024). News diffusion in social networks and stock market reactions. *Review of Financial Studies*, Forthcoming.
- Kogan, S., Moskowitz, T., Niessner, M. (2023). Social media and financial news manipulation. *Review of Finance*, 1229–1268.
- Loughran, T. and McDonald B. (2023) Measuring firm complexity. *Journal of Financial and Quantitative Analysis* 59(6), 2487–2514.
- Milgrom, P., Stokey, N. (1982). Information, trade and common knowledge. *Journal of Economic Theory*.
- Tetlock, P.C., Saar-Tsechansky, M., Macskassy, S. (2008). More than words: Quantifying language to measure firms' fundamentals. *Journal of Finance* 63 (3), 1437–1467.
- van Binsbergen, J.H., Bryzgalova, S., Mukhopadhyay, M., Sharma, V. (2024). (Almost) 200 years of news-based economic sentiment. NBER Working Paper.

Figure 1. Number of tweets around stock manipulations



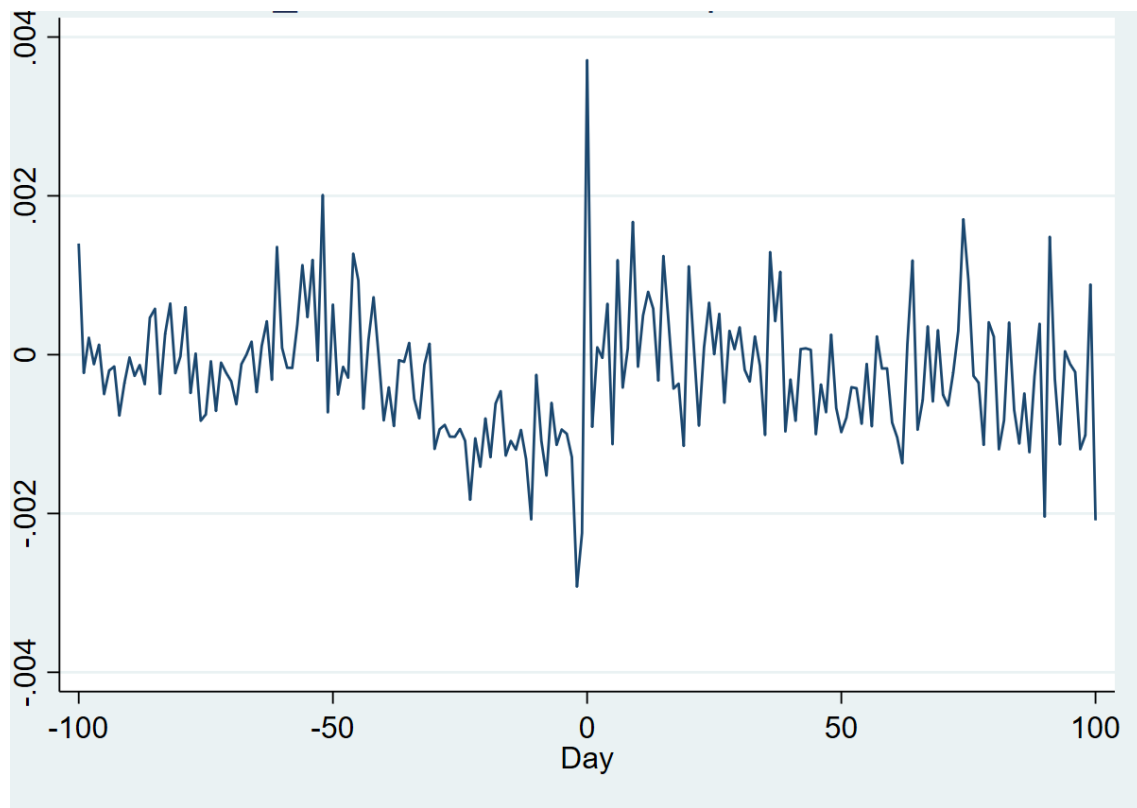
Note: This figure plots the average number of tweets with any of 190 keywords around manipulation events i.e. day 0. The sampled period is from 1st January 2010 to 31st December 2018.

Figure 2. Trading volume around stock manipulations



Note: This figure plots the average log-trading volume around manipulation events i.e. day 0. The sampled period is from 1st January 2010 to 31st December 2018.

