Pump-and-dump or news? Stock market manipulation on social media

Thomas Renault[a,b]

[a]IÉSEG School of Management
[b]Université Paris 1 Panthéon Sorbonne

Abstract

Social media can help investors gathering and sharing information about stock markets but also presents opportunities for fraudsters to send false or misleading statements to the marketplace. Analyzing millions of messages sent on the social media Twitter about securities traded in OTC markets, we find that an abnormally high message activity on social media is associated with a large price increase on the event day, followed by a sharp price reversal over the next week. Our results are consistent with a pump-and-dump scheme, where fraudsters use social media to temporarily inflate the price of small capitalization stocks. To disentangle an overoptimism effect due to the presence of noise traders from an illegal pump-and-dump scheme, we investigate social interactions between users through the use of network theory. We identify several clusters of users with suspicious online activity (stock promoters, fake accounts, automatic postings), favoring the manipulation/promotion hypothesis over the behavioral hypothesis.

Keywords: Asset Pricing, Market Manipulation, Social Media

JEL classification: C18, D80, G12, G14.

1. Introduction

Market manipulation is at least as old as trading on organized exchanges (Putniņš, 2012). More than three centuries have passed since the first stories of market manipulation have been reported by Joseph de la Vega in 1688, but despite substantial academic researches, our understanding of market manipulation is still very incomplete. While theoretical models have been developed to address trade-based manipulation (Allen and Gale, 1992) or information-based manipulation (Bommel, 2003), empirical studies remain 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 consist of touting a company's stock through false or misleading statements to 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 has existed for many decades, the emergence of the Internet and social media has created a new fertile ground for fraudsters. Spreading false or misleading information to a large number of potential investors can now be done with minimum effort, anonymously and at a relatively low cost.[1] According to the Security and Exchange Commission[2], "investors who learn of investing opportunities from social media should always be on the lookout for fraud."

To have a better understanding of pump-and-dump schemes, we first start by focusing 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 called "micro-cap" or "penny stocks", listed on the OTC Markets (previously known as Pink Sheets). Regarding tools used by fraudsters, pump-and-dump schemes often combine a false or misleading press release with a touting of the stock on spam e-mails, websites or bulletin boards. We find that fraudsters specifically use Twitter to manipulate stock prices in two cases. In the first case, 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 (see Appendix A). In the second case, a Scottish trader sent tweets falsely announcing that two companies were under investigation causing sharp drops in the stock prices of the targeted companies (see Appendix B).

While empirical proofs of market manipulation on small capitalization stocks have al-

[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

ready been identified using data from stock spams (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 tested. We extend the literature on indirect empirical evidence of market manipulation by analyzing data from one of the largest worldwide social media: Twitter. Analyzing data from Twitter could provide new insights compared to previous studies as interactions between users are directly observable. This feature allows researchers to cluster users based on common characteristics in order to detect groups of users (insiders, stock promoters, traders...) with suspicious behaviors. Furthermore, as data are collected in real-time, our analysis is not affected by the survivorship bias appearing when data are collected "ex-post". As in any illegal activity, we should expect market manipulators to delete their messages or accounts after conducting a fraud in order to decrease the probability of being caught by the Security and Exchange Commission. Collecting data in real-time on Twitter solve this issue.

We conduct event-studies to examine the relation between a spike in posting activity on Twitter and stock returns. We find that an abnormally high message activity on social media about a company listed on OTC Markets is associated with a large price increase on the event day, followed by a sharp price reversal on the next week. This price reversal pattern is consistent with a pump-and-dump scheme (manipulation hypothesis) but could also simply be caused by overoptimistic noise traders (behavioral hypothesis). While judicial inquiries would be needed to assess precisely if large increase/decrease in stock prices are caused by fraudsters or by irrational unsophisticated traders, we investigate social interactions between users through the use of network theory to identify suspicious online behaviors. Clustering users based on Twitter mentions and retweets, we identify few groups of users with behaviors that could be related to frauds (multi-account posting, automatic posting, scheduled posting activity), favoring the manipulation hypothesis over the behavioral hypothesis. Overall, our finding shed light on the need for a higher control of the information published on social media and a higher education for investors looking for trading opportunities on the Internet.

Our paper is structured as follows. Section 2 presents briefly the theoretical literature on market manipulation and reviews the empirical literature using data from the Internet. Section 3 describes the database we construct by analyzing SEC litigation releases and justifies our focus on OTC Markets. Section 4 presents the OTC markets and data extracted from Twitter. Section 5 shows the event-studies results. Section 6 proposes a methodology to identify potential fraudsters by analyzing interactions between users and discusses how to avoid frauds on social media. Section 7 concludes.

2. Related Literature

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). But, despite the importance of ensuring fair and transparent markets, little is know about the prevalence and impact of market manipulation (Putniņš, 2012). Theoretical papers have shown that informed or uninformed traders can generate profits through trade-based manipulation (Allen and Gale, 1992) or information-based manipulation (Bommel, 2003). However, as any illegal behavior, market manipulation is not directly observable and empirical studies remain very scarce. Due to the lack of data available, a 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 stocks manipulated are penny stocks with low trading volume and market capitalization that trade in OTC Markets, such as the OTC Bulletin Board and the Pink Sheets.[3] Regarding techniques used by fraudsters, more than 55% of cases involves the spread of rumors or false information. Manipulators also frequently use wash trades and nominee accounts to create artificial trading activity. However, only a small fraction of manipulation is detected and prosecuted (Comerton-Forde and Putniņš, 2014). Furthermore, focusing on reported cases tend to create a selection bias towards poor manipulation and is affected by regulators agenda (Bonner et al., 1998). Thus, another strand of the literature focus on indirect evidences by studying abnormal market behaviors (for trade-based manipulation) or by trying to detect suspicious behaviors outside the market (for information-based manipulation).

Analyzing intraday volume and order imbalance, Ben-Davis et al. (2013) provide evidences suggesting that some hedge funds manipulate stock prices on critical reporting dates. Their findings are consistent with Carhart et al. (2002) on end-of-quarter manipulation by mutual funds. Closer to our paper, a nascent strand of the literature focus 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), Hanke and Hauser (2008) and Nelson et al. (2013) all find a significant positive short-run price impact after a stock spam touting, followed by a price reversal on the following days. Similar patterns have also been identified by Sabherwal et al. (2011) who use Internet message boards activity to identify pump-and-dump scheme on small stocks without fundamental news .

[3] Only 17% of the reported cases occurs in the NYSE, the AMEX or the NASDAQ.

In this paper, we follow both approaches. We first start by analyzing reported manipulation cases before conducting an empirical investigation of 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 2001[4] and then manually classify all cases by looking at 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 better understand who are the people involved in frauds (insiders, promoters, traders...) and which tools are used to send false or misleading information to 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 available in 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 year. In the remainder of this paper, we focus our analysis on the category "Market Manipulation", including the sub-category "Newsletter/Touting". Each case is included in only one category following SEC classification, even though many cases involve multiple allegations and may fall under more than one category.

Overall, market manipulations represented 9.60% of all civil actions initiated by the SEC between 1996 and 2015. During our sample period, the SEC has demonstrated its commitment to prosecute market manipulation occurring in Cyberspace on numerous occasions. For example in October 1998 (fiscal year 1999), the SEC launched a nationwide "sweep" on purveyors of fraudulent spam, online newsletters, message board postings and websites caught in "effort to clean up the Internet", leading to 23 enforcement actions against 44 individuals and companies.[5] In 2000, the fourth nationwide internet fraud sweep leads to 15

[4]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".

[5]"SEC Conducts First Ever Nationwide Internet Securities Fraud Sweep, Charges 44 Stock Promoters in 23 Enforcement Actions"

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

	Broker Dealer	Insider Trading	Securities Offering	Market Manipulation	Others 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. Category "Others Actions" includes: Investment Advisors/Companies, Delinquent Fillings, Civil Contempt, Transfer Agents and Miscellaneous. Category "Market Manipulation" includes "Newsletter/Touting", a category initiated by the SEC in 1999 and re-integrated into Market Manipulation in 2003.

