

Market manipulation and suspicious stock recommendations on social media

Thomas Renault^{*a,b}

^aIESEG School of Management, Paris, France

^bUniversité Paris 1 Panthéon Sorbonne, Paris, France

14 April, 2018

Abstract

Social media can help investors gather and share information about stock markets. However, it also presents opportunities for fraudsters to spread false or misleading statements in the marketplace. Analyzing millions of messages sent on the social media platform Twitter about small capitalization firms, we find that an abnormally high number of messages on social media is associated with a large price increase on the event day and followed by a sharp price reversal over the next trading week. Examining users' characteristics, and controlling for lagged abnormal returns, press releases, tweets sentiment and firms' characteristics, we find that the price reversal pattern is stronger when the events are generated by the tweeting activity of stock promoters or by the tweeting activity of accounts dedicated to tracking pump-and-dump schemes. Overall, our findings are consistent with the patterns of a pump-and-dump scheme, where fraudsters/promoters use social media to temporarily inflate the price of small capitalization stocks.

Keywords: Asset Pricing, Market efficiency, Market manipulation, Pump-and-dump scheme, Stock promotion, Small capitalization stocks, Social media, Twitter, Event study, Security and Exchange Commission

JEL classification: G12, G14.

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

1. Introduction

Market manipulation is as old as trading on organized exchanges (Putniņš, 2012). However, despite their long prevalence and considerable academic research on the topic, our understanding of the phenomenon is far from adequate. While theoretical models have been developed to address trade-based manipulation (Allen and Gale, 1992) or information-based manipulation (Bommel, 2003), empirical studies continue to be very scarce. This paper contributes to the emerging empirical literature on market manipulation by focusing on a specific type of illegal price manipulation: pump-and-dump schemes.

Pump-and-dump schemes involve touting a company’s stock through false or misleading statements in the marketplace in order to artificially inflate (pump) the price of a stock. Once fraudsters stop hyping the stock and sell their shares (dump), the price typically falls. Although pump-and-dump schemes have existed for many decades, the emergence of the Internet and social media has provided a fertile new ground for fraudsters. False or misleading information can now be disseminated to a large number of potential investors with minimum effort, anonymously, and at a relatively low cost.¹ According to the Security and Exchange Commission (SEC)², “investors who learn of investing opportunities from social media should always be on the lookout for fraud.”

To gain a better understanding of pump-and-dump schemes, we first focus on reported manipulation cases by analyzing all SEC litigation releases published between 1996 and 2015. We construct a database of pump-and-dump frauds, extending previous findings from Aggarwal and Wu (2006). We find that pump-and-dump schemes mainly target small capitalization stocks with low liquidity, also known as “micro-cap” or “penny stocks”, traded in the Over-The-Counter (OTC) market. We also find that stock promoters play a predominant role in pump-and-dump schemes, with nearly 50% of reported manipulation involving a stock promoter. Channels of communication

¹ “Investor Alert: Social Media and Investing - Avoiding Fraud” - Security and Exchange Commission, January 2012

² “Updated Investor Alert: Social Media and Investing - Stock Rumors” - Security and Exchange Commission, November 2015

used by fraudsters to disseminate false information to the marketplace include false or misleading press release, stock spam e-mails, websites, bulletin boards and, in recent years, social media.

While empirical proofs of market manipulation on small capitalization stocks have been identified using data from stock spam (e-mails) recommendations (Böhme and Holz, 2006; Frieder and Zittrain, 2007; Hanke and Hauser, 2008; Nelson et al., 2013) and messages boards (Sabherwal et al., 2011), pump-and-dump schemes on social media have, to the best of our knowledge, never been empirically studied. We extend the literature on indirect empirical evidence of market manipulation by analyzing data from one of the largest worldwide social media platforms: Twitter. To do so, we construct a novel database by collecting, in real-time, all messages posted on Twitter containing the ticker of a list of more than 5,000 small-capitalization stocks. We end up with a total of 7,196,307 tweets posted during a 11-month period (October 2014 to September 2015) by 248,748 distinct users.

We first conduct an event study to analyze the impact of a spike in posting activity on Twitter on the returns of small capitalization stocks. We find that an abnormally high message activity on social media about a company is associated with a sharp increase in abnormal trading volume from two days before the event up to five days after the event. We also find significant positive abnormal returns on the event day and on the day before the event, and a significant price reversal during the next five trading days. The price and volume patterns could be consistent with a pump-and-dump scheme (manipulation hypothesis) but could also be caused by overoptimistic noise traders (behavioral hypothesis) or by news / press releases (overreaction to news). To disentangle between the three hypotheses, we conduct cross-sectional regressions. For each event, we examine users' characteristics to assess if the event was (partially) generated by the tweeting activity of stock promoters or by the tweeting activity accounts dedicated to tracking pump-and-dump-schemes. After controlling for company characteristics (market capitalization, market type, percentage of non-trading days...), press releases, lagged returns and total tweeting activity, we find that the price reversal pattern is stronger when the events are generated by the tweeting activity of stock

promoters and by the tweeting activity accounts dedicated to tracking pump-and-dump-schemes. Overall, our findings favor the manipulation hypothesis over the behavioral hypothesis, and shed light on the need for a higher control of the information published on social media and better education for investors seeking trading opportunities on the Internet.

Our paper is organized as follows. Section 2 briefly reviews the theoretical and the empirical literature on market manipulation. Section 3 examines SEC litigation releases and justifies our focus on OTC stocks and on stock promoters. Section 4 describes our database. Section 5 shows the results of the event study. Section 6 examines the relation between users' characteristics and stock returns, using both contemporaneous and predictive cross-sectional regressions. Section 7 presents some robustness checks. Section 8 concludes.

2. Related literature and hypothesis

Market manipulation undermines economic efficiency both by making prices less accurate as signals for efficient resource allocation and by making markets less liquid for risk transfer (Kyle and Viswanathan, 2008). Despite the importance of fair and transparent markets, little is known about the prevalence and impact of market manipulation (Putniņš, 2012). Theoretical studies have shown that traders can generate profits through trade-based manipulation (Allen and Gale, 1992) or information-based manipulation (Bommel, 2003). However, like any illegal behavior, market manipulation is not directly observable, and empirical studies remain very scarce. Owing to this lack of available data, our first strand of the literature focus on reported manipulation cases.

Studying all cases pursued by the Security and Exchange Commission from January 1990 to October 2001, Aggarwal and Wu (2006) present an extensive review of stock market manipulation in the United States. They find that around 50% of the manipulated stocks are small capitalization stocks (penny stocks) quoted in the OTC markets, such as the OTC Bulletin Board and the Pink Sheets. With regard to techniques used by fraudsters, more than 55% of cases involve spreading rumors or false information. Manipulators also frequently use wash trades and nominee accounts

to create artificial trading activity. More recently, and using a database of 421 German pump-and-dump schemes between 2002 and 2015, Leuz et al. (2017) find that information-based market manipulation are quite common (6% of investors have participated in at least one pump-and-dump in their sample) and involve sizable losses for market participants (average loss of 30%).

However, only a small fraction of manipulation is detected and prosecuted (Comerton-Forde and Putniņš, 2014). Further, focusing on reported cases tends to create a selection bias toward unsophisticated manipulation and is affected by the regulators' agenda (Bonner et al., 1998). Hence, another strand of the literature focuses on indirect evidences by studying abnormal market behaviors (for trade-based manipulation) or by detecting suspicious behaviors outside the market (for information-based manipulation).

Analyzing intraday volume and order imbalance, Ben-Davis et al. (2013) show evidence suggesting that some hedge funds manipulate stock prices on critical reporting dates. Their findings are consistent with those of Carhart et al. (2002) on end-of-quarter manipulation by mutual funds. In line with this study, a nascent strand of the literature focuses on information-based manipulation by analyzing new datasets of stock spams (newsletters) sent by fraudsters trying to pump the value of a stock. Böhme and Holz (2006), Frieder and Zittrain (2007), and Hanke and Hauser (2008) find a significant positive short-run price impact after a stock spam touting, followed by a price reversal over the following days. Similar patterns have been observed when Internet message board activity is used to identify pump-and-dump scheme on small stocks without fundamental news (Sabherwal et al., 2011). In contrast, and using detailed data on the information contained in the spam messages, Nelson et al. (2013) find that stock spams lead to trading activity, but not a median price reaction.

In this paper, we first start by analyzing reported manipulation cases to improve our understanding of pump-and-dump schemes and to justify our focus on small capitalization stocks and on stock promoters. Then, focusing on data from Twitter, we explore the relation between the social media activity and small capitalization stock returns. We hypothesize that a high number

of messages about a company on Twitter should be associated with a contemporaneous increase in the price of the stock, followed by a price reversal over the next trading days. While this pattern could be consistent with market manipulation, it could also simply be caused by overoptimistic sentiment-driven noise trader. We thus hypothesize that if the price pattern is related to market manipulation or to stock promotion (pump-and-dump), the price reversal should be greater when the abnormal activity on Twitter is due to the tweeting activity of stock promoters or to the tweeting activity of accounts tracking pump-and-dump-schemes.

3. SEC litigation

Before extending the literature on indirect empirical evidence by analyzing information-based market manipulation on Twitter, we construct an updated database of SEC civil enforcement actions. Our work is closely related to Aggarwal and Wu (2006) who collect all SEC litigation releases containing keywords related to market manipulation published between 1990 and 2001³ and then manually classify all cases by the type of stocks targeted (listed on NYSE, AMEX, NASDAQ, OTC Markets...) and the type of people involved (insiders, brokers, shareholders...). We complement their findings by (1) extending the sample period, (2) using the new SEC classification, and (3) examining specifically pump-and-dump schemes to identify those involved in frauds (insiders, promoters, traders...) and the tools used to send false or misleading information in the marketplace (press releases, spam e-mails, websites, message boards, social media...).