Figure 3. Abnormal return around stock manipulations



Note: This figure plots the average abnormal return around manipulation events i.e. day 0. We estimate the market model abnormal return using the last 6-month period. The sampled period is from 1st January 2010 to 31st December 2018.

Table 1: Summary statistics

	(1) N	(2) Mean	(3) Median	(4) SD	(5) 5 th pct.	(6) 95 th pct.
Manipul.Dummy	11,419,791	0.01	0.00	0.12	0.00	0.00
Manipul.Count	11,419,791	0.02	0.00	0.20	0.00	0.00
Manipul.Value	11,419,791	0.20	0.00	1.67	0.00	0.00
Manipul.Val.Ratio	11,419,791	0.00	0.00	0.03	0.00	0.00
N_tweets _t	11,419,791	0.99	0.00	7.75	0.00	4.00
$\Delta N_{tweets}_{[t-10, t-1]}$	11,348,270	0.00	0.00	0.56	-0.92	0.92
$\Delta Positiveness_{[t-10, t-1]}$	11,229,489	-0.01	0.00	0.25	0.00	0.00
$\Delta Disagreement_{[t-10, t-1]}$	11,230,136	0.05	0.00	0.20	0.00	0.61
Log Market Cap.	7,377,497	6.79	6.82	2.05	3.37	10.19
Market-to-Book	7,354,428	3.08	1.94	5.90	0.31	10.81
Quick Ratio	6,419,561	2.82	1.97	3.07	0.67	7.79
Log Revenues	8,243,240	7.99	7.59	0.95	7.25	10.08
Return (%)	11,112,802	0.04	0.00	3.26	-3.94	3.97
Volatility	11,115,027	0.03	0.02	0.04	0.01	0.09
Log Trading Volume	11,115,039	12.01	12.20	2.30	8.19	15.31
Institutional Ownership	7,543,013	0.15	0.12	0.16	0.00	0.43
S&P 500 Return (%)	11,117,944	0.04	0.05	0.94	-1.57	1.45
VIX	11,316,042	16.97	15.52	5.82	10.50	29.46

This table reports the summary statistics of variables used in this paper. Our final data includes 11,499,386 tweets on the StockTwist platform during the sample period. Each of these tweets contain at least one of the keywords for stock recommendation. In addition, each of these tweets also mentions a stock(s) that have been flagged as manipulative or suspicious in the NASDAQ Market Surveillance database. The sample includes daily observations for 8,816 companies listed in NYSE and NASDAQ during the period is from 1st January 2010 to 31st December 2018. Columns (1) – (6) report the number of daily observations, the mean value, the median value, the standard deviation, the 5th percentile and the 95th percentile, respectively.

Table 2: Event Study for numbers of tweets around manipulation

Time Window	(1) [-22,-10]	(2) [-10,-5]	(3) [-5,-1]	(4) [-1,1]	(5) [1,5]	(6) [5,10]	(7) [10,22]
Panel A: Event Study for numbers of tweets around all manipulation dates							
ΔN_{tweets}	-0.0029	0.1579***	1.2638***	1.8972***	-2.7632***	-0.1286***	-0.0190
t-stat	-0.1771	9.9508	44.1029	52.1857	-61.8413	-4.9299	-0.6495
Obs.	167,688	168,790	169,296	169,783	169,897	169,813	169,544
Panel B: Event Study for numbers of tweets around manipulation dates without earning announcements							
ΔN_{tweets}	-0.0068	0.1403***	1.2224***	1.7099***	-2.5079***	-0.1223***	-0.0241
t-stat	-0.4083	8.5467	41.0595	46.5602	-55.3617	-4.4883	-0.7872
Obs.	160,098	161,154	161,657	162,142	162,256	162,172	161,903
Panel C: Event Study for numbers of tweets around manipulation dates without earning announcements within the last two weeks i.e. during window [-10,0]							
ΔN_{tweets}	-0.0017	0.1330***	1.0109***	1.8241***	-2.3010***	-0.1600***	-0.1229***
t-stat	-0.1011	7.5102	32.1115	41.9211	-44.0431	-4.6733	-3.4350
Obs.	117,302	118,334	119,321	119,321	119,435	119,351	119,082

This table reports event study of numbers of tweets with transaction recommendations around all manipulation dates. A manipulation date, day 0, is defined as if there is any manipulation alert during the date. Column (1) shows results t-tests during the time window $[-22, -10]$ i.e. 1-month to 2-week prior to manipulation events. Column (2) and (3) present corresponding results for 2-week to 1-week and 1-week prior to manipulation events, respectively. Columns (5) to (7) show the corresponding results for the symmetric time windows i.e. post-manipulation events. Panel A reports the results based on all manipulation events. Panels B and C drop manipulation events surrounding earning announcements.