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 large 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 SEC agenda and may be biased towards poor manipulation, studying the different civil actions still help us understand which type of stocks are manipulated, who are the people involved and which tools and techniques are used by fraudsters. Starting in 2002, the SEC starts releasing detailed complaints for a great majority of civil actions they initiate. While litigation releases only summarize the enforcement case in approximately one page, complaint reports gives a lot of detail about the fraudulent scheme and the exact role of each defendants, in ten to thirty pages. Using this new report, we manually analyze all complaints classified as "Market Manipulation" or "Newsletter/Touting". Over the 362 "Market Manipulation" civil actions initiated by the SEC between 2002 and 2015, we manage to collect detailed complaint reports for 273 cases, of which 150 are related to pump-and-dump schemes. Table 2 summarizes respectively, year by year, the type of stocks targeted by fraudsters, the type of people involved in the manipulation scheme and the tools used to disseminate false or misleading information to the marketplace.

We find that 86% of pump-and-dump schemes target stocks listed on OTC Markets. The most common channel of communication used by fraudsters to send false or misleading information to the marketplace are press releases (73.3%), followed by spam emails / newsletters (34%), websites (32%), fax blast (12.6%) and message boards (10.6%).[6] People involved in frauds are mostly company's insiders (CEO, CFO...) (60.7%), stock promoters paid in cash or in shares to pump the price of a stock (49.3%)[7] and traders / shareholders (37.3%).

We identify two cases where manipulators specifically use Twitter to manipulate stock prices. In the first case, Twitter was used by a stock promoter to tout the value of a penny

[6]The sum is not equal to 100% as fraudsters often combine multiple channel of communication to increase the outreach and visibility of their messages.

[7]Stock promotion (investor relation) are not per se illegal. 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. According to the Securities Act of 1933, Section 17(b), "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".

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

	Stock Targeted		People Involved				Tools Used								
	OTC	$Other_m$	Insider	Pro-moter	Trader Share-holder	$Other_p$	Press Release	Email	Website	Fax	Tele-phone	Mailer	Message Board	Social Media	$Others_t$
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
Total (%)	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, regarding the type of stocks targeted by fraudsters, the type of people involved in the fraud and the tools uses by fraudsters to manipulate market. Each case may involve multiple stocks, multiple people and multiple tools. OTC Markets includes both the OTC Bulletin Board and the Pink Sheets. "$Other_m$" includes NASDAQ, NYSE, AMEX and "unknown". Insiders are in great majority CEO and CFO's. Promoters are paid investor relationship companies using various channel of communication to send false or misleading information to the marketplace. "$Other_p$" includes broker-dealer, attorneys and analysts. Press releases are disseminated through online wires like "PR Newswires" or "Business Wire". Spam e-mail includes specialized newsletter and blast unsolicited e-mails. Website includes both companies' websites and promoters' websites. Messages boards include Yahoo! Finance Message Boards, the Raging Bull and Investor Hub. Social media includes Twitter, Facebook and LinkedIn. "$Other_t$" includes fake analyst reports and false filings sent to the SEC or FINRA.

stock before selling shares at an inflated price.[8] In the second case, a Scottish trader dupe the public by mimicking the existing securities research firm Citron Research's Twitter page to send multiple false statements regarding two companies and to profit from the sharp price swings.[9] Given those recent cases and SEC renewed attention towards risks created by social media communication, we believe that analyzing data from Twitter could provide new insights into the empirical literature on stock market manipulation.

4. Data

4.1. The OTC Markets Group

Given our preliminary analysis on SEC litigation, we choose to focus our analysis on stocks listed in the OTC Markets. The OTC Markets Group is an electronic inter-dealer quotation and trading system providing marketplaces for around 10,000 over-the-counter (OTC) securities. 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.

1. The OTCQX marketplace: Companies must meet high financial standards, be current in their disclosure and receive third party advisory.
2. The 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.
3. The OTC Pink marketplace: Open for all companies. The OTC Pink is then 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. For companies listed on OTC Pink, 814 are "Current" regarding information provided, 403 only provide "Limited" information and 1,951 are classified as "No Information". Companies in the OTC Pink "No

[8]Litigation Release No. 21580 - https://www.sec.gov/litigation/litreleases/2010/lr21580.htm

[9]Litigation Release No. 23401 - https://www.sec.gov/litigation/litreleases/2015/lr23401.htm - The fake Twitter account created was "@Citreonresearc" (without an "h") whereas the official Twitter account of the securities research team is "@citronresearch"

Information" category should, according to OTC Markets "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 our sample period, the vast majority of stock experienced a sharp decrease in price with a numerous number of stocks loosing nearly all their value. This finding is consistent with Ang et al. (2013) who identify that, over a long period, comparable listed-stocks tend to overperform OTC stocks by nearly 9% per year. However, few stocks experienced impressive returns over the sample period. For example, Micro Imaging Technology stock price increases from \$0.0229 to \$0.45 between October 2014 and October 2015 (+1,800%). As documented by Eraker and Ready (2015), OTC Markets return are negative on average and highly positively skewed, with few "lottery-like" stocks doing extremely well while many of the stocks became 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 millions messages are posted on Twitter. We develop a computer program in the Python programming language to collect data in real-time using Twitter Search and Stream Application Programming Interface (API). Following Da et al. (2011) and Drake et al. (2012), we identify a stock using its ticker symbol. More precisely, given Twitter "cashtag" feature introduced in 2012, we extract all message containing a "\$" sign followed by the ticker name as in Sprenger et al. (2014). Figure 1 shows three tweets containing the ticker \$AACS from our database about company "American Commerce Solutions".[10]

During our sample period from October 5th 2014 to September 1st 2015, we collect a total of of 7,196,307 tweets containing at least one "cashtag" of a stock listed on the OTC Markets. Among the 5,087 companies, around 50% of companies received a very low level of attention (between 0 and 20 tweets). On the other side, four companies got 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 over our sample period. Overall, we find that Twitter activity is higher for companies listed in OTC Pink marketplace, with a low stock price (penny stocks) and a small market capitalization.

Analyzing all messages published on Twitter about the five most discussed companies in our sample, we identify a list of 255 fake Twitter accounts posting exactly the same type of

[10] \$AACS is the first company in our sample, in alphabetical order, with a significant number of tweets.

Fig. 1. Example of tweets containing the cashtag $AACS

Notes: Figure 1 shows three tweets containing the cashtag $AACS published on January 21-22th 2014. We extract from each tweet the exact time stamp, the name of the user who sent the tweet and the content of the message.