Since 1996, each enforcement action is classified by the SEC into a unique category and the classification is shared via the “SEC annual reports” and the “Select SEC and Market Data reports”. Our database of SEC litigation releases contains 4,918 civil actions from 1996 to 2015, of which 471 are related to market manipulation, a slightly higher number than in Aggarwal and Wu (2006) for comparable years. Table 1 shows the distribution of SEC civil actions by category and by fiscal

³ More precisely, they search for the keywords “manipulation” and “9(a)” or “10(b)” (which refer to the two articles of the Securities and Exchange Act of 1934). Mei, Wu, and Zhou (2004) use the same list plus the keyword “pump-and-dump”.

year. In the remainder of this paper, we focus our analysis on the category “market manipulation”, which includes the subcategory “newsletter/touting”. Each case is included in only one category following the SEC classification, even though many cases involve multiple allegations and may fall under more than one category.

[**Insert Table 1 about here**]

In total, market manipulations account for 9.60% of all civil actions initiated by the SEC between 1996 and 2015. Over the last two decades, the SEC demonstrated its commitment to prosecuting market manipulation occurring in cyberspace on numerous occasions. For example in October 1998 (fiscal year 1999), the SEC launched a nationwide “sweep” for purveyors of fraudulent spam, online newsletters, message board postings and websites caught in an “effort to clean up the Internet”, which led to 23 enforcement actions against 44 individuals and companies.⁴ In 2000, the fourth nationwide Internet fraud sweep led to 15 enforcement actions against 33 companies and individuals who used the Internet to defraud investors. More recently, in July 2013, the SEC launched the Microcap Fraud Task Force to target abusive trading and fraudulent conduct in securities issued by microcap companies. This announcement was followed by a steep increase in the number of cases related to market manipulation on microcap companies between July and September 2013 (fiscal year 2013).

Even if the absolute number of reported cases is affected by the SEC agenda and may be biased toward poor manipulation, studying the different civil actions can still help us understand which type of stocks are manipulated, who are the people involved and which tools and techniques are used by fraudsters. Since 2002, the SEC has been releasing detailed complaints about a great majority of civil actions they initiate. While litigation releases only summarize the enforcement case in approximately one page, complaint reports provide more details about the fraudulent scheme and the exact role of each defendant, in ten to thirty pages. Using this new report, we manually

⁴ “SEC Conducts First Ever Nationwide Internet Securities Fraud Sweep, Charges 44 Stock Promoters in 23 Enforcement Actions” - <https://www.sec.gov/news/press/pressarchive/1998/98-117.txt>

analyze all complaints classified as “market manipulation” or “newsletter/touting”. From the 362 “market manipulation” civil actions initiated by the SEC between 2002 and 2015, we managed to collect detailed complaint reports for 273 cases, of which 150 are related to pump-and-dump schemes. Table 2 summarizes, year by year, the type of stocks targeted by fraudsters, the type of people involved in the manipulation schemes, and the tools used to disseminate false or misleading information in the marketplace.

[**Insert Table 2 about here**]

We find that 86% of pump-and-dump schemes target stocks traded on OTC markets. The most common channel of communication⁵ used by fraudsters to send false or misleading information in the marketplace is press releases (73.3%), followed by spam e-mails/newsletters (34%), websites (32%), fax blast (12.6%), and message boards (10.6%). Prosecuted cases mostly target frauds performed by company insiders (CEO, CFO) (60.7%), and/or by stock promoters paid in cash or in shares to pump the price of a stock⁶ (49.3%), and/or by traders/shareholders (37.3%).

We find one case where fraudsters specifically use Twitter to conduct a pump-and-dump scheme (see Appendix A). In 2009, a Canadian couple used their website (PennyStockChaser), Facebook, and Twitter to pump up the stock of microcap companies and sold their shares after the pump. Given this recent case and the SEC’s renewed attention toward risks created by social media communication, we explore in the remainder of this paper the relation between the content published on social media and small-capitalization stock returns.

⁵ The sum is not equal to 100% as fraudsters often combine multiple channels of communication to increase the outreach and visibility of their messages.

⁶ Stock promotion (investor relation) are not illegal per se. If promoters provide full disclosure of their compensation (type, amount, person paying the compensation) in all their communication, and if the information provided is neither false nor misleading, stock promotion can be legal. The Securities Act of 1933, Section 17(b) states the following: “It shall be unlawful for any person, by the use of any means or instruments of transportation or communication in interstate commerce or by the use of the mails, to publish, give publicity to, or circulate any notice, circular, advertisement, newspaper, article, letter, investment service, or communication which, though not purporting to offer a security for sale, describes such security for a consideration received or to be received, directly or indirectly, from an issuer, underwriter, or dealer, without fully disclosing the receipt, whether past or prospective, of such consideration and the amount thereof”.

4. Data

4.1. *The OTC Markets Group*

On the basis our preliminary analysis into SEC litigation, we choose to focus on stocks quoted by the OTC Markets Group. The OTC Markets Group is an electronic inter-dealer quotation and trading system providing marketplaces for around 10,000 OTC securities. The OTC Markets Group organizes securities into three tiered marketplaces: OTCQX, OTCQB, and OTC Pink. The marketplace on which a company trades reflects the integrity of its operations, its level of disclosure, and its degree of investor engagement.

- OTCQX marketplace: Companies must meet high financial standards, be current in their disclosure, and receive third party advisory.
- OTCQB marketplace: Companies must be current in their reporting, meet a minimum bid test of \$0.01, and undergo an annual verification and management certification process.
- OTC Pink marketplace: Open to all companies. The OTC Pink is divided into three sub-categories based on the quantity and quality of information provided to investors: current information, limited information, and no information.

We download the list of all Common Stock and Ordinary Shares of companies incorporated in the United States, excluding American Depository Receipts, ETF, Funds, and Warrants. Our sample consists of 5,087 companies: 61 (1.20%) are quoted on OTCQX, 1,858 (36.52%) on OTCQB, and 3,168 (62.28%) on OTC Pink. Among the companies listed on OTC Pink, 814 provide current information, 403 provide limited information, and 1,951 provide no information. Companies in the last category should, according to the OTC Markets Group “be treated with suspicion and their securities should be considered highly risky.”

We use Bloomberg to download daily price data, traded volume data, and market capitalization for all 5,087 stocks. During the sample period, the vast majority of the stocks experienced a sharp

decrease in price with a number of stocks losing nearly all their value. This finding is consistent with the finding reported by Ang et al. (2013) that over a long period, comparable listed-stocks tend to overperform OTC stocks by nearly 9% per year. However, a few stocks also showed impressive returns over the sample period. For example, the price of Micro Imaging Technology increased from \$0.0229 to \$0.45 between October 2014 and October 2015 (+1,865%). As documented by Eraker and Ready (2015), the returns of OTC stocks are negative on average and highly positively skewed, with a few “lottery-like” stocks doing extremely well while many of the stocks become worthless.

4.2. *Twitter data*

Twitter is a micro-blogging platform that enables users to send and read short 140-character messages called “tweets”. Every day, more than 500 million messages are posted on Twitter. We develop a computer program in the Python programming language to collect data in real time using the Twitter Search Application Programming Interface (API). Precisely, using the Twitter “cashtag” feature, introduced in 2012, we extract all the messages containing a “\$” sign followed by the ticker name, as in Sprenger et al. (2014b). As in Sprenger et al. (2014a), we query the API with each keyword (company cashtag) every hour to ensure that we collect all tweets during our sample period.

In the course of our sample period from October 5, 2014, to September 1, 2015, we collected a total of 7,196,307 tweets. Among the 5,087 companies, around 50% received a very low level of attention (between 0 and 20 tweets). On the other hand, four companies featured in more than 100,000 tweets: Tykhe Corp (*\$HALB*), Cardinal Energy Group (*\$CEGX*), Sterling Consolidated (*\$STCC*) and Arrayit Corp (*\$ARYC*). Table 3 presents descriptive statistics for the top 10 most discussed companies in the sample period. Overall, we find that Twitter activity is higher for companies listed on the OTC Pink marketplace, with a low stock price (penny stocks) and a small market capitalization.

[**Insert Table 3 about here**]

By analyzing the Twitter messages for the ten most discussed companies in our sample, we identified many fake Twitter accounts (bots) posting exactly the same type of messages at different periods, simply by replacing a ticker with another and changing a few keywords over time. After a certain period of abnormally high posting activity, the number of tweets reduced to a level close to zero. While it is difficult to ascertain if those bursts in social media activity are directly linked with attempts to manipulate the market, the use of multiple fake accounts to recommend buying a stock is at least suspicious.

The case of Wholehealth Products, Inc. (*\$GWPC*), the eight most-discussed stock in our sample, is especially interesting. On November, 20, 2014, the Security Exchange Commission suspended trading on *GWPC* because of concerns regarding the accuracy and adequacy of publicly disseminated information by the company, including information about the relationship between the company's business prospects and the current Ebola crisis.⁷ By examining the number of messages containing the ticker *\$GWPC* posted on Twitter before the SEC halt, we identify a sharp increase in posting activity starting on October 27th (Figure 1). A total of 2,774 tweets were sent on that day, compared to an average of less than 30 messages per day on the week before. The spike in posting activity on Twitter was followed by a one-week increase in stock price and a sharp price reversal afterward.

[**Insert Figure 1 about here**]

This anecdotal example is typical of a pump-and-dump scheme. A false piece of information is shared on Twitter to generate a spike in the social media activity about a given company. Stock price increases (pump) over a short period, and decreases sharply (dump) afterward. In the next section, we conduct an event study to analyze if the price reversal pattern identified anecdotally in the *\$GWPC* case can be generalized. We do so by analyzing the link between an abnormally high activity on social media and OTC stocks returns.