Table 3: Abnormal changes in numbers of tweets before manipulative/suspicious events

	(1) Manipul.Dummy _t	(2) Log.Manipul.Count _t	(3) Manipul.Value _t	(4) Manipul.Val.Ratio _t
$\Delta \text{Ntweets}_{[t-10, t-1]}$	0.0529*** (9.19)	0.0603*** (8.43)	0.0557*** (9.65)	0.0383*** (7.76)
Log Market Cap. _{t-1}	0.0267*** (5.17)	0.0253*** (5.16)	0.0321*** (6.39)	0.0120** (2.89)
Market-to-Book _{t-1}	-0.0004 (-0.67)	-0.0003 (-0.61)	-0.0004 (-0.73)	-0.0003 (-0.48)
Quick Ratio _{t-1}	-0.0014 (-1.05)	-0.0012 (-1.00)	-0.0013 (-1.02)	-0.0012 (-0.99)
Log Revenues _{t-1}	0.0247** (2.98)	0.0329*** (3.68)	0.0385*** (4.29)	0.0225** (2.79)
Return _{t-1}	-0.0093*** (-7.77)	-0.0103*** (-8.37)	-0.0091*** (-8.18)	-0.0075*** (-8.03)
Volatility _{t-1}	0.0105*** (4.15)	0.0081*** (3.57)	0.0095*** (4.09)	0.0054** (3.29)
Log Trading Volume _{t-1}	0.0470*** (9.57)	0.0490*** (8.08)	0.0491*** (9.62)	0.0317*** (7.07)
S&P 500 Return _{t-1}	-0.0074*** (-4.18)	-0.0054*** (-4.29)	-0.0074*** (-4.34)	-0.0038** (-2.50)
VIX _{t-1}	-0.0085 (-1.45)	-0.0115* (-2.08)	-0.0089 (-1.56)	-0.0114* (-2.08)
Firm FEs	Yes	Yes	Yes	Yes
Time FEs	Yes	Yes	Yes	Yes
Observations	6,226,978	6,226,978	6,226,978	6,226,978
Adj. R ²	0.012	0.013	0.013	0.007

Note: This table reports estimation results of stock manipulation by abnormal changes in numbers of tweets during the 2-week period prior to manipulation events i.e., [t-10, t-1]. The dependent variable is a manipulative event dummy in column (1), a logarithm of manipulative counts plus one in column (2), manipulative trading value in column (3), a ratio of manipulative trading value to total trading turnover in column (4). The key independent variable is change in log number of tweets during 10-day period prior to the day t i.e. [t-10, t-1]. Control variables include (lagged) firm characteristics and trading indicators, and month, year, firm fixed effects. Standard errors are clustered within firms and years.

Table 4: Abnormal changes in social media indicators before manipulation events

	(1) Manipul.Dummy _t	(2) Log.Manipul.Count _t	(3) Manipul.Value _t	(4) Manipul.Val.Ratio _t
$\Delta N_{\text{tweets}}_{[t-10, t-1]}$	0.0545*** (7.48)	0.0624*** (7.00)	0.0572*** (7.83)	0.0396*** (6.60)
$\Delta \text{Positiveness}_{[t-10, t-1]}$	-0.0057* (-2.14)	-0.0069* (-2.19)	-0.0057* (-2.11)	-0.0045* (-2.28)
$\Delta \text{Disagreement}_{[t-10, t-1]}$	0.0032 (1.14)	0.0038 (1.16)	0.0045 (1.50)	0.0021 (0.90)
Controls	Yes	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes	Yes
Time FEs	Yes	Yes	Yes	Yes
Observations	6,078,145	6,078,145	6,078,145	6,078,145
Adj. R ²	0.012	0.013	0.013	0.008

This table reports estimation results of manipulation by changes in aggregate social media indicators during the 2-week period prior to manipulation events i.e., [t-10, t-1]. The dependant variables are a dummy for manipulation, logarithm number of manipulative transactions plus one, manipulative value and manipulative value ratio on day t in columns (1) - (4), respectively. The key independent variables are changes in the log number of tweets, net positive narratives in all tweets, and degree of disagreement between social media users. All these three variables are calculated during 10-day period prior to the day t i.e. [t-10, t-1]. Control variables include (lagged) firm characteristics and trading indicators, and month, year, individual fixed effects. Standard errors are clustered within firms and years.