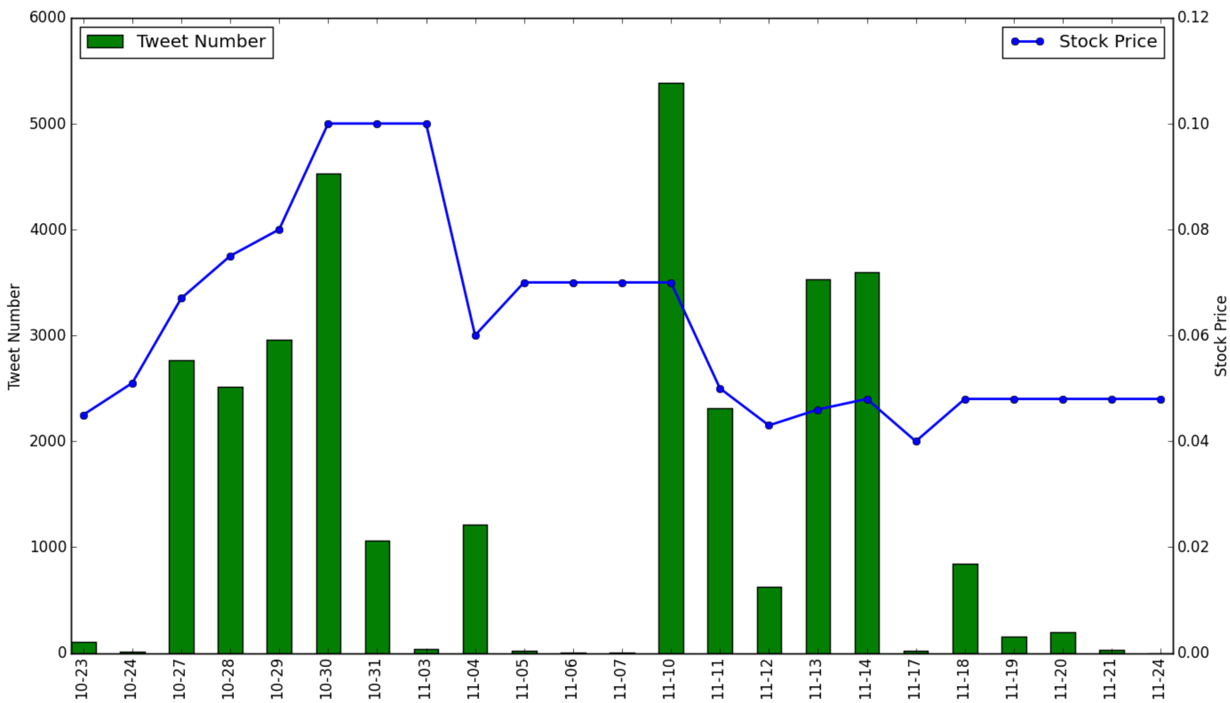
messages at different period, simply replacing a ticker by another and changing few keywords overtime. After a certain period of abnormally high posting activity, the number of tweets came back to a level close to zero. While we cannot assess if those bursts in social media activity are directly linked with attempts to manipulate market, the use of multiple fake accounts to post messages to urge buying a stock is at least suspicious. We will study more in details users' behavior on the last part of this paper.

The case of Wholehealth Products, Inc. ($GWPC), the eight most-discussed stock in our sample, is also especially interesting. On November 20th 2014, the Security Exchange Commission have suspended the trading on GWPC as questions have arisen concerning the accuracy and adequacy of publicly disseminated information, including information about the relationship between the company's business prospects and the current Ebola crisis.[11] Looking at the number of messages containing the ticker *$GWPC* posted on Twitter before SEC halt, we identify a sharp increase in posting activity starting on October 26th (see Figure 2). 2,768 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 afterwards.

This anecdotal example is typical of a pump-and-dump scheme. A false information is shared on Twitter, creating a large increase in social media activity about a given company. Stock price increases (pump) over a short period, and decreases sharply (dump) afterwards. In the next section, we will conduct an event-study to analyze if the price reversal pattern

[11] "SEC Suspends Trading in Companies Touting Operations Related to Prevention or Treatment of Ebola"

Fig. 2. Wholehealth Products, Inc ($GWPC) - Stock Price and Twitter Activity



Notes: Figure 2 shows $GWPC daily stock price (right-axis) and the daily number of messages containing the cashtag $GWPC published on Twitter between October 23th 2014 and November 24th 2014 (left-axis). Due to SEC investigation, stock price is constant at $0.048 between November 20th and November 24th. $GWPC stock price drops to $0.0001 when trading resumes on December 23rd 2014.

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		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		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 5th 2014 and September 1st 2015 for the 10 most discussed stocks. Stock price (in USD) and market capitalization (in million USD) as of October 1st, 2014. #NA represents values non-available on Bloomberg.

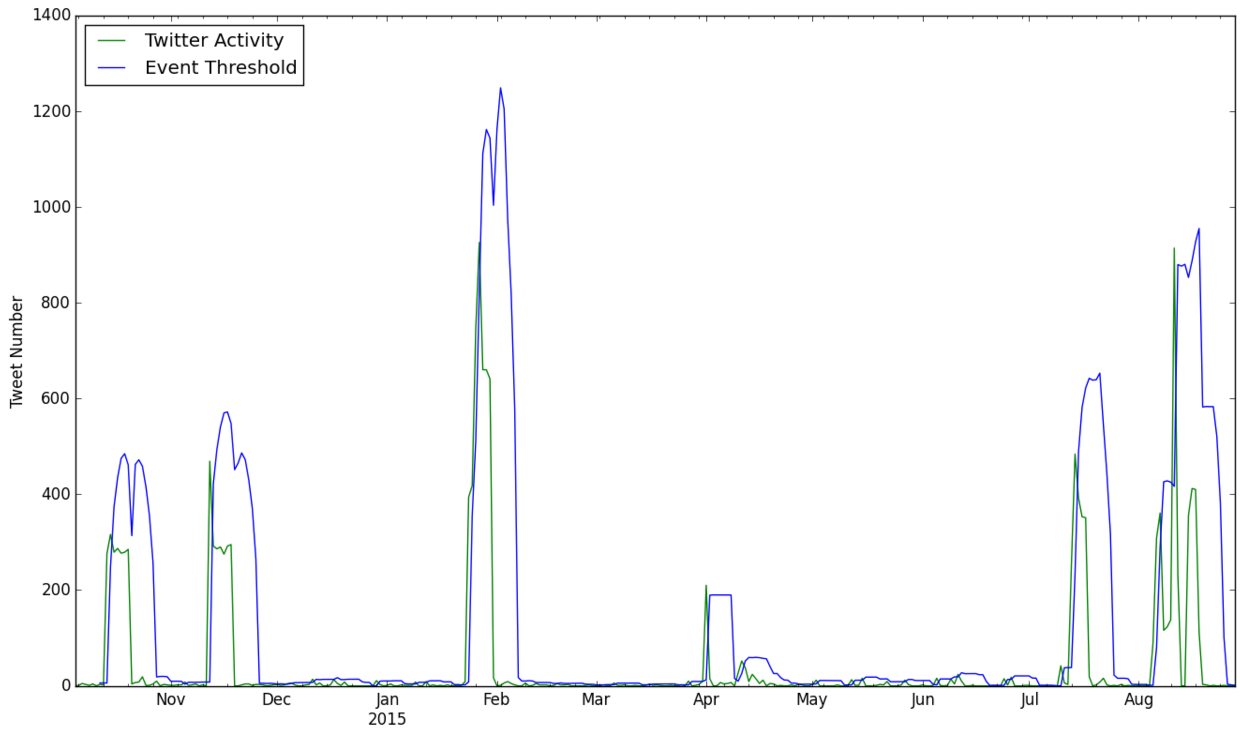
identified anecdotally in the $GWPC case can be generalized. We will do so by analyzing the link between an abnormally high activity on social media and stock prices of companies listed on OTC markets.

5. Event-Study

Following Tumarkin and Whitelaw (2001) and Leung and Ton (2015), we define an event when the number of messages posted on Twitter about company i during a given day t exceeds the 7 previous days average number of messages plus two-standard deviations. We consider all messages sent between 4pm on day *t-1* and 4pm on day t to only consider tweets sent before market close on day t. We impose a minimum of 20 tweets from 20 distinct users to avoid having our results driven by few active users. If an event is detected on a non-trading days, we consider the next trading day as being the event day. To include an event in our event-study, we impose a minimum stock price of $0.1 and a market capitalization greater than $1,000,000 at the beginning of the event-window. As in Ang et al. (2013) and Eraker and Ready (2015), we test various threshold for minimum price, minimum market capitalization and minimum percentage of non-trading days to avoid having our results driven by illiquid or non-tradable stocks. Our results are robust to a $0.01 and $1 minimum price, a $100,000 and $10,000,000 minimum market capitalization, and minimum non-trading-days from 25% to 75%.