⁷ "SEC Suspends Trading in Companies Touting Operations Related to Prevention or Treatment of Ebola", November 20, 2016

5. Event study

Following Tumarkin and Whitelaw (2001) and Leung and Ton (2015), we define event days as follows: when the number of messages posted on Twitter about company i during a given day t exceeds the yearly average plus two standard deviations. To account for the regular operational hours of the exchanges, we consider all messages sent between 4 p.m. on day $t-1$ and 4 p.m. on day t as pertaining to day t . To reduce error introduced by stocks with very low tweeting activity or very illiquid stocks, we impose as event criteria a minimum of 20 tweets⁸ and a minimum market capitalization of \$1,000,000. If an event is detected on a non-trading day, we consider the next trading day as the event day. We also impose a minimum of 20 trading days between two events to avoid contagion on the event window.

The following example illustrates our methodology using a specific company: UBIQ Inc, ($UBIQ$). During the sample period, a total of 1,048 messages containing the ticker $UBIQ$ were posted on Twitter. We identify two events for $UBIQ$ company: on February 6, 2015 (385 messages) and August 27, 2015 (53 messages). Table 4 shows a sample of tweets related to February 6, 2015, event. The activity on Twitter on that day is typical of a stock promotion scheme, where tweets are sent by bots and stock promoters through multiple accounts. Examining manually the 47 distinct users who sent tweets on that day, we identify tweets sent by stock promoters (i.e., @JetPenny, from the website <http://www.jetlifepennystocks.com/>, and @WallStreetPenni, from the website <http://www.wallstreetpennies.com/>) and tweets sent by users tracking pump-and-dump schemes (@PUMPSandDumps and @ThePumpTracker). This example is typical of a pump-and-dump scheme, where fraudsters or stock promoters try to temporarily inflate the price of a small capitalization stocks by using Twitter as a new channel of communication. Two years later, on March 20, 2017, the Security and Exchange Commission temporarily suspended trading in the securities of Ubiquity due to “a lack of current and accurate information about the company because

⁸ Tumarkin and Whitelaw (2001) and Leung and Ton (2015) impose a minimum of 10 message postings. We impose a higher threshold as the level activity is much higher on Twitter than on messages boards. Our results are robust to the threshold used.

Ubiquity is delinquent in its requisite periodic filings with the Commission.”⁹

[**Insert Table 4 about here**]

Using this methodology for all stocks in our sample, we identify a total of 635 events for 315 distinct companies. Then, for each event detected, we compute abnormal return for the estimation window [-260:-11] (L1 = 250 days) and event window [-10:+10] (L2 = 21 trading days).¹⁰ To define abnormal return, we consider three models of normal return: a market return model, a constant mean return model, and a capital asset pricing model. We use the NASDAQ MicroCap Index as a benchmark of market return. We also consider raw returns, as in Nelson et al. (2013). For clarity, we will report our results only for the market return model, as we find that our results are robust to all models of normal return.¹¹

We test the significance of abnormal return during the event window by conducting a non-parametric Corrado (1989) rank test, making no assumption about the normality of the underlying data. We transform each abnormal return $AR_{i,t}$ to a rank variable $K_{i,t}$, by assigning to the day with the highest return over the complete window (estimation and event window) a rank of +271, to the day with the second highest return a rank of +270, and so on until we assign the lowest return a rank of 1. Tied ranks are treated by the method of midranks. To allow for missing returns, ranks are standardized by dividing by one plus the number of non-missing returns.

$$K_{i,t} = \frac{rank(AR_{i,t})}{(1 + M_i)} \quad (1)$$

where M_i is the number of non-missing values for security i in L1 and L2. This yields order statistics for the uniform distribution with an expected value of one-half. The rank test statistic

⁹ <https://www.sec.gov/litigation/suspensions/2017/34-80275.pdf>

¹⁰ On unreported robustness check, we find that our results are robust to a 6-month estimation window and a 11-day event window.

¹¹ The bad-model problem is less serious for event study with short window as daily expected returns are close to zero (Fama, 1998).

for day t (T_t) is equal to

$$T_t = \frac{1}{\sqrt{N}} \sum_{i=1}^N (U_{i,t} - 0.5) / S(U) \quad (2)$$

where N is equal to the number of events. The estimated standard deviation $S(U)$ is defined on the estimation ($L1$) and event ($L2$) window as¹²

$$S(U) = \sqrt{\frac{1}{L_1 + L_2} \sum_t \left[\frac{1}{\sqrt{N_t}} \sum_{i=1}^{N_t} (U_{i,t} - 0.5) \right]^2} \quad (3)$$

where N_t represents the number of non-missing returns in the cross-section of N -firms on day t .

We test the statistical significance of abnormal return on each day of the event window and on each 5-day rolling interval to identify a price reversal over a one week period. We also consider abnormal trading volume, defined as volume on a given day divided by the average volume on the estimation window. Figure 2 presents abnormal return (AR) and cumulative abnormal return (CAR) during the $[-10:+10]$ event window, where day 0 is defined as a day of abnormally high activity on Twitter. Figure 3 presents abnormal volume (AV). Tables 5 and 6 summarize the results.

[**Insert Figure 2, Figure 3, Table 5 and Table 6 here**]

As in Kim and Kim (2014), we identify a strong contemporaneous relationship between Twitter activity and stock price on the event day ($t0$). We find a significant abnormal return of +4.10% on the day before the event and a significant abnormal return of +6.88% on the the event day. This finding is consistent with that of Sabherwal et al. (2011) who reported an increase of +13.93% on the event day (and +4.91% on the day before the event), when an event was defined as an abnormal number of messages on the financial message board “TheLion.com”. According to Sabherwal et al. (2011), the abnormal return on the day before the event suggests a two-day price momentum, which is consistent with a pre-event two-day pumping hypothesis. While we acknowledge that

¹² We also consider a multi-day version by multiplying by the inverse of the square root of the period’s length.

price momentum is a plausible explanation, this pattern could also be related to a lagged reaction of the tweeting activity. For example, we find some cases where users on Twitter start talking about a stock that has experienced a strong price increase on day t (from 9.30 a.m. to 4 p.m.) only after market close (after 4 p.m.). In that situation, our methodology will identify the event on day $t+1$, leading to an abnormal return on the day before the event.

More interestingly, we find a significant post-event price reversal. Cumulative abnormal return is statistically significant and negative during the five days following the event ($[+1:+5]$ window), with a post-event cumulative decrease in abnormal return of -3.11%. This finding is in line with Sabherwal et al. (2011) who observed a significant post-event decrease in stock price of -5.4% over the five trading days following the event day. Three non-exclusive hypotheses can explain the price reversal pattern and the deviation from the efficient market hypothesis. First, we conjecture that social media can be used as a proxy of investor overoptimism. In a market driven by unsophisticated traders with limits to arbitrage, price can deviate temporarily from its fundamental values in the presence of irrational sentiment-driven noise traders. In such a case, the price reversal identified on OTC stocks is simply caused by “standard” investor sentiment, as explained by Tetlock (2007). Second, the price reversal could be driven by an overreaction to news. For example, if the increase in tweeting activity is contemporaneous to the publication of a press release by a company, then the increase on the event day can be caused by the news, and the price reversal could be related to an overreaction to the news on the event day. Last but not least, the sharp price increase on the event day might be caused by fraudsters or stock promoters pumping the price of targeted stocks, before dumping it on the following days after having made an illegal profit.

To partially isolate one hypothesis from the other, we conduct cross-section contemporaneous and predictive cross-sectional regressions, controlling for stock characteristics, lagged returns, sentiment and press releases.

6. Cross-Sectional Regressions

In this section, we examine if the tweeting activity can (1) explain the increase in contemporaneous return, and, more importantly, (2) predict the price reversal on the week following the event. To examine if price changes are rooted in manipulative promotion rather than over-optimism by investors, we test the three following hypotheses.

- The price increase on the event day is higher when at least one tweet was sent by a stock promoter. This would be consistent with stock promoters' activity pumping the price of a stock on the event day by sending false or misleading information on Twitter.
- The price reversal after the event is higher when at least one tweet was sent by a stock promoter on the event day. This would be consistent with stock promoters selling the stock at an artificially inflated prices after the pumping period, and stock price reverting to its fundamental value.
- The price increase on the event day is lower when at least one tweet was sent by a user tracking pump-and-dump schemes on the event day. This would be consistent with pump-and-dump trackers' activity mitigating the price increase on the event day by alerting users on Twitter of a potential pump-and-dump scheme.

Regarding the first hypothesis, we define manually a list of 156 stock promoters / paid advertisers by analyzing the tweeting activity of all users from our database with a minimum of 100 tweets (7,069 users). We define as a stock promoter all users with a reference to a promotion website or a promotional newsletter in their description. We also analyze all tweets containing the following keywords: “newsletter”, “paid”, “tout”, “promote”, “promotion”, “promoter”, “compensate”, “compensation”, “advert”, “advertiser”, “disclaimer”, “disclosure”, “investor relation”, and we manually flag all stock promoters.¹³ For each event, we create a dummy variable denoted

¹³ The list is available upon request.

$StockPromoter_t^i$ equal to 1 if at least one tweet was sent by a stock promoter from our list, and equal to 0 otherwise. Considering all events in our sample (635), we identify at least one stock promoter for 403 events (63.46%).