Table 5: Can we explain (abnormal) return and trading volume by changes in social media?

	(1) Return _t	(2) Abn.Return _t	(3) Log.Trade.Volum _t	(4) Abn.Trade.Volum _t
$\Delta \text{Ntweets}_{[t-10, t-1]}$	-0.0005 (-0.24)	0.0012 (0.78)	0.0079*** (6.30)	0.0619*** (19.45)
$\Delta \text{Positiveness}_{[t-10, t-1]}$	0.0077*** (8.66)	0.0068*** (8.60)	0.0012** (2.97)	0.0030 (1.64)
$\Delta \text{Disagreement}_{[t-10, t-1]}$	-0.0011 (-0.79)	-0.0050*** (-3.65)	0.0214*** (12.79)	0.0256*** (11.97)
Controls	Yes	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes	Yes
Time FEs	Yes	Yes	Yes	Yes
Observations	5,711,792	4,151,852	5,711,823	5,200,974
Adj. R ²	0.003	0.001	0.864	0.249

This table reports GLS estimation of (abnormal) stock return and trading volume by changes in social media during the period from day -10 to day -1. The dependant variables are stock return, abnormal stock return, (log) trading volume, and abnormal trading volume on day t in columns (1) - (4), respectively. Abnormal stock return equal raw return minus expected return which is estimated using the market model during lagged 6-month period. Abnormal trading volume is log-difference between day t volume and the average volume in the last six months. The key independent variables are changes in social media measurements from day t-10 to day t-1. The measures include log-number of tweets, aggregate positiveness, and aggregate disagreement in social media. Control variables include (lagged) firm characteristics and trading indicators, and month, year, individual fixed effects. Standard errors are clustered within firms and years.

Table 6: Placebo tests**Panel A: Placebo manipulations - social media surges**

	(1) Manipul.Dummy _t	(2) Manipul.Dummy _t
$\Delta \text{Ntweets}_{[t-10, t-1]}$	-0.6252 (-1.42)	-0.4015 (-0.99)
$\Delta \text{Positiveness}_{[t-10, t-1]}$		-0.3276 (-0.69)
$\Delta \text{Disagreement}_{[t-10, t-1]}$		0.2542 (0.65)
Controls	Yes	Yes
Firm FE	Yes	Yes
Month FE	Yes	Yes
Year FE	Yes	Yes
Observations	4,935,432	4,422,965
Pseudo R ²	0.020	0.019

Note: We randomly generate 10,000 fake manipulation events and re-estimate equation (1).

Panel B: Placebo manipulations – returns and volume

	(1) Return _t	(2) Abn.Return _t	(3) Log.Trade.Volum _t	(4) Abn.Trade.Volum _t
$\Delta \text{Ntweets}_{[t-10, t-1]}$	0.0128 (0.45)	0.0027 (0.10)	0.0190 (1.64)	0.0884* (1.91)
$\Delta \text{Positiveness}_{[t-10, t-1]}$	0.0165 (0.75)	0.0362 (1.32)	-0.0068 (-1.02)	0.0079 (0.49)
$\Delta \text{Disagreement}_{[t-10, t-1]}$	0.0112 (0.41)	-0.0002 (-0.01)	0.0131 (1.67)	0.0074 (0.39)
Controls	Yes	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes	Yes
Time FEs	Yes	Yes	Yes	Yes
Observations	3,762	2,841	3,762	3,353
Adj. R ²	-0.006	-0.001	0.160	0.135

Note: We randomly generate 10,000 fake manipulation events and re-estimate equation (4).

Panel C: Impact of social media following SEC investigation of media coverage of the fake online articles scandal February – March 2014.