The following example illustrates our methodology on a specific company: SinglePoint

Fig. 3. SinglePoint, Inc ($SING) - Twitter Activity and Event Detection



Notes: Figure 3 shows the daily number of messages sent on Twitter containing the ticker $SING and one-week rolling-mean plus two standard deviations threshold. We define event-days all dates at which the number of messages (in blue) exceeds the one-week rolling-mean plus two standard deviations (in green).

Inc, (*$SING*). During our sample period, a total of 15,188 messages containing the ticker *$SING* where posted on Twitter. Figure 3 shows the daily number of messages on Twitter, and the threshold level we use for event detection. Using this methodology, we identify six events for *$SING* company, on October 14th 2014, November 12th 2014, January 24th 2015, April 1st 2015, July 13th 2015 and August 7th 2015. Table 4 shows a selected sample of tweets related to October 14th 2014 event. Activity on Twitter on that day is typical of a stock promotion scheme, where tweets are sent by bots using regular timing schedule and multi-accounts. All the account promoting the stock are owned by ”Stock Talk 101”, a firm ”engaged in the business of marketing and advertising companies for monetary compensation”. As a disclaimer is clearly visible on stock promoter Twitter accounts, the scheme is not per se illegal. However, this example is an illustrative case of how Twitter can be used by stock promoters as a new channel of communication.

For each event, we analyze all tweets sent on that day using Renault (2016) investor social lexicon. We convert each tweet into a quantitative sentiment variable and we aggregate individual message sentiment to derive an event sentiment. We find that 82.41% of event-day are associated with a positive sentiment. As already documented on the literature (see

Table 4: Messages about $SING sent on October 13th 2014 on Twitter

Date	User	Message content
2014-10-13 17:12:05	ckelly3	RT @majegivudys: $SING - SinglePoint's product suite will provide Medical Cannabis dispensaries a user-friendly platform [...]
2014-10-13 17:36:19	_Singlepoint_	$SING working to finish acquiring GreenStar Payment Solutions in short order
2014-10-13 17:44:57	badnewsbruno	RT @_Singlepoint_: $SING working to finish acquiring GreenStar Payment Solutions in short order
2014-10-13 18:54:40	_Singlepoint_	$SING increasing number of terminals every week on track to hit sales targets
2014-10-13 20:00:01	JayBugster	When completed GreenStar Payment Solutions, Inc. will be a wholly owned subsidiary of SinglePoint. - $SING
2014-10-13 20:00:02	BoardwalkPennyS	When completed GreenStar Payment Solutions, Inc. will be a wholly owned subsidiary of SinglePoint. - $SING
2014-10-13 20:00:02	Micro_Cap_Pro	When completed GreenStar Payment Solutions, Inc. will be a wholly owned subsidiary of SinglePoint. - $SING
2014-10-13 20:00:02	MicroCapUnivers	When completed GreenStar Payment Solutions, Inc. will be a wholly owned subsidiary of SinglePoint. - $SING
2014-10-13 20:00:03	PennyStockMach	When completed GreenStar Payment Solutions, Inc. will be a wholly owned subsidiary of SinglePoint. - $SING
2014-10-13 20:00:03	StockShocks	When completed GreenStar Payment Solutions, Inc. will be a wholly owned subsidiary of SinglePoint. - $SING
2014-10-13 20:00:03	PennyStockExcel	When completed GreenStar Payment Solutions, Inc. will be a wholly owned subsidiary of SinglePoint. - $SING
2014-10-13 20:00:03	DaddyHotStocks	When completed GreenStar Payment Solutions, Inc. will be a wholly owned subsidiary of SinglePoint. - $SING
2014-10-13 20:00:03	StockUltraman	When completed GreenStar Payment Solutions, Inc. will be a wholly owned subsidiary of SinglePoint. - $SING
2014-10-13 20:00:03	HotStockCafe	When completed GreenStar Payment Solutions, Inc. will be a wholly owned subsidiary of SinglePoint. - $SING
2014-10-13 20:00:03	Penny_Hotsocks	When completed GreenStar Payment Solutions, Inc. will be a wholly owned subsidiary of SinglePoint. - $SING
2014-10-13 20:00:07	PlatinumPennys	When completed GreenStar Payment Solutions, Inc. will be a wholly owned subsidiary of SinglePoint. - $SING
2014-10-13 20:00:11	Virmmac	When completed GreenStar Payment Solutions, Inc. will be a wholly owned subsidiary of SinglePoint. - $SING
2014-10-13 20:00:13	IonPennyStocks	When completed GreenStar Payment Solutions, Inc. will be a wholly owned subsidiary of SinglePoint. - $SING
2014-10-13 21:00:03	JayBugster	Did you read the $SING LOI news? http:t.cogTEykD6oSz
2014-10-13 21:00:03	Virmmac	Did you read the $SING LOI news? http:t.coAxEUNqLoOk
2014-10-13 21:00:04	BoardwalkPennyS	Did you read the $SING LOI news? http:t.coarilkU2OIX
2014-10-13 21:00:04	StockShocks	Did you read the $SING LOI news? http:t.coixKf03Rm9C
2014-10-13 21:00:04	PlatinumPennys	Did you read the $SING LOI news? http:t.coSUg48Rf392
2014-10-13 21:00:04	StockUltraman	Did you read the $SING LOI news? http:t.coWDy1STjZ2B
2014-10-13 21:00:04	Micro_Cap_Pro	Did you read the $SING LOI news? http:t.co8LlCPwiBLV
2014-10-13 21:00:06	HotStockCafe	Did you read the $SING LOI news? http:t.cof3EKncSwQq
2014-10-13 21:00:07	Penny_Hotsocks	Did you read the $SING LOI news? http:t.coNrH2obv7ti http:t.coLn1HqzOnok
2014-10-13 21:00:11	PennyStockMach	Did you read the $SING LOI news? http:t.coHx4mynnUDr
2014-10-13 21:00:12	JayBugster	Did you read the $SING LOI news? http:t.cob6brmQtl5t http:t.coeqgcRTP9NG
2014-10-13 21:00:12	Penny_Hotsocks	Did you read the $SING LOI news? http:t.cor8Au3hcEaL
2014-10-13 21:00:13	IonPennyStocks	Did you read the $SING LOI news? http:t.coWJ5vFU7CHB
2014-10-13 21:04:54	aheadsupotc	$SING received a new alert.See why at http:t.coFwU0sYdHLW $MSEZ #Penny #pennystocks [...]
2014-10-13 21:06:11	ckelly3	RT @aheadsupotc: $SING received a new alert.See why at http:t.coFwU0sYdHLW $MSEZ #Penny #pennystocks [...]
2014-10-13 22:00:00	PennyStockExcel	SinglePoint, Inc. Signs LOI to Acquire 100% of GreenStar Payment Solutions - $SING
2014-10-13 22:00:00	StockUltraman	SinglePoint, Inc. Signs LOI to Acquire 100% of GreenStar Payment Solutions - $SING
2014-10-13 22:00:01	Micro_Cap_Pro	SinglePoint, Inc. Signs LOI to Acquire 100% of GreenStar Payment Solutions - $SING
2014-10-13 22:00:01	IonPennyStocks	SinglePoint, Inc. Signs LOI to Acquire 100% of GreenStar Payment Solutions - $SING
2014-10-13 22:00:02	HotStockCafe	SinglePoint, Inc. Signs LOI to Acquire 100% of GreenStar Payment Solutions - $SING
2014-10-13 22:00:04	MicroCapUnivers	SinglePoint, Inc. Signs LOI to Acquire 100% of GreenStar Payment Solutions - $SING
2014-10-13 22:52:29	ckelly3	RT @JayBugster: SinglePoint, Inc. Signs LOI to Acquire 100% of GreenStar Payment Solutions - $SING

Notes: This table presents a selected sample of messages containing the keyword "$SING" published on Twitter after market close on October 13th, 2014.