Examining users' characteristics, we also identify two users whose tweeting activity is dedicated to tracking pump-and-dump schemes: “@ThePumpTracker” and “@PUMPSandDUMPS”. Those two accounts are, according to their own description, dedicated to “track pump and dump’s, pump promotions, chatroom pumps, scams, ICO scams, crypto scams” and publish a message on Twitter to alert individual investors every time they suspect that a stock is under manipulation. For each event, we create a dummy variable denoted $PumpTracker_t^i$ equal to 1 if at least one tweet was sent by @ThePumpTracker or by @PUMPSandDUMPS on the event day, and equal to 0 otherwise. Considering all events in our sample (635), we identify at least one user tracking pump-and-dump schemes for 57 events (8.98%)

The contemporaneous relation is tested using the following model, considering each of the 635 events defined previously as an observation:

$$\begin{aligned}
AR_t^i = & \alpha + \beta_1 AR_{t-1}^i + \beta_2 MarketCap_t^i + \beta_3 Price_t^i + \beta_4 NonTradingDays_t^i \\
& + \beta_5 MarketType_t^i + \beta_6 Sentiment_t^i + \beta_7 News_t^i + \beta_8 NumberMessages_t^i \\
& + \beta_9 StockPromoter_t^i + \beta_{10} PumpTracker_t^i + \beta_{11} IRAccount_t^i + \epsilon_t
\end{aligned} \tag{4}$$

where, for each observation i , AR_t^i is the abnormal return on the event day, AR_{t-1}^i is the abnormal return on the day before the event, $MarketCap_t^i$ is the market capitalization at the beginning of the event window, $Price_t^i$ is the price of the stock at the beginning of the event window, $NonTradingDays_t^i$ is the percentage of non-trading days during the estimation window (percentage of days with zero trading volume), $MarketType_t^i$ is a dummy variable equal to 1 if the company is listed on the “Limited & No Information” OTC Pink marketplace, $News_t^i$ is a dummy variable equal to 1 if a press release was sent on the event day or on the day before the event,

$Sentiment_t^i$ is the average sentiment of messages sent during the event day, $NumberMessages_t^i$ is the total number of message sent during the event day, $StockPromoter_t^i$ is a dummy variable equal to 1 if at least one tweet was sent on the event day by an account on the “stock promoter list” defined previously, $PumpTracker_t^i$ is a dummy variable equal to 1 if at least one tweet was sent on the event day by an account dedicated to tracking pump-and-dump scheme, and $IRAccount_t^i$ is a dummy variable equal to 1 if the company had an official “Investor Relation” Twitter account at the date of the event.

Press releases are extracted from the “OTC Disclosure & News Service” on the OTC Market Group website. Sentiment is computed using Renault (2017) field-specific lexicon, derived from a large database of 750,000 classified messages published on the microblogging platform StockTwits. The lexicon contains a list of 543 positive terms and a list of 768 negative terms. Table 7 presents descriptive statistics for all variables. We find that 117 events are related to a day with a press release (18.43%). The 518 other event days (81.57%) are days without any fundamental news, as in the framework of Sabherwal et al. (2011). Regarding sentiment, we find that the average sentiment is positive for 539 events (84.88%) and negative or neutral for only 96 events (15.12%). As already documented in the literature (see, e.g., Kim and Kim, 2014; Avery et al., 2016), online investors are mostly bullish when sharing information about stock market on the Internet. Individual investors do not (typically) sell short, hold small portfolios and are net-buyer of attention-grabbing stocks (Barber and Odean, 2008). Thus, when individual investors talk about a stock on the Internet, they tend to post messages mainly about the stock they hold or the stock they want to buy using a bullish (positive) vocabulary. In this investigation, the bullishness bias can also be viewed as fraudsters trying to pump the price of a stock by sharing (false) positive information about a given company on social media.

Table 8 reports the results of the contemporaneous regression (Equation 4). We find that abnormal returns on the event day are positively related with social media sentiment and negatively related to market capitalization. These findings are similar to those of Sabherwal et al. (2011):

same-day stock returns are higher for stocks with higher sentiment and smaller size. We also find that the percentage of non-trading days is negatively related to abnormal returns on the event day: returns are on average higher for illiquid penny stocks. Abnormal return on the event day are lower for stocks that have experienced higher abnormal return on the day before the event. As the coefficient on *News* is not significantly different from zero, and as we demonstrate in the previous section that returns on the day before the event are large and significant, this finding could be related to the results of Savor (2012) who find that, after major price changes, no-information price events experience strong reversals. This situation would also be consistent with a lagged reaction of users on Twitter, as discussed previously.

We also find that adding variables related to user’s tweeting activity significantly improve the accuracy of the model. On one hand, the tweeting activity of stock promoters is positively related to abnormal return on the event day, consistent with the “pump” period of a pump-and-dump scheme. On the other hand, the tweeting activity of users tracking pump-and-dump scheme is negatively related to abnormal return on the event day. When individuals investors are alerted that a stock might be under promotion, the impact on stock price on the event day is lower. Those results confirm our hypothesis of a “pumping” period associated with an abnormal tweeting activity on social media.

To examine if a “dumping” period follows the “pumping” period identified on the event day, we examine the predictive relation by using a model similar to Equation 4, replacing AR_t by $CAR_{t+1,t+n}$, where $n = 1, 2, \dots, 10$. Table 9 reports the results of the predictive regressions (Equation 5).

$$\begin{aligned}
CAR_{t+1,t+n}^i &= \alpha + \beta_1 AR_t^i + \beta_2 MarketCap_t^i + \beta_3 Price_t^i + \beta_4 NonTradingDays_t^i \\
&+ \beta_5 MarketType_t^i + \beta_6 Sentiment_t^i + \beta_7 News_t^i + \beta_8 NumberMessages_t^i \\
&+ \beta_9 StockPromoter_t^i + \beta_{10} PumpTracker_t^i + \beta_{11} IRAccount_t^i + \epsilon_t
\end{aligned} \tag{5}$$

We find that on average, the price reversal is higher for stocks that have experienced a strong price increase on the event day, consistent again with Savor (2012). Other variables, such as market capitalization, sentiment or press releases, are not significant and do not help predicting cumulative abnormal returns after the event day. Consistent with a pump-and-dump scheme, we find that the price reversal pattern is significantly stronger when the events are generated by the tweeting activity of stock promoters or by the tweeting activity of accounts dedicated to tracking pump-and-dump schemes. This finding is true for $1 \leq n \leq 10$. We also find that the price reversal is significantly lower (at the 10% confidence level) for firms with corporate social media accounts. This result is consistent with those in Blankespoor et al. (2013); firms with corporate social media accounts might be able to control the spread of fake news and thus reduce information asymmetry.

Overall, all those results favor the manipulation/promotion hypothesis over the behavioral hypothesis.

7. Robustness Tests

To assess the robustness of our results, we first conduct another event study by splitting our 635 events into “Promoter Event” and “No Promoter Event”. We do so to analyze if the price reversal identified in Section 5 totally disappears for “No Promoter Event”, which would be consistent with the efficient market hypothesis. Then, we examine if the number of retweets or the number of tweets with an external links improve the accuracy of our model. We do so to examine if the tweeting activity by fake accounts (bots) affects the phenomenon, as bots mostly publish messages with an external links to their own website, while traders do not (and bots very rarely retweets messages from other users, while traders do). Last, and following a SEC Investor Alert warning investors about potential risks involving investments in marijuana-related companies in May 2014¹⁴, we analyze if our events are associated with marijuana-related companies and we examine the pattern of abnormal returns around events related to those companies.

¹⁴ https://www.sec.gov/oiea/investor-alerts-bulletins/ia_marijuana.html

As shown in the previous section, the price reversal is higher when the events are generated by the tweeting activity of stock promoters or by the tweeting activity of accounts dedicated to tracking pump-and-dump schemes. Thus, we conduct two event studies, using the same methodology as previously, but considering (1) all events without any tweets by a stock promoter nor by a user tracking pump-and-dump schemes (228 events - “No Promoter Event”), (2) all events with at least one tweet by a stock promoter or one tweet from a user tracking pump-and-dump schemes (407 events - “Promoter Event”). We do not identify any significant price reversal when we consider “No Promoter” events. We even find the opposite pattern, as abnormal return increases on average by 0.87% on the five days following the event. On the other hand, we find a very strong and significant price reversal when we focus on “Promoter Event”. Abnormal return decreases on average by 5.34% on the five days following the event. Table 10 presents the results.

We also analyze (1) if tweets containing links are more/less credible than tweets without links, (2) if tweets with a lot of retweets are more/less credible than other tweets. We denote $PctRetweet$ the number of retweets divided by the total number of tweets sent on the event day, and $PctLink$ the number of tweets containing a link divided by the total number of tweets sent on the event day. We also consider the number of distinct users and the number of tweets containing a mention. None of these variable improve the accuracy of the model, neither on the contemporaneous regression nor on the predictive regression. Table 11 presents the results.

Last, we examine if the abnormal return on the event day and the price reversal on the next trading week is more pronounced for stocks that have been classified “at risk” by the Security and Exchange Commission. On May 7, 2014, the SEC has released an investor alert to make investors aware about the potential risks of investments involving Bitcoin and other forms of virtual currency. On May 16, 2014, another investor alert was released about the risks of marijuana-related companies, and on November 20, 2014, about companies that claim their products or services relate to Ebola. We create three dummy variables $Bitcoin_t$, $Marijuana_t$ and $Ebola_t$, equal to 1 if at

least one tweet contains a keyword related to this topic¹⁵. We find that, while “bitcoin stocks” and “ebola stocks” only represent a minor percentage of the events (respectively 1.1% and 2.52%), the number of events related to “marijuana stocks” is very large (122 events out of a total of 635). Adding the variable $Marijuana_t$ to our model (we do not add $Bitcoin_t$ and $Ebola_t$, as the number of observations were too low), we find that the price reversal is significantly higher (at the 10% level) for marijuana-related stocks. Table 11 presents the results.

The two variables $StockPromoter$ and $PumpTracker$ are significant in all our robustness checks. Overall, our findings are consistent with the patterns of a pump-and-dump scheme, where fraudsters/promoters use social media to temporarily inflate the price of small capitalization stocks.

8. Conclusion

Social media can help investors gather and share information about stock markets. However, it also presents opportunities for fraudsters to send false or misleading statements in the marketplace. In that regard, the social media platform Twitter is a very attractive channel for manipulators or stock promoters as it allows them to target a wide unsophisticated audience, more prone to being scammed than sophisticated investors. The anonymity of Twitter and the ease with which fake accounts and/or bots can be used to spam the network also facilitate fraudsters’ activities.