	(1) Manipul.Dummy _t	(2) Log.Manipul.Count _t	(3) Manipul.Value _t	(4) Manipul.Val.Ratio _t
$\Delta Ntweets_{[t-10, t-1]}$	0.0599*** (4.69)	0.0663*** (4.30)	0.0639*** (4.83)	0.0427*** (4.96)
$\Delta Positiveness_{[t-10, t-1]}$	-0.0091* (-1.95)	-0.0113** (-2.15)	-0.0096** (-2.04)	-0.0078** (-2.42)
$\Delta Disagreement_{[t-10, t-1]}$	0.0024 (1.47)	0.0018 (1.27)	0.0034* (1.88)	-0.0003 (-0.13)
Post x $\Delta Ntweets_{[t-10, t-1]}$	0.0043 (0.54)	0.0066 (0.69)	0.0034 (0.42)	0.0031 (0.63)
Controls	Yes	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes	Yes
Time FEs	Yes	Yes	Yes	Yes
Observations	753,754	753,754	753,754	753,754
Adj. R ²	0.012	0.014	0.014	0.008

Note: This table reports estimation results of manipulation by changes in aggregate social media indicators during the 2-week period prior to manipulation events i.e., [t-10,t-1]. The dependent variables are a dummy for manipulation, logarithm of number of manipulative transactions plus one, manipulative value and manipulative value ratio on day t in columns (1) - (4), respectively. The key independent variables is the interaction term between changes in the log number of tweets and the Post dummy for 6-month period following SEC public investigation of online fake articles on a social media platform in February – March 2014. All these three variables are calculated during 10-day period prior to the day t i.e. [t-10,t-1]. Control variables include (lagged) firm characteristics and trading indicators, and month, year, individual fixed effects. This table uses a subsample of 6 months before and 6 months after March 2024. Standard errors are clustered within firms and years. Post is subsumed by time fixed effects.

Table 7: Explaining return and trading volume by interaction changes in social media with manipulation dummy

	(1) Return _t	(2) Abn.Return _t	(3) Log.Trading.Volum _t	(4) Abn.Trade.Volum _t
Δ Ntweets _[t-10, t-1]	-0.0017 (-0.89)	0.0003 (0.18)	0.0043*** (4.04)	0.0517*** (21.74)
Δ Positiveness _[t-10, t-1]	0.0059*** (7.69)	0.0051*** (6.83)	0.0016*** (4.10)	0.0041** (2.51)
Δ Disagreement _[t-10, t-1]	-0.0040** (-2.68)	-0.0081*** (-5.54)	0.0189*** (11.08)	0.0186*** (8.22)
Manipulation x Δ Ntweets _[t-10, t-1]	0.0007 (0.48)	-0.0005 (-0.26)	0.0151*** (20.57)	0.0437*** (29.35)
Manipulation x Δ Positiveness _[t-10, t-1]	0.0115*** (8.89)	0.0107*** (6.98)	-0.0025*** (-4.86)	-0.0075*** (-6.50)
Manipulation x Δ Disagreement _[t-10, t-1]	0.0192*** (17.15)	0.0198*** (8.52)	0.0155*** (67.84)	0.0430*** (69.78)
Controls	Yes	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes	Yes
Time FEs	Yes	Yes	Yes	Yes
Observations	5,711,792	4,151,852	5,711,823	5,200,974
Adj. R ²	0.003	0.002	0.865	0.254

This table reports GLS estimation of (abnormal) stock return and trading volume by changes in social media during the period from day -10 to day -1. The dependant variables are stock return, abnormal stock return, (log) trading volume, and abnormal trading volume on day t in columns (1) - (4), respectively. Abnormal stock return equal raw return minus expected return which is estimated using the market model during lagged 6-month period. Abnormal trading volume is log-difference between day t volume and the average volume in the last six months. The key independent variables are changes in social media measurements from day $t-10$ to day $t-1$ and the interaction terms between these and the manipulation dummy. Control variables include (lagged) firm characteristics and trading indicators, and month, year, individual fixed effects. Standard errors are clustered within firms and years.