Kim and Kim (2014) and Avery et al. (2016), among others), online investors are mostly bullish when sharing information about stock market on the Internet. Individual investor do not (often) 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 stock they hold or stock they want to buy using a bullish (positive) vocabulary. Furthermore, in our situation, the bullishness bias can also be related to fraudsters trying to pump the price of a stock by sharing (false) positive information about a given company on social media. Applying the previous methodology to all stocks listed on OTC Markets, and focusing our attention on event-day with a positive sentiment, we end up with a total of 567 events. The distribution of events over time does not exhibit any significant clustering on specific time period or day of the week.

To compute abnormal return, one has first to define a model for expected returns. However, choosing a daily model for normal returns of stock listed on OTC Markets is tricky, as even a five-factor model explains only 57.3% of the variation in the OTC market with monthly data (Ang et al., 2013). Furthermore, if a stock is under a manipulation scheme, we do not expect any model of normal return fitted during a non-manipulation estimation window to predict post-manipulation stock return. In order to define abnormal return, we thus use a simple market return model considering the NASDAQ MicroCap Index as benchmark of normal return.[12] 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 present our results using a 6-month estimation window and a 21 days event-window. On unreported robustness check, we find that our results are robust to a 12-month estimation window and a 11 days event-window.

For each event detected previously, we compute abnormal return for the estimation window [-130:-11] (L1 = 120 days) and event-window [-10:+10] (L2 = 21 trading days). We transform each abnormal return $AR_{i,t}$ to a rank variable $K_{i,t}$, by assessing to the day with the highest return over the complete window (estimation and event-window) a rank of +141, to the day with the second highest return a rank of +140, and so on until assigning to day with 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 in each firm's excess returns time series.

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

where M_i is the number of non-missing values for security i in L1 and L2. This yields

[12]Results are robust when using a constant mean return model or a capital asset pricing model.

order statistics for the uniform distribution with an expected value of one-half. The rank test statistic for day t (T_t) is equal to:

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

where N is equal to the number of event. The estimated standard deviation S(U) is defined on the estimation (L1) and event (L2) window as [13]

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

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

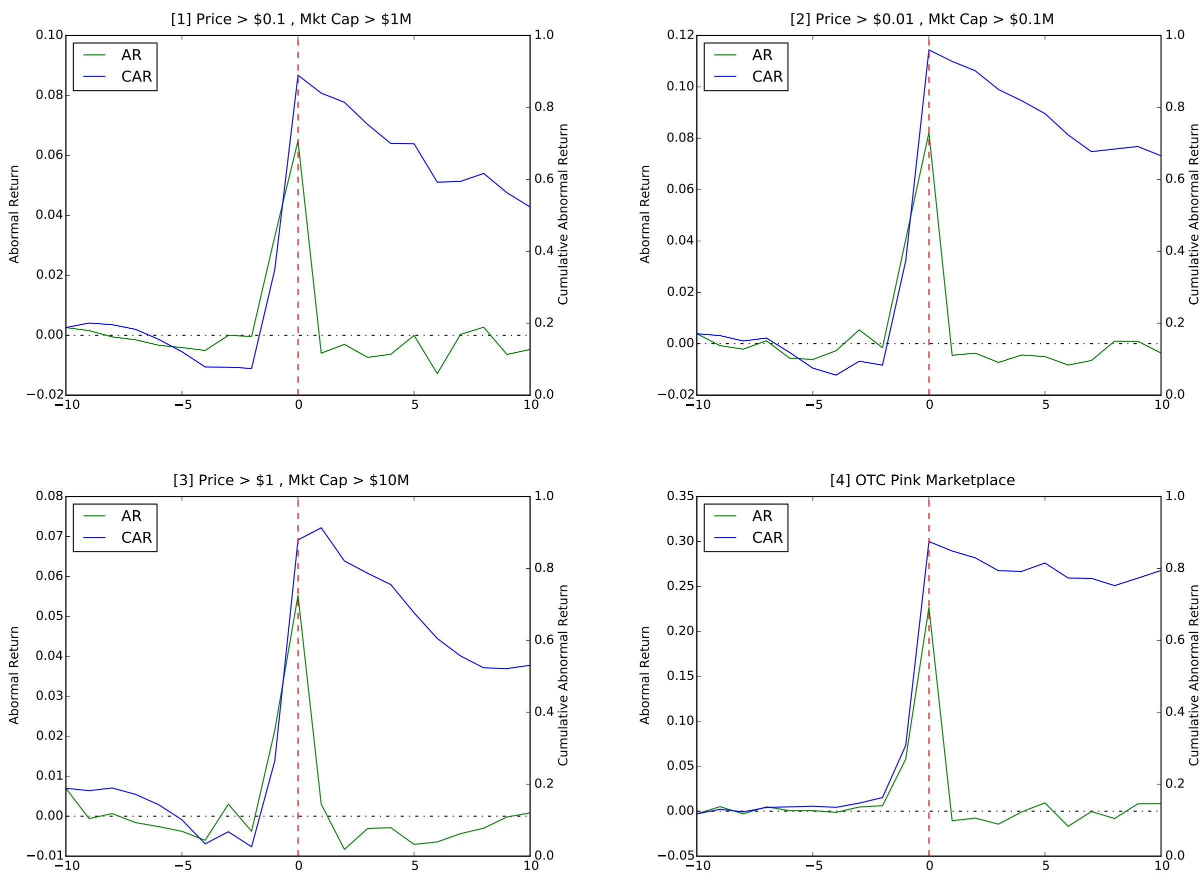
We conduct event-studies using various thresholds to include or not a stock depending on stock price, market capitalization and the percentage of non-trading days. More precisely, we use four filtering methods: [1] all stocks with a minimum price at the beginning of the event window of $0.1 and a market capitalization greater than $1,000,000 (as defined previously), [2] all stocks with a minimum price of $0.01 and a market capitalization greater than $100,000, [3] all stocks with a minimum price of $1 and a market capitalization greater than $10,000,000, and [4] all stocks listed on the OTC Pink marketplace with a price greater than $0.00001. We test the statistical significance of abnormal return on each day of the event-window and on each 5-day rolling interval to identify price reversal over a one week period. Table 5 summarizes results based on different filtering methods. Figure 4 presents abnormal return (AR) and cumulative abnormal return (CAR) during the [-10:+10] event-window, where day 0 is defined as an abnormally high activity on Twitter. Figure 5 shows the value of the non-parametric statistic computed by converting abnormal return to ranks on both the estimation and the event window.

As in Kim and Kim (2014), we identify a strong contemporaneous relationship between Twitter activity and stock price on the event day (*t0*). When analyzing stocks with a minimum price at the beginning of the event window of 0.1$ and a market capitalization greater than 1 million, we find an abnormal return of +6.49% on the event-day. This finding is consistent with Sabherwal et al. (2011) who find an increase of +13.93% on the event-day when defining event as an abnormal number of messages on the financial message board "TheLion.com". When focusing on stocks listed on OTC Pink Marketplace, we find a significant increase of +5.80% on the day before the event and +22.68% on the event-day.[14]

[13] We also consider multi-day version by multiplying by the inverse of the square root of the periods length.