In this paper, we provide, to the best of our knowledge, the first empirical evidence showing that fraudsters can use social media to artificially inflate the price of a stock. Defining an event as an abnormally high posting activity on Twitter about a company, we identify a large increase in stock price on the event day, followed by a sharp price reversal over the next five trading days. Examining users’ characteristics, and controlling for lagged abnormal returns, press releases, sentiment and firms’ characteristics, we find that the price reversal pattern is stronger when the events are generated by the tweeting activity of stock promoters or by the tweeting activity of

¹⁵ More precisely, we consider the root “bitcoin” and “crypto” for the variable $Bitcoin_t$, “marijuana” and “cannabis” for $Marijuana_t$, and “ebola” for $Ebola_t$

accounts dedicated to tracking pump-and-dump schemes.

While a judicial inquiry would be needed to assess if the promotion scheme were legal or not, our findings shed light on the need for higher control over the information published on social media and better education for investors seeking trading opportunities on the Internet. Given the risk of manipulation and the average negative return of OTC stocks, individual investors should be very cautious when choosing to invest on risky and illiquid small capitalization stocks.

Appendix A

U.S. SECURITIES AND EXCHANGE COMMISSION - Litigation Release No. 21580 / June 29, 2010

Securities and Exchange Commission v. Carol McKeown, Daniel F. Ryan, Meadow Vista Financial Corp., and Downshire Capital, Inc., Civil Action 10-80748-CIV-COHN (S.D. Fla. June 23, 2010)

The Securities and Exchange Commission announced today that it has obtained an emergency asset freeze against a Canadian couple who fraudulently touted penny stocks through their website, Facebook and Twitter. The SEC also charged two companies the couple control and obtained an asset freeze against them. According to the SEC's complaint, the defendants profited by selling penny stocks at or around the same time that they were touting them on www.pennystockchaser.com. The website invites investors to sign up for daily stock alerts through email, text messages, Facebook and Twitter.

The SEC alleges that since at least April 2009, Carol McKeown and Daniel F. Ryan, a couple residing in Montreal, Canada, have touted U.S. microcap companies. According to the SEC's complaint, McKeown and Ryan received millions of shares of touted companies through their two corporations, defendants Downshire Capital Inc., and Meadow Vista Financial Corp., as compensation for their touting. McKeown and Ryan sold the shares on the open market while PennyStockChaser simultaneously predicted massive price increases for the issuers, a practice known as "scalping." The SEC's complaint, filed in the U.S. District Court for the Southern District of Florida, also alleges McKeown, Ryan and one of their corporations failed to disclose the full amount of the compensation they received for touting stocks on PennyStockChaser. The SEC alleges that McKeown, Ryan and their corporations have realized at least \$2.4 million in sales proceeds from their scalping scheme.

The SEC's complaint charges McKeown, Ryan, Downshire Capital Inc. and Meadow Vista Financial Corp. with violating Section 17(a) of the Securities Act of 1933, Section 10(b) of the Securities Exchange Act of 1934, and Rule 10b-5 thereunder. The SEC's complaint also charges McKeown, Ryan and Meadow Vista Financial Corp. with violating Section 17(b) of the Securities Act of 1933. In addition to the emergency relief already granted by the U.S. District Court the Commission also seeks a preliminary injunction and permanent injunction, along with disgorgement of ill-gotten gains plus prejudgment interest and the imposition of a financial penalty, penny stock bars against the individuals and the repatriation of assets to the United States.

In the course of its investigation, the SEC worked with the Quebec Autorit des marches financiers (AMF), which was also investigating this matter. As a result of both ongoing investigations, the AMF obtained an emergency order freezing assets and a cease trade order against McKeown, Ryan, Downshire Capital Inc. and Meadow Vista

Financial Corp. The SEC appreciates the collaboration with the AMF. The SEC's case was investigated by Michael L. Riedlinger, Timothy J. Galdencio and Eric R. Busto of the Miami Regional Office. The SEC's litigation effort will be led by Christine Nestor, Amie R. Berlin and Robert K. Levenson. The SEC's investigation is continuing.

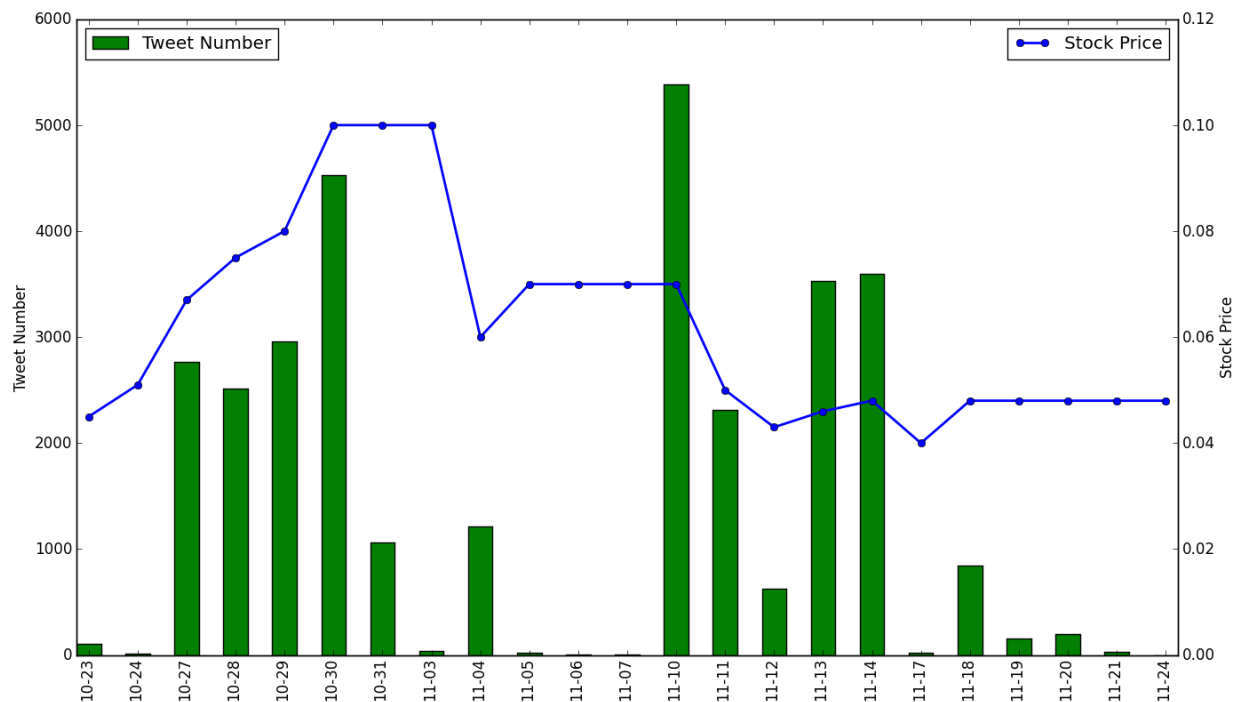
References

- Aggarwal, R., Wu, G., 2006. Stock market manipulations. *The Journal of Business* 79, 1915–1954.
- Allen, F., Gale, D., 1992. Stock-price manipulation. *Review of Financial Studies* 5, 503–529.
- Ang, A., Shtauber, A. A., Tetlock, P. C., 2013. Asset pricing in the dark: The cross-section of OTC stocks. *Review of Financial Studies* 26, 2985–3028.
- Avery, C. N., Chevalier, J. A., Zeckhauser, R. J., 2016. The “CAPS” prediction system and stock market returns. *Review of Finance* 20, 1363–1381.
- Barber, B. M., Odean, T., 2008. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies* 21, 785–818.
- Ben-Davis, I., Franzoni, F., Landier, A., Moussawi, R., 2013. Do hedge funds manipulate stock prices? *The Journal of Finance* 68, 2383–2434.
- Blankespoor, E., Miller, G. S., White, H. D., 2013. The role of dissemination in market liquidity: Evidence from firms’ use of Twitter. *The Accounting Review* 89, 79–112.
- Böhme, R., Holz, T., 2006. The effect of stock spam on financial markets. Working Paper, Technische Universität Dresden .
- Bommel, J. V., 2003. Rumors. *The Journal of Finance* 58, 1499–1520.
- Bonner, S. E., Palmrose, Z.-V., Young, S. M., 1998. Fraud type and auditor litigation: An analysis of SEC accounting and auditing enforcement releases. *Accounting Review* 73, 503–532.
- Carhart, M. M., Kaniel, R., Musto, D. K., Reed, A. V., 2002. Leaning for the tape: Evidence of gaming behavior in equity mutual funds. *The Journal of Finance* 57, 661–693.
- Comerton-Forde, C., Putniņš, T. J., 2014. Stock price manipulation: Prevalence and determinants. *Review of Finance* 18, 23–66.

- Corrado, C. J., 1989. A nonparametric test for abnormal security-price performance in event studies. *Journal of Financial Economics* 23, 385–395.
- Eraker, B., Ready, M., 2015. Do investors overpay for stocks with lottery-like payoffs? An examination of the returns of OTC stocks. *Journal of Financial Economics* 115, 486–504.
- Fama, E. F., 1998. Market efficiency, long-term returns, and behavioral finance. *Journal of financial economics* 49, 283–306.
- Frieder, L., Zittrain, J., 2007. Spam works: Evidence from stock touts and corresponding market activity. *Hastings Communications and Entertainment Law Journal* 30, 479.
- Hanke, M., Hauser, F., 2008. On the effects of stock spam e-mails. *Journal of Financial Markets* 11, 57–83.
- Kim, S.-H., Kim, D., 2014. Investor sentiment from Internet message postings and the predictability of stock returns. *Journal of Economic Behavior & Organization* 107, 708–729.
- Kyle, A. S., Viswanathan, S., 2008. How to define illegal price manipulation. *The American Economic Review* 98, 274–279.
- Leung, H., Ton, T., 2015. The impact of Internet stock message boards on cross-sectional returns of small-capitalization stocks. *Journal of Banking & Finance* 55, 37–55.
- Leuz, C., Meyer, S., Muhn, M., Soltes, E., Hackethal, A., 2017. Who falls prey to the Wolf of Wall Street? Investor participation in market manipulation. Working Paper 24083, National Bureau of Economic Research.
- Mei, J., Wu, G., Zhou, C., 2004. Behavior based manipulation: theory and prosecution evidence. Working Paper, New York University .