Table 8: Role of institutional investors in manipulations**Panel A: Interaction with Δ Ntweets**

	(1) Manipul.Dummy _t	(2) Log.Manipul.Count _t	(3) Manipul.Value _t	(4) Manipul.Val.Ratio _t
Δ Ntweets _[t-10, t-1]	0.0586*** (12.73)	0.0649*** (10.07)	0.0608*** (13.35)	0.0412*** (8.61)
Δ Positiveness _[t-10, t-1]	-0.0052*** (-4.33)	-0.0060*** (-4.11)	-0.0054*** (-4.49)	-0.0042*** (-3.83)
Δ Disagreement _[t-10, t-1]	0.0081** (3.03)	0.0104** (3.01)	0.0095*** (3.43)	0.0064* (2.28)
High Institution	-0.0028** (-3.12)	-0.0029*** (-3.54)	-0.0029*** (-3.40)	-0.0026*** (-3.66)
High Institution x Δ Ntweets _[t-10, t-1]	0.0102*** (3.76)	0.0129*** (3.49)	0.0107*** (3.57)	0.0086** (3.23)
Controls	Yes	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes	Yes
Time FEs	Yes	Yes	Yes	Yes
Observations	5,769,613	5,769,613	5,769,613	5,769,613
Adj. R ²	0.014	0.015	0.015	0.008

This table reports estimation results of manipulative trading by the interaction term between High-Institutional Holding dummy and changes in aggregate social media indicators during the 2-week period prior to manipulation events i.e., [-10,-1]. The dependent variables are a dummy for manipulation, logarithm number of manipulative transactions plus one, manipulative value and manipulative value ratio on day t in columns (1) - (4), respectively. The key independent variables are the interaction term between High-Institutional Holding dummy and changes in the log number of tweets. All these three variables are calculated during 10-day period prior to the day t i.e. [t-10,t-1]. Control variables include (lagged) firm characteristics and trading indicators, and month, year, individual fixed effects. Standard errors are clustered within firms and years.

Panel B: Interactions with other social media indicators

	(1) Manipul.Dummy _t	(2) Log.Manipul.Count _t	(3) Manipul.Value _t	(4) Manipul.Val.Ratio _t
$\Delta Ntweets_{[t-10, t-1]}$	0.0589*** (12.67)	0.0653*** (10.03)	0.0612*** (13.27)	0.0416*** (8.45)
$\Delta Positiveness_{[t-10, t-1]}$	-0.0053*** (-4.59)	-0.0062*** (-4.89)	-0.0056*** (-4.79)	-0.0048*** (-4.32)
$\Delta Disagreement_{[t-10, t-1]}$	0.0064** (3.05)	0.0079** (3.19)	0.0075*** (3.52)	0.0050** (2.42)
High Institution	-0.0039*** (-3.84)	-0.0045*** (-4.19)	-0.0042*** (-4.12)	-0.0034** (-3.33)
High Institution x $\Delta Ntweets_{[t-10, t-1]}$	0.0097*** (3.73)	0.0121*** (3.48)	0.0100*** (3.49)	0.0080** (3.25)
High Institution x $\Delta Positiveness_{[t-10, t-1]}$	0.0002 (0.25)	0.0003 (0.44)	0.0004 (0.53)	0.0010 (1.37)
High Institution x $\Delta Disagreement_{[t-10, t-1]}$	0.0030** (2.36)	0.0045** (2.46)	0.0036** (2.58)	0.0024 (1.72)
Controls	Yes	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes	Yes
Time FEs	Yes	Yes	Yes	Yes
Observations	5,769,613	5,769,613	5,769,613	5,769,613
Adj. R ²	0.014	0.015	0.015	0.008

This table reports estimation results of manipulative trading by the interaction term between High-Institutional Holding dummy and changes in aggregate social media indicators during the 2-week period prior to manipulation events i.e., [-10,-1]. The dependent variables are a dummy for manipulation, logarithm of number of manipulative transactions plus one, manipulative value and manipulative value ratio on day t in columns (1) - (4), respectively. The key independent variables are the interaction term between High-Institutional Holding dummy and changes in the log number of tweets. All these three variables are calculated during 10-day period prior to the day t i.e. [t-10,t-1]. Control variables include (lagged) firm characteristics and trading indicators, and month, year, individual fixed effects. Standard errors are clustered within firms and years.