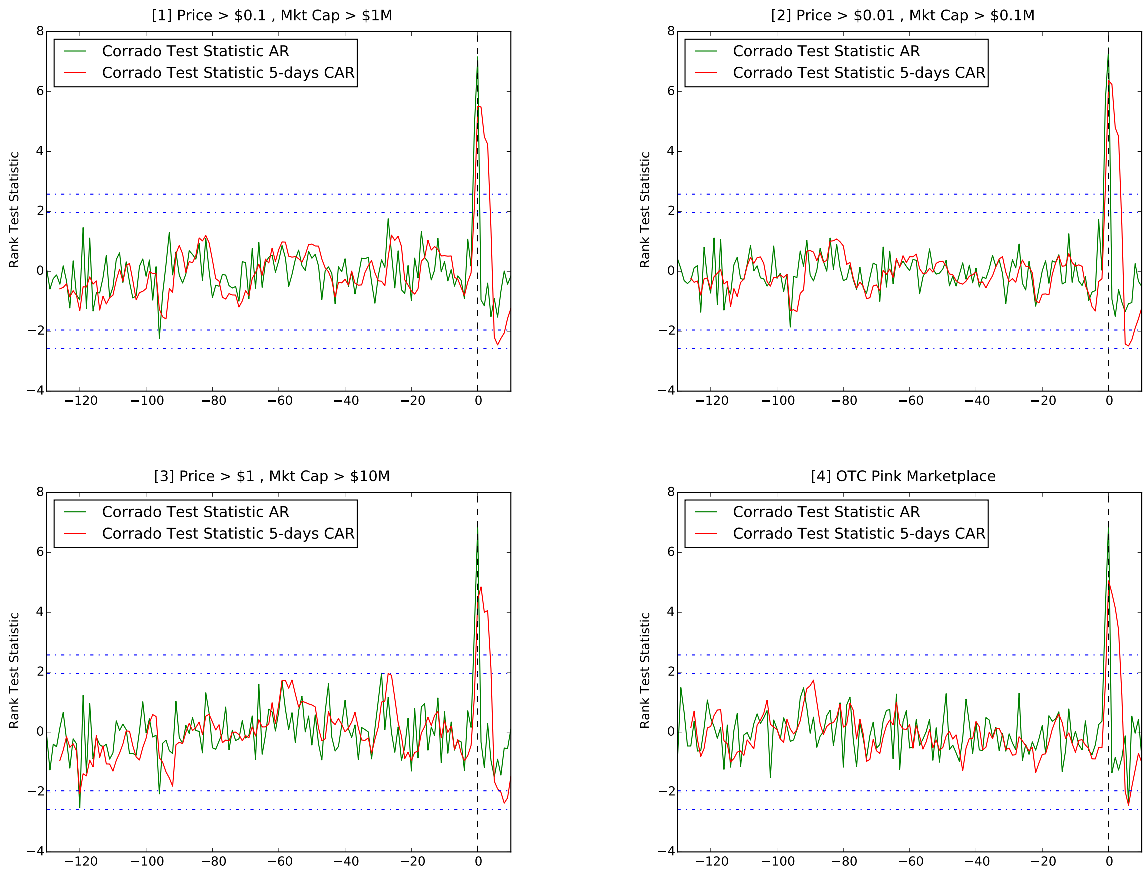
[14] On unreported robustness checks, we find that results are robust when we remove events with return on

Fig. 4. Event-Study - Abnormal Returns and Cumulative Abnormal Returns (%)



Notes: Figure 4 shows the one-day abnormal return and the cumulative abnormal return on a [-10:+10] days event-window around an abnormally high posting activity on social media about a stock listed on OTC markets. Results are presented for [1] stocks with a price greater than $0.1 and a market capitalization greater than $1,000,000, [2] stocks with a price greater than $0.01 and a market capitalization greater than $100,000, [3] stocks with a price greater than $1 and a market capitalization greater than $10,000,000, [4] all stocks listed on OTC Pink Marketplace with a price greater than $0.00001.

Fig. 5. Event-Study - Abnormal Returns and Cumulative Abnormal Returns - Rank Test



Notes: Figure 5 shows the one-day standardized average rank (green) and the 5-day rolling average rank (red) for both the estimation window [-130:-11] and the event-window [-10:+10]. Horizontal dashed blue lines represents significance threshold at the 5% level and 1% level. Results are presented for [1] stocks with a price greater than $0.1 and a market capitalization greater than $1,000,000, [2] stocks with a price greater than $0.01 and a market capitalization greater than $100,000, [3] stocks with a price greater than $1 and a market capitalization greater than $10,000,000, [4] all stocks listed on OTC Pink Marketplace with a price greater than $0.00001.

Table 5: Abnormal Returns and Cumulative Abnormal Returns (5-days) (%)

	[1]		[2]		[3]		[4]	
	AR	CAR	AR	CAR	AR	CAR	AR	CAR
t-10	0.0025	nan	0.0039	nan	0.0070	nan	-0.0029	nan
t-9	0.0015	nan	-0.0007	nan	-0.0006	nan	0.0051	nan
t-8	-0.0006	nan	-0.0021	nan	0.0007	nan	-0.0029	nan
t-7	-0.0015	nan	0.0012	nan	-0.0016	nan	0.0049	nan
t-6	-0.0034	nan	-0.0056	nan	-0.0026	nan	0.0007	nan
t-5	-0.0041	-0.0081	-0.0061	-0.0134	-0.0038	-0.0079	0.0007	0.0085
t-4	-0.0051	-0.0147	-0.0027	-0.0154	-0.0060	-0.0133	-0.0013	0.0021
t-3	-0.0001	-0.0142	0.0054*	-0.0079	0.0030	-0.0110	0.0047	0.0097
t-2	-0.0004	-0.0130	-0.0015	-0.0106	-0.0038	-0.0131	0.0062	0.0110
t-1	0.0330***	0.0233**	0.0406***	0.0356***	0.0215***	0.0110	0.0580***	0.0683
t0	0.0649***	0.0923***	0.0822***	0.1239***	0.0553***	0.0700***	0.2268***	0.2944***
t1	-0.0060	0.0914***	-0.0045	0.1221***	0.0030	0.0791***	-0.0105	0.2851***
t2	-0.0031	0.0884***	-0.0037	0.1130***	-0.0083	0.0678***	-0.0076	0.2728***
t3	-0.0074	0.0814***	-0.0073	0.1073***	-0.0031	0.0685***	-0.0144	0.2523***
t4	-0.0064	0.0421	-0.0044	0.0624	-0.0029	0.0441**	-0.0007	0.1936
t5	-0.0001	-0.0229**	-0.0050	-0.0248**	-0.0070	-0.0182	0.0093	-0.0239**
t6	-0.0128	-0.0297**	-0.0083	-0.0286**	-0.0064	-0.0277*	-0.0167**	-0.0301**
t7	0.0002	-0.0264**	-0.0065	-0.0315**	-0.0044	-0.0238**	-0.0003	-0.0228*
t8	0.0027	-0.0163**	0.0010	-0.0232*	-0.0030	-0.0237**	-0.0081	-0.0165
t9	-0.0064	-0.0164	0.0010	-0.0178	-0.0002	-0.0210**	0.0083	-0.0076
t10	-0.0048	-0.0211	-0.0036	-0.0164	0.0008	-0.0131	0.0086	-0.0083
Event	567		929		260		892	

Notes: This table shows the daily abnormal returns, relative to the social event day *t0*, on a [-10,+10] event-window. CAR is equal the sum of abnormal return on day t and the four previous days. $***$, $**$ and $*$ represent abnormal returns significance respectively at the 1%, 5%, 10% level using a Corrado rank test. Results are presented for [1] stocks with a price greater than $0.1 and a market capitalization greater than $1,000,000, [2] stocks with a price greater than $0.01 and a market capitalization greater than $100,000, [3] stocks with a price greater than $1 and a market capitalization greater than $10,000,000, [4] all stocks listed on OTC Pink Marketplace with a price greater than $0.00001.

More interesting, we find a significant post-event price reversal. Cumulative abnormal return is statistically significant and negative on an [+1:+5] window, with a post-event cumulative decrease in stock price between 2.5% and 3%. Again, this finding is consistent with Sabherwal et al. (2011) who identify a significant post-event decrease in stock price of -5.4% over the next 5 trading days. Two non-exclusive hypotheses can explain the price reversal pattern and the deviation from the efficient market hypothesis. First, we could conjecture that social media can be used as a proxy of investor over-optimism. In a market driven by unsophisticated trader with limits to arbitrage, price can deviate temporarily from its fundamental values in the presence of irrational sentiment-driven noise trader. In that case, the price reversal identified on OTC Markets is simply caused by "standard" investor sentiment, as in Tetlock (2007). Another explanation would be to conjecture that the sharp increase on the event day is caused by fraudsters or stock promoters pumping the price of targeted stocks, before dumping it on the following days after having made illegal profit. This hypothesis would also be consistent with the price reversal pattern identified on our event-studies.

the event-day greater or equal to +50%.