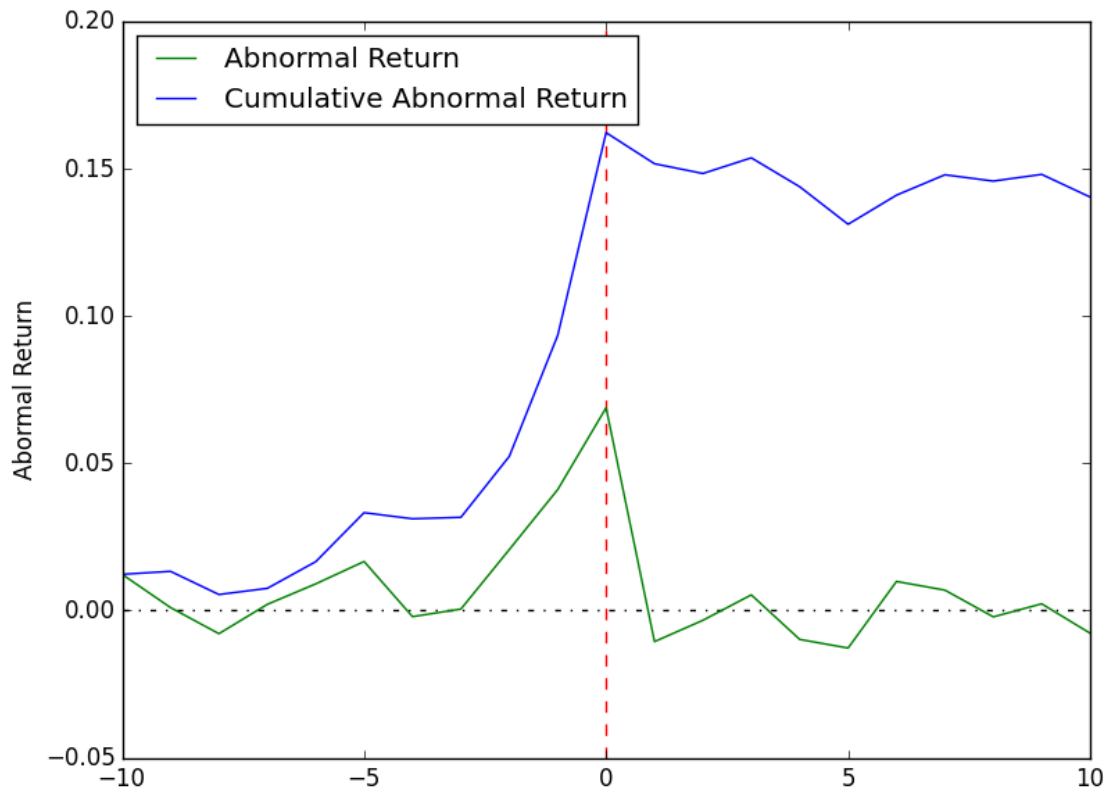
- Nelson, K. K., Price, R. A., Rountree, B. R., 2013. Are individual investors influenced by the optimism and credibility of stock spam recommendations? *Journal of Business Finance & Accounting* 40, 1155–1183.
- Putniņš, T. J., 2012. Market manipulation: a survey. *Journal of Economic Surveys* 26, 952–967.
- Renault, T., 2017. Intraday online investor sentiment and return patterns in the US stock market. *Journal of Banking & Finance* 84, 25–40.
- Sabherwal, S., Sarkar, S. K., Zhang, Y., 2011. Do Internet stock message boards influence trading? Evidence from heavily discussed stocks with no fundamental news. *Journal of Business Finance & Accounting* 38, 1209–1237.
- Savor, P. G., 2012. Stock returns after major price shocks: The impact of information. *Journal of Financial Economics* 106, 635–659.
- Sprenger, T. O., Sandner, P. G., Tumasjan, A., Welpe, I. M., 2014a. News or noise? Using Twitter to identify and understand company-specific news flow. *Journal of Business Finance & Accounting* 41, 791–830.
- Sprenger, T. O., Tumasjan, A., Sandner, P. G., Welpe, I. M., 2014b. Tweets and trades: The information content of stock microblogs. *European Financial Management* 20, 926–957.
- Tetlock, P. C., 2007. Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance* 62, 1139–1168.
- Tumarkin, R., Whitelaw, R. F., 2001. News or noise? Internet postings and stock prices. *Financial Analysts Journal* 57, 41–51.
- White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the Econometric Society* pp. 817–838.

Fig. 1. Wholehealth Products, Inc (*\$GWPC*) - Stock price and Twitter activity



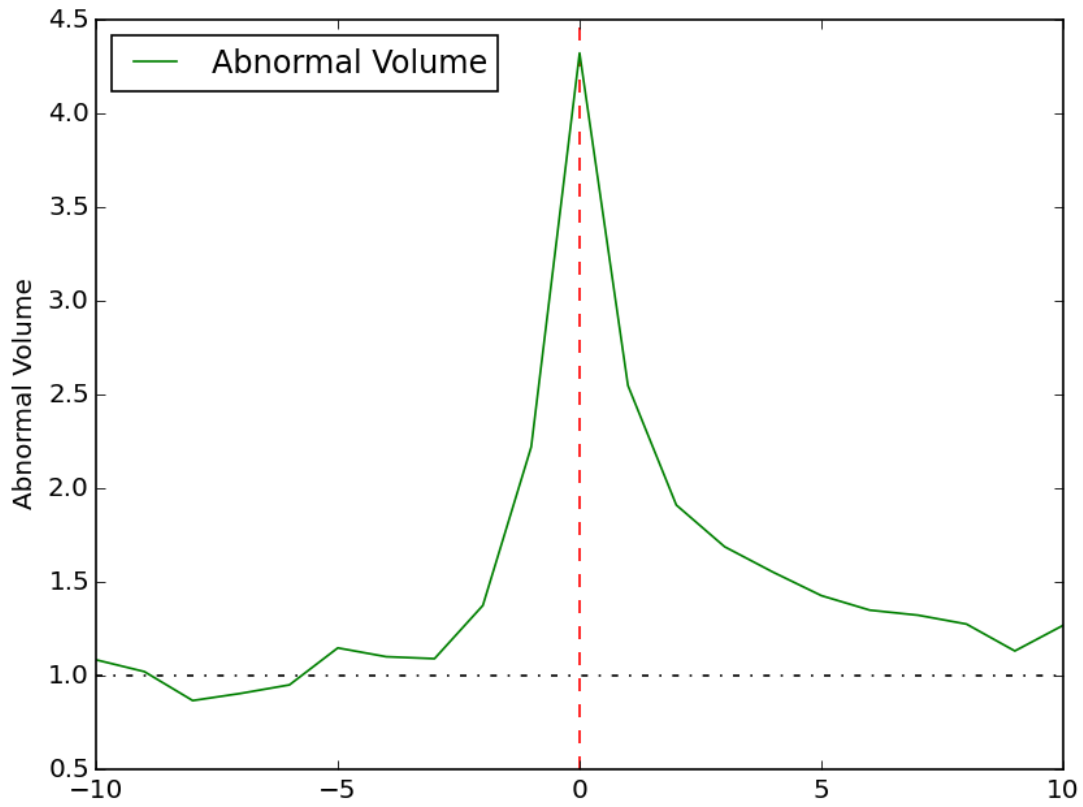
Notes: This figure shows the price of the Wholehealth Products (*\$GWPC*) shares (right-axis) and the daily number of messages containing the cashtag *\$GWPC* posted on Twitter between October 23, 2014, and November 24, 2014 (left-axis). Due to the SEC investigation, *\$GWPC* stock price is flat at \$0.048 between November 20, and November 24. *\$GWPC* stock price drops to \$0.0001 when trading resumes on December 23, 2014.

Fig. 2. Event Study - Abnormal returns and cumulative abnormal returns



Notes: This figure shows the abnormal returns and the cumulative abnormal returns on a [-10:+10] days event window. Events are defined as an abnormally high tweeting activity on social media. Average abnormal returns and cumulative average abnormal returns are computed on a total of 635 events about 315 distinct companies.

Fig. 3. Event Study - Abnormal volume



Notes: This figure shows the abnormal volume on a [-10:+10] days event window. Events are defined as an abnormally high tweeting activity on social media. Abnormal volume is computed on a total of 635 events about 315 distinct companies.

Table 1: Number of SEC civil actions by category and by fiscal year

	Broker-Dealer	Insider Trading	Securities Offering	Market Manipulation	Other Civil Actions	Total
1996	23	29	76	4	81	213
1997	19	36	66	11	72	204
1998	20	38	82	18	71	229
1999	15	51	67	26	78	237
2000	20	36	70	34	99	259
2001	13	47	56	17	104	237
2002	16	52	80	26	143	317
2003	32	37	70	18	156	313
2004	17	32	59	17	139	264
2005	20	42	34	30	138	264
2006	6	37	45	22	107	217
2007	59	31	44	27	101	262
2008	67	37	67	39	75	285
2009	26	42	106	34	104	312
2010	7	34	73	24	117	255
2011	21	48	82	29	86	266
2012	16	52	73	34	95	270
2013	7	43	76	23	58	207
2014	7	40	52	11	35	145
2015	4	26	58	28	46	162
Total	415	790	1,336	472	1,905	4,918
Total (%)	8.44%	16.06%	27.16%	9.60%	38.74%	100%

Notes: This table reports the number of SEC civil actions by category and by fiscal year. The category “Other Civil Actions” includes: “Investment Advisors/Companies”, “Delinquent Filings”, “Civil Contempt”, “Transfer Agents” and “Miscellaneous”. The category “Market Manipulation” includes “Newsletter/Touting”, a category initiated by the SEC in 1999 and re-integrated into “Market Manipulation” in 2003.