Appendix

Table A.1. Description of variables

Variable	Description
Manipul.Dummy _t	A dummy variable for manipulative flag which takes one if there is suspicious transaction alert(s) on a given stock in the NASDAQ market surveillance database on day t , zero otherwise.
Manipul.Count _t	The number of manipulative / suspicious flags on a given stock in day t .
Log.Manipul.Count _t	Logarithm of the number of manipulative / suspicious alerts on a given stock in day t .
Manipul.Value _t	Log of one plus all manipulative / suspicious trade value on a given stock in day t .
Manipul.Val.Ratio _t	Ratio between manipulative trade value over the total daily trading value of a given stock in day t .
N_tweets _t	Numbers of tweets with recommendation keywords containing \$ followed by company stock ticker in day t .
Δ Ntweets _[t-10, t-1]	Log-change in numbers of tweets during the 2-week period from $t-10$ to $t-1$.
Positiveness _t	This is an aggregate measure of the net positiveness in social media on a company in day t . $Positiveness_{i,t} = \frac{N_{i,t,positive} - N_{i,t,negative}}{N_{i,t,positive} + N_{i,t,negative}}$
Δ Positiveness _[t-10, t-1]	Change in the positiveness measurement during the 2-week period from $t-10$ to $t-1$.
Disagreement _t	This measures the degree of disagreement between social media users whether they should be bullish or bearish about a given company in day t . $Disagreement_{i,t} = \sqrt{1 - \left(\frac{N_{i,t,positive} - N_{i,t,negative}}{N_{i,t,positive} + N_{i,t,negative}} \right)^2}$
Δ Disagreement _[t-10, t-1]	Change in the disagreement measurement during the 2-week period from $t-10$ to $t-1$.
Log Market Cap.	Logarithm of market value total
Market-to-Book	Ratio of market to book value
Quick Ratio	Current asset total plus one divided by current liabilities total plus one
Log Revenues	Logarithm of normalised revenues plus one i.e. logarithm of the company's total revenues minus the min value of these revenues during the sampled period plus one.
Return	Logarithm of stock return
Abn.Ret	Abnormal return estimated by the market model using last 6-month data
S&P 500 Return	S&P 500 return
VIX	VIX index
Volatility	Stock volatility estimated using Parkinson (1980) method i.e. Log-difference between high and low scaled by log(2).
Log.Trade.Volum	Logarithm of one plus trading volume on a given stock in day t .
Abn.Trade.Volum	Log.Trade.Volum _t minus the average value of Log.Trade.Volum in the last six months.
Institutional ownership	Percentage of long-term holding
High Institution	High-Institutional ownership dummy variable. This variable takes 1 if the long-term holding proportion is larger than the mean value, which is 15% for our sample, takes 0 otherwise.

Table A.2. Robustness check for Table 3 columns (1) – (2)

	(1)	(2)	(3)	(4)
	Manipul.Dummy _t	dy/dx	Manipul.Count _t	dy/dx
$\Delta \text{Ntweets}_{[t-10, t-1]}$	2.1363*** (104.92)	0.0086*** (6.26)	1.5169*** (8.08)	0.0142*** (8.08)
Controls	Yes		Yes	
Firm FEs	Yes		Yes	
Time FEs	Yes		Yes	
Observations	6163256		6163256	
Pseudo R ²	0.028		0.094	

This table reports estimation results of manipulative trading by changes in aggregate social media indicators during the 2-week period prior to manipulation events i.e., $[t-10, t-1]$. The dependent variables are a dummy for manipulation in columns (1) - (2) and number of manipulative transactions in columns (3) – (4). Columns (1) – (2) use the Logit model while (3) – (4) employ the Poisson model. Columns (1), (3) report the estimated coefficients of the respective models while columns (2), (4) present the corresponding marginal effects. Control variables include (lagged) firm characteristics and trading indicators. Standard errors are clustered within firms and years for the Poisson models. Logit model does not allow for clustered standard errors.

Table A.3. Robustness check for Table 4 columns (1) – (2)

	(1)	(2)	(3)	(4)
	Manipul.Dummy _t	dy/dx	Manipul.Count _t	dy/dx
$\Delta \text{Ntweets}_{[t-10, t-1]}$	2.8588*** (119.09)	0.0120*** (6.02)	2.0645*** (10.38)	0.0137*** (10.38)
$\Delta \text{Positiveness}_{[t-10, t-1]}$	-0.2562*** (-11.24)	-0.0023*** (-5.32)	-0.1960*** (-4.87)	-0.0028*** (-4.87)
$\Delta \text{Disagreement}_{[t-10, t-1]}$	0.0465* (1.81)	0.0006* (1.71)	0.0734 (0.73)	0.0015 (0.73)
Controls	Yes		Yes	
Firm FEs	Yes		Yes	
Time FEs	Yes		Yes	
Observations	5,769,613	5,769,613	5,769,613	5,769,613
Pseudo R ²	0.035		0.107	

This table reports estimation results of manipulative trading by changes in aggregate social media indicators during the 2-week period prior to manipulation events i.e., $[t-10, t-1]$. The dependent variables are a dummy for manipulation in columns (1) - (2) and number of manipulative transactions in columns (3) – (4). Columns (1) – (2) use the Logit model while (3) – (4) employ the Poisson model. Columns (1), (3) report the estimated coefficients of the respective models while columns (2), (4) present the corresponding marginal effects. Control variables include (lagged) firm characteristics and trading indicators. Standard errors are clustered within firms and years for the Poisson models. Logit model does not allow for clustered standard errors.