To disentangle partially between those two hypotheses, we conduct a network analysis by examining interactions between users in order to identify (if any) suspicious behaviors on Twitter.

6. Network Analysis

6.1. Analyzing interaction between users

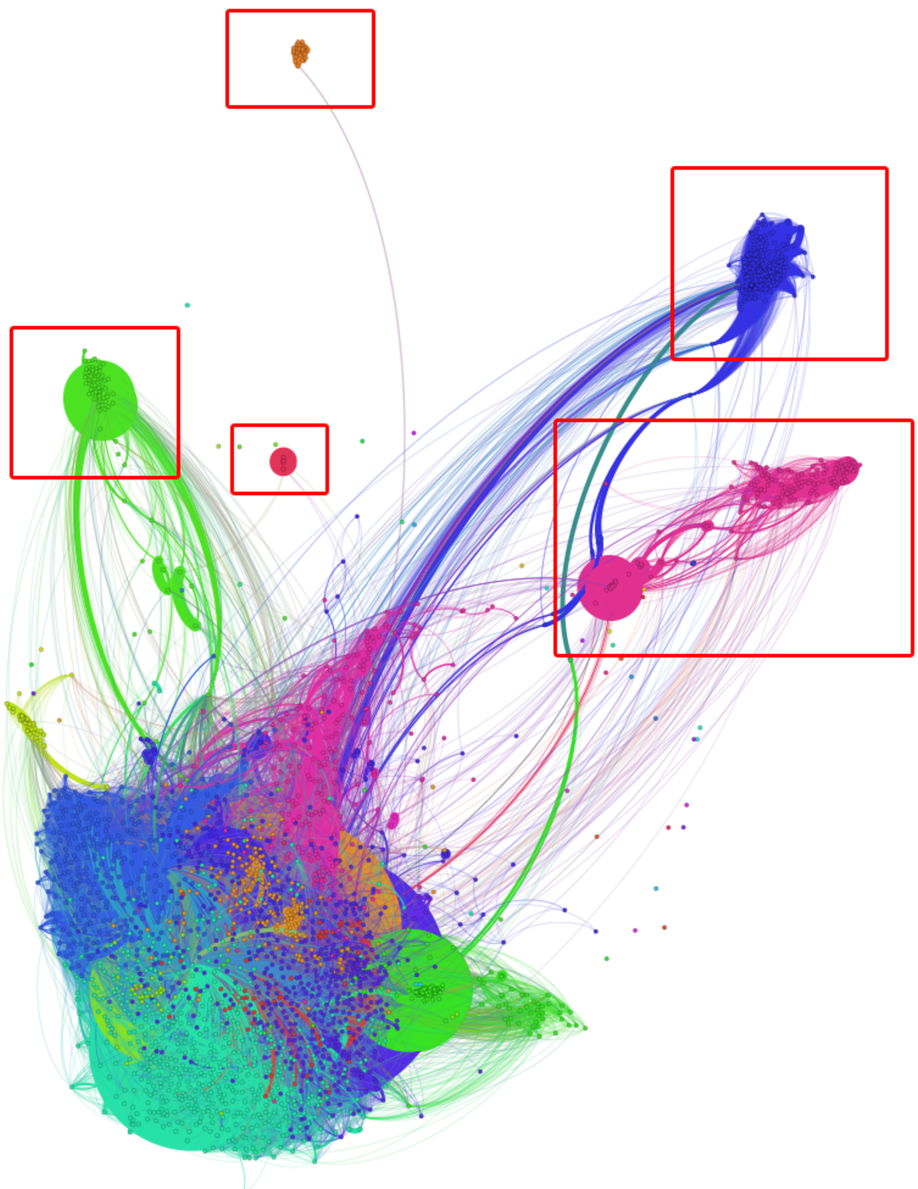
Disentangling an overoptimism effect from manipulation is a challenging issue, as the goal of a pump-and-dump scheme is exactly to exacerbate positive sentiment. However, analyzing directed interactions between users on a network can help identifying suspicious behaviors, as shown by Diesner et al. (2005) using the Enron email corpus. In that regards, Twitter offers an interesting framework, as interactions are directly observable through the functions "mention" and "retweet"

Twitter "retweet" function allows any user to share with its own list of followers any message created by another user. The "mention" function allows users to "tag" other members on a tweet to start a conversation with this user. The action of "retweeting" or "mentioning" can be considered as an interaction between two users. When user "A" who choose to retweet the original message posted by user "B" or to mention user "B" in a tweet, we can represent this interaction in a graph as a directed link between node A and node B. Then, as in any directed network, we can cluster users based on interactions similarities in order to identify potential suspicious behaviors. For example, if user "A" retweets all contents published by user "B", those two users would be clustered together. While clustering can also be caused by natural interactions or real friendships, an automatic approach helps identifying suspicious patterns before manually analyzing interactions to confirm (or invalidate) our hypothesis.

Over the 7,196,307 tweets from our database, we identify a total of 2,011,315 users' interactions (retweets or mentions). Removing all users with less than 50 directed entrant or outbound links, we end up with a network of 8,961 users and 205,093 directed links. Figure 6 shows Twitter network, where each nodes represents a user from our database and each directed link is a retweet or a mention from a user to another one. Clustering is based on directed links similarity and node size depends on the number of entrant links. Colors depend on modularity, an optimization methods for detecting community structure in networks (see Brandes et al. (2008)).

We identify five clusters characterized by a very high level of interactions between themselves and a low level of interactions with all the other clusters. The first cluster, in blue on

Fig. 6. Network Analysis of the Twittersphere based on retweets and mentions



Notes: Figure 6 shows interactions between all users on Twitter. Each node represents a user and each link (edge) an interaction between two users. Clustering is based on directed links similarities and node size depends on the number of entrant links. Colors depend on modularity, an optimization methods for detecting community structure in networks. Suspicious clusters are framed in red.

the top right corner of Figure 6 is composed of 481 users. This group is organized around stock promoters from the website http://stockmarketnews.co/ and the official account of a stock listed on OTC Markets (*$GPDB*, The Green PolkaDot Box). Analyzing all tweets containing the ticker *$GPDB* sent during our sample period, we find clear evidences of stock promotion.[15] For example, on July 21st 2015, 9,533 tweets containing the ticker *$GPDB* where published on Twitter by a total of 1,162 distinct users, without any specific news on that day. Analyzing tweet content, we find only 31 distinct messages, showing that a promotion scheme involving fake accounts and automatic posting was implemented. Activity on Twitter remains abnormally high on the next day (2,176 tweets) before collapsing afterwards (no tweet containing the ticker *$GPDB* where sent between July 23rd and July 29th). Another promotion campaign started on July 30th, with a total of 601 tweets.

The second cluster, in pink on the right side of Figure 6, is composed of accounts sharing an interest for cryptocurrencies. The user with the highest number of entrant links from this cluster is a user called CannabisCoins, a ”medical marijuana-backed digital currency” listed on OTC Markets (*$CANN*). Analyzing all tweets related to *$CANN* company, we identify various suspicious posting behaviors. For example, on October 9th 2014, between 3:13am and 3:22am, a tweet by CannabisCoin company announcing a future event was retweeted 735 times in 10 minutes. In this situation, the large peak in social activity was caused by a list of fake accounts retweeting automatically CannabisCoin's message to increase message's outreach and visibility.[16]

Analyzing other clusters, we find similar patterns. The most common anomaly is a very large peak in volume on social media caused by a large number of (fake) accounts posting or retweeting a message on Twitter about a given company. We also identify a very large number of users declaring themselves as stock promoters in their Twitter description. Some groups of promoters tend to act together on various occasion to tout a stock with a spamming method close to spam blast or fax blast.