Table 2: Distribution of pump-and-dump manipulation cases

	Stock targeted		People involved					Tools used							
	OTC	Other mar-kets	In-sider	Promoter	Trader Shareholder	Other People	Press release	E-mail	Web-site	Fax	Tele-phone	Mailer	Mes-sage board	Social media	Other tools
2002	12	4	8	9	5	5	11	8	10	3	2	0	1	0	4
2003	8	1	6	4	3	2	6	4	4	2	0	0	2	0	4
2004	2	3	3	2	2	1	4	2	0	1	0	0	2	0	1
2005	11	7	12	8	4	2	12	5	5	3	3	0	1	0	5
2006	10	1	8	7	2	3	7	3	5	4	1	0	1	0	3
2007	7	2	6	3	4	1	8	3	2	2	0	0	2	0	3
2008	16	0	14	3	4	0	15	2	2	0	0	3	1	0	3
2009	12	0	6	7	5	2	8	3	3	3	0	2	0	0	4
2010	11	2	9	5	5	0	11	4	6	0	1	0	1	2	6
2011	7	0	3	5	4	2	5	4	3	1	1	0	0	0	0
2012	11	0	7	7	4	2	7	3	2	0	1	1	2	0	3
2013	3	0	2	3	0	1	3	3	0	0	1	0	1	1	1
2014	6	1	2	4	6	1	5	3	3	0	0	0	1	1	0
2015	13	0	5	7	8	2	8	4	3	0	0	3	1	0	2
Total	129	21	91	74	56	24	110	51	48	19	10	9	16	4	39
(%)	86.00	14.00	60.67	49.33	37.33	16.00	73.33	34.00	32.00	12.67	6.67	6.00	10.67	2.67	26.00

Notes: This table reports the distribution of pump-and-dump schemes from 2002 to 2015. Each case may involve multiple stocks, multiple people and multiple tools. "OTC" includes both the OTC Bulletin Board and the Pink Sheets. "Other markets" includes NASDAQ, NYSE, AMEX. "Insiders" are in great majority CEOs and CFOs. "Promoters" are paid investor relationship companies. "Other people" includes broker-dealers, attorneys and analysts. "Press releases" are disseminated through online wires like "PR Newswires" or "Business Wire". "E-mail" includes specialized newsletters and blast unsolicited e-mail. "Website" includes both companies' websites and promoters' websites. "Messages board" includes Yahoo! Finance, the Raging Bull, Investor Hub. "Social media" includes Twitter, Facebook and LinkedIn. "Other tools" includes mainly fake analyst reports and false filings sent to the SEC or to the FINRA.

Table 3: Top 10 most discussed OTC Markets stocks on Twitter

Ticker	Company	Market	Disclosure	Tweet Number	Stock Price	Market Cap
\$HALB	Tykhe Corp	OTC Pink	Current	397,098	0.01	#NA
\$CEGX	Cardinal Energy Group	OTC Pink	Current	169,263	0.8	28.14
\$STCC	Sterling Consolidated	OTC Pink	Limited	143,572	0.045	1.81
\$ARYC	Arrayit Corp	OTCQB	#NA	104,683	0.1624	6.43
\$GPDB	Green Polkadot Box	OTC Pink	Current	93,352	1.85	19.75
\$MINE	Minerco Resources	OTC Pink	Current	80,330	0.7813	19.04
\$MYEC	MyEcheck	OTC Pink	Current	49,940	0.0202	81.00
\$GWPC	Wholehealth Products	OTC Pink	Limited	36,500	0.25	19.92
\$PUGE	Puget Technologies	OTCQB	#NA	32,797	0.0556	2.36
\$CELH	Celsius Holdings	OTC Pink	Current	31,041	0.5283	9.78

Notes: This table presents the number of messages published on Twitter between October 5, 2014, and September 1, 2015, for the 10 most discussed stocks in our sample. Stock price (in USD) and market capitalization (in million USD) as of October, 1, 2014. #NA is used to indicate when information is not provided by Bloomberg.

Table 4: Messages containing the cashtag \$UBIQ posted on Twitter on February 6, 2015.

Date	User	Message content
2015-02-05 16:40:35	pickoftheday1	\$UBIQ Nice days here today Ubiquity Inc. QB UBIQ 0.729 + 0.159 27.89% Volume: 157 453 @ 4:05:55 PM ET [url]
2015-02-05 16:42:44	cabroncita	UBIQ Nice green close today!! Put this one on your radar! \$UBIQ [url]
2015-02-05 16:46:14	cabroncita	\$UBIQ Company Website [url]
2015-02-05 16:51:04	JetPenny	Cabroncita: \$UBIQ Company Website [url]
2015-02-05 16:59:37	iHangout4	cherrob: The Hunt for the Next 10 Bagger: \$UBIQ Nice days here today Ubiquity Inc. QB [url]
2015-02-05 16:59:38	iHangout4	cherrob: BREAKOUTS..RUNNERS AND HOT PENNIES: \$UBIQ Nice days here today Ubiquity Inc. QB [url]
2015-02-05 17:01:46	Hot_pennyalert	budfohub: \$UBIQ +27% big close today bull run starting! [url]
2015-02-05 17:20:54	cabroncita	\$UBIQ Check out their patents! http://t.co/ckhN8B36yB
2015-02-05 17:47:36	ThePUMPTracker	Alert : http://t.co/zVxyqFLK7 announces \$UBIQ as 02/06/2015 promo.Get the detailed promo [url]
2015-02-05 17:50:46	PennyStocksBlog	RT @ThePUMPTracker: Alert : [url] announces \$UBIQ as 02/06/2015 promo.Get the detailed promo alert at [url]
2015-02-05 19:31:49	Pubcos	\$UBIQ seems to be building. Looks like there could be some momentum. Let s see...
2015-02-05 19:35:16	Pubcos	\$UBIQ new entertainment amp communications experience creators amp viewers together combining television [url].
2015-02-05 19:38:08	Pubcos	Information is power. You may want to put this company on your radar: [url] \$UBIQ... Pubcos UBIQ MoneyTeam
2015-02-05 19:42:48	Pubcos	\$UBIQ +27% growth today. Congrats....
2015-02-06 09:01:32	PUMPSandDUMPS	Before you get taken by today s \$UBIQ pump you better review the ticker s Past Performances [url]
2015-02-06 09:25:39	StockPromoters	\$UBIQ StockPromoters.com http://t.co/v3yEIMowcM
2015-02-06 09:27:31	hecklerhouse	\$UBIQ chatter chatter
2015-02-06 10:43:25	WallStreetPenni	Mick Dodge: \$UBIQ is working hard to create a diverse [url]
2015-02-06 10:43:26	WallStreetPenni	Affix Trader: \$UBIQ Ubiquity is focused in five specific areas.... [url]
2015-02-06 10:51:09	JetPenny	Affix Trader: \$UBIQ Don t miss their patent portfolio overview! IP [url]
2015-02-06 15:15:29	explodeprofits	DITRstocks: \$UBIQ Short Term Indicators Barchart Opinion [url] [url]
2015-02-06 15:17:05	JetPenny	DITRstocks: \$UBIQ Short Term Indicators Barchart Opinion [url]

Notes: This table presents a sample of messages containing the cashtag \$UBIQ posted on Twitter between February 5, 2015 (after market close) and February 6, 2015 (before market close). The messages include tweets sent by stock promoters (i.e., @JetPenny, from the website <http://www.jetlifepennystocks.com/>), and @WallStreetPenni, from the website <http://www.wallstreetpennies.com/>) and tweets sent by users tracking pump-and-dump schemes (@PUMPSandDumps and @ThePumpTracker). This example is typical of a pump-and-dump scheme, where fraudsters or stock promoters try to temporarily inflate the price of a small capitalization stocks by using Twitter as a new channel of communication.

Table 5: Event-Study Results - Abnormal Return and Abnormal Volume

Day	Abnormal Return (%)	Abnormal Volume
-10	1.2288	1.0843
-9	0.1003	1.0213
-8	-0.7860	0.8668
-7	0.2093	0.9053
-6	0.9044	0.9505
-5	1.6631	1.1477
-4	-0.2072	1.1008
-3	0.0497	1.0903
-2	2.0686	1.3745
-1	4.1035***	2.2205
0	6.8796***	4.3223
1	-1.0551*	2.5479
2	-0.3310*	1.9102
3	0.5308	1.6875
4	-0.9802*	1.5522
5	-1.2729	1.4268
6	0.9869	1.3493
7	0.6920	1.3223
8	-0.2144	1.2750
9	0.2277	1.1315
10	-0.7696	1.2683
Event Number	635	635

Notes: This table shows the abnormal returns and the abnormal volumes on a [-10:+10] days event window around the event day (t_0). ***, ** and * represent abnormal returns significance, respectively at the 1%, 5%, and 10% level using a Corrado rank test.

Table 6: Event-Study Results - Cumulative Abnormal Returns (5-day)

Dates	5-day CAR (%)
[-10 : -6]	1.6569
[-9 : -5]	2.0912
[-8 : -4]	1.7837
[-7 : -3]	2.6193
[-6 : -2]	4.4786
[-5 : -1]	7.6776***
[-4 : 0]	12.8942***
[-3 : 1]	12.0462***
[-2 : 2]	11.6656***
[-1 : 3]	10.1278***
[0 : 4]	5.0441
[1 : 5]	-3.1085***
[2 : 6]	-1.0664**
[3 : 7]	-0.0434*
[4 : 8]	-0.7886***
[5 : 9]	0.4193**
[6 : 10]	0.9226**
Event Number	635

Notes: This table shows the cumulative abnormal returns at 5-day intervals in a [-10:+10] days event window around the event day (t_0). ***, ** and * represent abnormal returns significance respectively at the 1%, 5%, and 10% level using a Corrado rank test.