Table A.4. Robustness check for Table 6 Panel C columns (1) – (2)

	(1)	(2)	(4)	(5)
	Manipul.Dummy _t	dy/dx	Manipul.Count _t	dy/dx
$\Delta \text{Ntweets}_{[t-10, t-1]}$	0.3820*** (30.57)	0.0899*** (6.47)	0.4826*** (4.60)	0.0153*** (4.60)
$\Delta \text{Positiveness}_{[t-10, t-1]}$	-0.1321*** (-6.64)	-0.0311** (-4.69)	-0.2002** (-2.28)	-0.0063** (-2.28)
$\Delta \text{Disagreement}_{[t-10, t-1]}$	-0.0635** (-1.98)	-0.0150* (-1.89)	-0.1454*** (-3.25)	-0.0046*** (-3.25)
Post x $\Delta \text{Ntweets}_{[t-10, t-1]}$	0.0850 (1.37)	0.0194 (1.02)	0.1082 (1.48)	0.0034 (1.48)
Controls	Yes		Yes	
Firm FEs	Yes		Yes	
Time FEs	Yes		Yes	
Observations	644,443	644,443	614,628	614,628
Pseudo R ²	0.024		0.081	

This table reports estimation results of manipulative trading by changes in aggregate social media indicators during the 2-week period prior to manipulation events i.e., $[t-10, t-1]$. The dependent variables are a dummy for manipulation in columns (1) - (2) and number of manipulative transactions in columns (3) – (4). Columns (1) – (2) use the Logit model while (3) – (4) employ the Poisson model. Columns (1), (3) report the estimated coefficients of the respective models while columns (2), (4) present the corresponding marginal effects. Control variables include (lagged) firm characteristics and trading indicators. Standard errors are clustered within firms and years for the Poisson models. Logit model does not allow for clustered standard errors.

Table A.5. Robustness check for Table 8 columns (1) – (2)

	(1)	(2)	(4)	(5)
	Manipul.Dummy _t	dy/dx	Manipul.Count _t	dy/dx
Δ Ntweets _[t-10, t-1]	2.6579*** (92.42)	0.0114*** (6.02)	0.494*** (10.90)	0.0127*** (10.90)
Δ Positiveness _[t-10, t-1]	-0.2538*** (-11.14)	-0.0024*** (-5.31)	-0.110*** (-4.98)	-0.0028*** (-4.98)
Δ Disagreement _[t-10, t-1]	0.0479* (1.87)	0.0006* (1.76)	0.0576 (0.74)	0.0014 (0.74)
High Institution	-0.2312*** (-6.41)	-0.0018*** (-4.31)	-0.0992*** (-6.12)	-0.0025*** (-6.13)
High Institution x Δ Ntweets _[t-10, t-1]	0.3202*** (12.67)	0.0023*** (5.46)	0.112*** (6.51)	0.0029*** (6.51)
Controls	Yes		Yes	
Firm FEs	Yes		Yes	
Time FEs	Yes		Yes	
Observations	5,769,613	5,769,613	5,769,613	5,769,613
Pseudo R ²	0.035		0.107	

This table reports estimation results of manipulative trading by changes in aggregate social media indicators during the 2-week period prior to manipulation events i.e., [t-10, t-1]. The dependent variables are a dummy for manipulation in columns (1) - (2) and number of manipulative transactions in columns (3) – (4). Columns (1) – (2) use the Logit model while (3) – (4) employ the Poisson model. Columns (1), (3) report the estimated coefficients of the respective models while columns (2), (4) present the corresponding marginal effects. Control variables include (lagged) firm characteristics and trading indicators. Standard errors are clustered within firms and years for the Poisson models. Logit model does not allow for clustered standard errors.