As in the case of wash trade manipulation, where fraudsters use nominee accounts to create a fake trading activity by being both on the buy side and on the sell side, creating fake activity in social media can be a sign of a future stock market manipulation. While it is true that promoting a tweet can also helps a firm increases its sales or improve brand awareness (without any manipulation), we believe that investors should always be very cautious of any information about OTC stocks posted on social media. According to the Security and

[15] The vast majority of the accounts from this cluster have now been suspended by Twitter for spam or inappropriate behaviors.

[16] Creating fake attention by buying followers or retweeters is very easy on Twitter as some companies propose, without respecting Twitter Terms of Services, to use fake accounts to automatically retweet a message for a cost around $5 for 1,000 retweets.

Exchange Commission, "fraudsters can set up new accounts specifically designed to carry out their scam while concealing their true identities" and investors should "be skeptical of information from social media accounts that lack a history of prior postings or sending messages". We would add to the SEC recommendation that investors should also be skeptical of information published by any non-verified accounts and should watch carefully previous tweets from the user to try detecting any anomalies highlighted below (scheduled automatic postings, previous tweets not related to financial market, abnormal followers/retweets ratio...).

On the other hand, we also detect several users who dedicate their tweets to the detection of pump-and-dump schemes, often using proprietary algorithm to detect anomalies and to inform market participants. For example, a user named "ThePumpTracker" (now "theOTCtoday") used to publish alerts on Twitter when detecting that a stock was under promotion. Matching alerts from this Twitter account with days for which we identify an abnormal activity on Twitter, we find that around 10% of our events where also identified by "ThePumpTracker" as being related to a stock promotion. We would advice investor to also look out for users tracking pump-and-dump schemes and stock promotions before investing in OTC stocks.

While further research would be needed to better understand how information is disseminated on Twitter, we believe that our analysis improves our understanding of techniques used by fraudsters on social media and could help investors avoid penny stock scams on the Internet. Our finding reinforces the SEC recommendation that "investors who learn of investing opportunities from social media should always be on the lookout for fraud".

7. Conclusion

Social media can help investors gathering and sharing information about stock markets. But, at least for stocks listed on the OTC Markets, it also presents opportunities for fraudsters to send false or misleading statements to the marketplace. In this paper, we first analyze all SEC litigation releases by focusing our attention on pump-and-dump schemes. We find that information-based stock market manipulation mainly target small capitalization stocks listed on the OTC Markets. Fraudsters use various channel of communication to send false or misleading information to the marketplace, such as press releases, spam e-mails and websites. In that regards, and even if the number of reported case involving directly false information released on social media is for now relatively low, Twitter represents a very interesting channel for manipulators or stock promoters as it allows them to target a wide unsophisticated audience more prone to being scammed that 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.

We thus complement the literature on indirect empirical evidences of market manipulation by analyzing a novel dataset of more than seven millions messages published on Twitter during a one-year period. We provide empirical evidences showing that manipulators can use social media to artificially inflate the price of a stock. Defining event as an abnormally high posting activity on Twitter about a given company, we identify a large increase in stock price on the event-day, followed by a sharp price decrease over the next 5 trading days. Examining interactions between users (retweets and mentions), we identify suspicious clusters of Twitter users using fake accounts, automatic postings or scheduled retweets. While a judicial inquires would be needed to assess if the promotion is legal or not, our findings shed light on the need for a higher control of the information published on social media and a higher education for investors looking for trading opportunities on the Internet. Given the risk of manipulation and the average negative return of OTC stocks documented on the literature, we think that individual investors should be very cautious when choosing to invest on risky and illiquid small capitalization stocks.

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.

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.

Brandes, U., Delling, D., Gaertler, M., Gorke, R., Hoefer, M., Nikoloski, Z., Wagner, D., 2008. On modularity clustering. IEEE Transactions on Knowledge and Data Engineering 20, 172–188.

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.

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

Diesner, J., Frantz, T. L., Carley, K. M., 2005. Communication networks from the Enron email corpus it's always about the people. enron is no different. Computational & Mathematical Organization Theory 11, 201–228.

Drake, M. S., Roulstone, D. T., Thornock, J. R., 2012. Investor information demand: Evidence from google searches around earnings announcements. Journal of Accounting Research 50, 1001–1040.

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.

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.

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., 2016. Intraday online investor sentiment and return patterns in the US stock market. Working Paper, Université Paris 1 Panthéon-Sorbonne .

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.

Sprenger, T. O., Sandner, P. G., Tumasjan, A., Welpe, I. M., 2014. News or noise? Using twitter to identify and understand company-specific news flow. Journal of Business Finance & Accounting 41, 791–830.

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.

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 marchs 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.

Appendix B

U.S. SECURITIES AND EXCHANGE COMMISSION - Litigation Release No. 23401 / November 6, 2015

Securities and Exchange Commission v. James Alan Craig, Civil Action No. 3:15-cv-05076) (N.D. Cal.)

On November 5, 2015, the Securities and Exchange Commission filed securities fraud charges against a Scottish trader whose false tweets caused sharp drops in the stock prices of two companies and triggered a trading halt in one of them.

According to the SEC's complaint filed in federal court in the Northern District of California, James Alan Craig of Dunragit, Scotland, tweeted multiple false statements about the two companies on Twitter accounts that he deceptively created to look like the real Twitter accounts of well-known securities research firms. Also yesterday, the U.S. Attorney's Office for the Northern District of California filed criminal charges.

The SEC's complaint alleges that Craig's first false tweets caused one company's share price to fall 28 percent before Nasdaq temporarily halted trading. The next day, Craig's false tweets about a different company caused a 16 percent decline in that company's share price. On each occasion, Craig bought and sold shares of the target companies in a largely unsuccessful effort to profit from the sharp price swings.

The SEC's investigation also determined that Craig later used aliases to tweet that it would be difficult for the SEC to determine who sent the false tweets because real names weren't used. According to the SEC's complaint:

- On Jan. 29, 2013, Craig used a Twitter account he created to send a series of tweets that falsely said Audience, Inc. was under investigation. Craig purposely made the account look like it belonged to the securities research firm Muddy Waters by using the actual firm's logo and a similar Twitter handle. Audience's share price plunged and trading was halted before the fraud was revealed and the company's stock price recovered.
- On Jan. 30, 2013, Craig used another Twitter account he created to send tweets that falsely said Sarepta Therapeutics, Inc. was under investigation. In this case Craig deliberately made the Twitter account seem like it belonged to the securities research firm Citron Research, again using the real firm's logo and a similar Twitter handle. Sarepta's share price dropped 16 percent before recovering when the fraud was exposed.

The Commission's complaint charges that Craig committed securities fraud in violation of Section 10(b) of the Securities Exchange Act of 1934 and Rule 10b-5 thereunder. The complaint seeks a permanent injunction against future violations, disgorgement and a monetary penalty from Craig.

The SEC has issued an Investor Alert titled "Social Media and Investing - Stock Rumors" prepared by the Office of Investor Education and Advocacy. The alert aims to warn investors about fraudsters who may attempt to manipulate share prices by using social media to spread false or misleading information about stocks, and provides tips for checking for red flags of investment fraud.

The SEC's investigation was conducted by staff in the Market Abuse Unit including Elena Ro, John Rymas, and Steven D. Buchholz. The case was supervised by Joseph G. Sansone, Co-Chief of the Market Abuse Unit. The SEC's litigation will be led by Ms. Ro and John S. Yun of the SEC's San Francisco Regional Office. The SEC acknowledges the assistance of the U.S. Department of Justice and the Federal Bureau of Investigation.