Table 7: Descriptive Statistics

Variable	Mean	Median	Min	Max	Std-Dev
AR_t	0.0688	0.0142	-0.7731	1.9837	0.2386
$CAR_{t+1;t+5}$	-0.0311	-0.0306	-0.9673	1.4782	0.2365
$MktCap_t$	196.2422	9.3357	1.0011	15037.7915	1278.4197
$Price_t$	2.1342	0.1225	0.0001	168.75	8.3172
$NonTradingDays_t$	0.0508	0.0	0.0	0.8571	0.1269
$MktType_t$	0.126	0.0	0.0	1.0	0.3321
$Sentiment_t$	0.3063	0.2727	-0.9436	1.0	0.317
$News_t$	0.1843	0.0	0.0	1.0	0.388
$MessageNumber_t$	192.5465	48.0	20.0	5250.0	514.5267
$StockPromoter_t$	0.6346	1.0	0.0	1.0	0.4819
$PumpTracker_t$	0.0898	0.0	0.0	1.0	0.2861
$IRAccount_t$	0.4961	0.0	0.0	1.0	0.5004

Notes: This table presents summary statistics for all variables used in the contemporaneous and predictive regressions. The sample includes a total of 635 observations (events), where t represents the event day (i.e. days with an abnormally high tweeting activity about company i).

Table 8: Contemporaneous Regression

Model	[1]		[2]		[3]	
Variable	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
α	0.1296***	5.9237	0.0413	1.1532	0.0034	0.0890
AR_{t-1}	-0.2188***	-3.9294	-0.2268***	-4.0162	-0.2425***	-4.3619
$MktCap$	-0.0159***	-3.0639	-0.0136***	-2.6893	-0.0124**	-2.4930
$Price$	-0.0011*	-1.8117	-0.0011**	-2.0287	-0.0010	-1.4964
$NonTradingDays$	-0.1404**	-2.3708	-0.1149*	-1.9156	-0.0848	-1.4392
$MarketType$	-0.0213	-0.7907	-0.0097	-0.3643	-0.0128	-0.4909
$Sentiment$			0.0911***	3.3576	0.0865***	3.2003
$News$			0.0305	1.2926	0.0239	1.0336
$MessageNumber$			0.0110	1.4499	0.0109	1.4398
$StockPromoter$					0.0612***	3.4007
$PumpTracker$					-0.0948***	-3.2784
$IRAccount$					0.0144	0.7864
Adj- R^2 (%)	4.62		6.23		8.23	
Observations	635		635		635	

This table reports the results of the equation $AR_t^i = \alpha + \beta_1 AR_{t-1}^i + \beta_2 MarketCap_t^i + \beta_3 Price_t^i + \beta_4 NonTradingDays_t^i + \beta_5 MarketType_t^i + \beta_6 Sentiment_t^i + \beta_7 News_t^i + \beta_8 NumberMessages_t^i + \beta_9 StockPromoter_t^i + \beta_{10} PumpTracker_t^i + \beta_{11} IRIccount_t^i + \epsilon_t$. Standard errors are computed using White (1980) heteroskedasticity robust standard errors. Superscripts ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively. The regression includes 635 observations.

Table 9: Predictive Regressions

Model	[n=5]		[n=5]		[n=5]	
	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
α	-0.0239	-1.1423	-0.0009	-0.0212	0.0205	0.5100
AR_t	-0.1805***	-3.6333	-0.1814***	-3.5638	-0.1848***	-3.7125
$MktCap$	0.0026	0.5494	0.0026	0.5360	-0.0018	-0.3665
$Price$	0.0001	0.2700	0.0001	0.1700	-0.0001	-0.2062
$NonTradingDays$	0.0133	0.1995	0.0132	0.1966	0.0132	0.2017
$MarketType$	-0.0172	-0.5488	-0.0176	-0.5635	-0.0253	-0.8074
$Sentiment$			0.0161	0.5002	0.0097	0.3030
$News$			-0.0013	-0.0626	-0.0030	-0.1407
$MessageNumber$			-0.0065	-0.8569	-0.0035	-0.4615
$StockPromoter$					-0.0430**	-2.2386
$PumpTracker$					-0.0845***	-2.8325
$IRAccount$					0.0315*	1.6703
Adj- R^2 (%)	2.72		2.40		4.39	
Observations	635		635		635	

Model	[n=1]		[n=2]		[n=10]	
	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
α	-0.0045	-0.1982	0.0458	1.3415	0.1382**	2.2166
AR_t	-0.0665**	-2.2536	-0.1122***	-2.8144	-0.0909	-1.5180
$MktCap$	-0.0006	-0.2504	-0.0040	-1.0680	-0.0087	-1.1945
$Price$	0.0008**	2.2017	-0.0001	-0.2789	0.0012	0.4891
$NonTradingDays$	-0.0386	-1.0564	0.0058	0.1073	0.3447	1.6212
$MarketType$	-0.0274*	-1.7096	-0.0603**	-2.4085	0.0338	0.5947
$Sentiment$	-0.0008	-0.0448	0.0044	0.1933	-0.0096	-0.2396
$News$	-0.0017	-0.1360	-0.0132	-0.9674	0.0134	0.5047
$MessageNumber$	0.0059	1.2633	-0.0011	-0.1933	-0.0254**	-2.3697
$StockPromoter$	-0.0365***	-3.4556	-0.0414***	-2.6506	-0.1043***	-3.2928
$PumpTracker$	-0.0240*	-1.7447	-0.0729***	-3.9165	-0.0809**	-2.0453
$IRAccount$	0.0097	0.9502	0.0081	0.5626	0.0519*	1.6506
Adj- R^2 (%)	3.51		4.41		5.04	
Observations	635		635		635	

This table reports the results of the equation $CAR_{t+1,t+n}^i = \alpha + \beta_1 AR_t^i + \beta_2 MarketCap_t^i + \beta_3 Price_t^i + \beta_4 NonTradingDays_t^i + \beta_5 MarketType_t^i + \beta_6 Sentiment_t^i + \beta_7 News_t^i + \beta_8 NumberMessages_t^i + \beta_9 StockPromoter_t^i + \beta_{10} PumpTracker_t^i + \beta_{11} IRAccount_t^i + \epsilon_t$ for $n=1$, $n=2$, $n=5$ and $n=10$. Standard errors are computed using White (1980) heteroskedasticity robust standard errors. Superscripts ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively. The regression includes 635 observations.

Table 10: Event-Study Results - Cumulative Abnormal Returns (5-day)

Dates	Promoter Event (%)	No Promoter Events (%)
[-10 : -6]	2.3640	0.3947
[-9 : -5]	2.8565	0.7251
[-8 : -4]	2.3922	0.6974
[-7 : -3]	3.2668	1.4635
[-6 : -2]	6.1218	1.5454
[-5 : -1]	10.4945*	2.6493**
[-4 : 0]	17.2805***	5.0642***
[-3 : 1]	14.9200***	6.9163***
[-2 : 2]	14.0037***	7.4918***
[-1 : 3]	11.4892***	7.6975***
[0 : 4]	5.3457	4.5058
[1 : 5]	-5.3418**	0.8782*
[2 : 6]	-3.0545	2.4826**
[3 : 7]	-2.4034	4.1694
[4 : 8]	-3.0681***	3.2805
[5 : 9]	-1.8124**	4.4032
[6 : 10]	-1.3852**	5.0422
Number of Events	407	228

Notes: This table shows the cumulative abnormal returns at 5-day intervals in a [-10:+10] days event window around the event day (t_0). “Promoter Event” are all events with at least one tweet by a stock promoter or one tweet from a user tracking pump-and-dump schemes (407 events). “No Promoter Event” are all events without any tweets by a stock promoter nor by a user tracking pump-and-dump schemes (228 events). Superscripts ***, ** and * represent abnormal returns significance, respectively at the 1%, 5%, and 10% level using a Corrado rank test.

Table 11: Predictive Regressions - Control Variables

Model	[1]		[2]		[3]	
	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
α	0.0051	0.1018	0.0161	0.4038	-0.0025	-0.0507
AR_t	-0.1806***	-3.6134	-0.1836***	-3.6842	-0.1787***	-3.5737
$MktCap$	-0.0016	-0.3116	-0.0020	-0.3999	-0.0017	-0.3311
$Price$	-0.0001	-0.1910	-0.0002	-0.4267	-0.0002	-0.4066
$NonTradingDays$	0.0112	0.1713	-0.0061	-0.0899	-0.0084	-0.1257
$MarketType$	-0.0235	-0.7477	-0.0233	-0.7464	-0.0216	-0.6892
$Sentiment$	0.0096	0.3004	0.0070	0.2191	0.0070	0.2194
$News$	-0.0012	-0.0547	0.0055	0.2533	0.0073	0.3340
$NumberMessages$	-0.0053	-0.6731	-0.0011	-0.1397	-0.0025	-0.3091
$StockPromoter$	-0.0426**	-2.2075	-0.0383**	-1.9695	-0.0382*	-1.9537
$PumpTracker$	-0.0806***	-2.6390	-0.0803***	-2.6645	-0.0767**	-2.4907
$IRAccount$	0.0315*	1.6708	0.0301	1.6019	0.0301	1.6031
$LinkPct$	0.0198	0.6300			0.0223	0.7117
$RetweetPct$	0.0355	0.7957			0.0334	0.7479
$Marijuana$			-0.0414*	-1.8793	-0.0414*	-1.8759
Adj- R^2 (%)	4.23		4.65		4.48	
Observations	635		635		635	

This table reports the results of the equation $CAR_{t+1,t+n}^i = \alpha + \beta_1 AR_t^i + \beta_2 MarketCap_t^i + \beta_3 Price_t^i + \beta_4 NonTradingDays_t^i + \beta_5 MarketType_t^i + \beta_6 Sentiment_t^i + \beta_7 News_t^i + \beta_8 NumberMessages_t^i + \beta_9 StockPromoter_t^i + \beta_{10} PumpTracker_t^i + \beta_{11} IRAccount_t^i + \beta_{12} LinkPct_t^i + \beta_{13} RetweetPct_t^i + \beta_{14} Cannabis_t^i + \epsilon_t$ for $n=5$. Standard errors are computed using White (1980) heteroskedasticity robust standard errors. Super-scripts ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively. The regression includes 635 observations.