

The International Journal of Banking and Finance, Volume 8 (Number 4) 2011: pages 1-18

THE IMPACT OF MANIPULATION IN INTERNET STOCK MESSAGE BOARDS

Jean-Yves Delort, Bavani Arunasalam, Henry Leung and Maria Milosavljevic
Google, Capital Markets CRC Ltd, University of Sydney and Australian Government
Office, Canberra

Abstract

Internet message boards are often used to spread information in order to manipulate financial markets. Although this hypothesis is supported by many cases reported in the literature and in the media, the real impact of manipulation in online forums on financial markets remains an open question. This paper is on the effect of manipulation in internet stock message boards on financial markets by employing a unique corpus of moderated messages to investigate market manipulation. Internet message board administrators use the process of moderation to restrict market manipulation. We find that manual supervision of stock message boards by moderators does not effectively protect Internet users against manipulation. By focusing on messages that have been moderated as manipulative due to ramping, we show ramping is positively related to market returns, volatility and volume. Stocks with higher turnover, lower price level, lower market capitalization and higher volatility are more common targets of ramping.

Key Words: Ramping, Market Manipulation, Internet Stock Message Boards, Event Study

JEL classification: D82, G14, G22

1. Introduction

Internet stock message boards provide an effective medium for investors to communicate, to disseminate and discover new information. However, the value and *integrity* of internet stock message boards are often criticized and investors are warned not to trade on the information provided (Fraser, 2007). In order to address these issues, regulators often dictate that message board providers conform to specific guidelines.¹ Nevertheless, people continue to use Internet forums for their own purposes, and financial surveillance analysts

¹ For example, in Australia, see Australian Securities & Investments Commission (ASIC) Regulatory Guide 162, <http://www.asic.gov.au>.

at regulators and exchanges are known to use message boards for investigating suspicious trading activities.

HotCopper² is the most popular internet stock message board in Australia and forms the focus of our study. This site operates under a strict code of conduct which prohibits unethical or illegal use of the forum. Administrators moderate postings which do not conform to the guidelines, and can revoke access for users who consistently disregard the rules. Moderated posts are labeled with the reason for moderation, for example *ramping*, *flaming* or *profanity*. These posts serve to proxy manipulative behaviour. Using an event study methodology, we investigate whether ramping is related to stock price volatility, trade volume and returns. Further, we study the firm characteristics targeted by rampers.

First we find that manual moderation of messages cannot effectively prevent ramping to take place in internet stock messages. This result is supported by the high number of posts that are moderated due to ramping every day, which indicates that ramping is a common phenomenon. In addition, we find that the lead time to moderation is approximately 8 hours, allowing sufficient time for these posts to impact the market. Our results also demonstrate that ramping is positively related to market returns, volatility and volume. Firms with higher turnover, lower price level, lower market capitalization and higher volatility also receive a higher proportion of ramping posts suggesting that they are more common targets of ramping.

The remainder of this study is organized as follows. In Section 2 we present an overview of previous works. This serves as motivation for the development of our hypotheses in section 3. In section 4, we describe the data and empirical methodology used in our analysis. The results are reported and discussed in section 5, and section 6 concludes by providing some areas for further work.

2. Literature review

Stock market manipulation is generally defined as the creation and exploitation of arbitrage

² <http://www.hotcopper.com.au>.

opportunities (Zigrand, 2006). Manipulation of stock prices may occur through direct trading strategies or indirectly via the dissemination of distorted price sensitive information. The former may be observed when trades are intentionally placed in the wrong direction or short term losses are being undertaken to move prices in the desired direction (Chakraborty and Ylmaz, 2004). The latter may occur through mediums such as investment blogs, spam emails and online stock message forums (Antweiler and Frank, 2004; Das and Sisk, 2005; Aggarwal and Wu, 2006). However, the detection of both types of manipulative behaviour remains difficult due to the lack of proxy for the occurrence of manipulation.

The impact of trade strategy manipulation can be observed in the market response to such strategies. For example, Aggarwal and Wu (2006) demonstrate that stock market manipulation cases pursued by the U.S. Securities and Exchange Commission (SEC) during 1990 through October 2001 were associated with greater stock volatility, greater liquidity and higher returns. Other evidence of stock market reaction to manipulation is documented by Hillion and Suominen (2004), who highlight a general rise in volatility, trading volume and one-minute returns in the closing price of stocks in Paris Bourse during January and April, 1995. Further, the proportion of partially hidden orders increased during this period.

Evidence of manipulation has also been found in certain types of derivatives markets such as the futures market. It is shown that uninformed investors earn positive profits by creating a futures position and simultaneously trading in the spot market. As a result, arbitrage free profit opportunities are exploited to derive profitable cash settlement at the time of delivery (Kumar and Seppi, 1992; Merrick et al., 2005).

Additional evidence of trade strategy manipulation include those found by Guo et al. (2008), who uses low earnings quality firms to proxy for firms' intention to manipulate. They find that these firms tend to use stock splits to manipulate equity values before acquisition announcements. On the other hand, Eom et al. (2009) find that investors strategically place spoofing intraday orders in the Korean Exchange. Spoofing orders have little chance of being executed, and are used to mislead other traders of an order book imbalance and thereby influence the direction of stock price movement. They conclude that stocks targeted for manipulation had higher return volatility, lower market capitalization, lower price level, and lower managerial transparency. Finally, large traders

have also been found to carry out manipulation due to its scale of operation and its ability to better sequence and time trades (Gastineau and Jarrow, 1991; Allen et al., 2006).

Manipulative behaviour involving the intentional misinterpretation and dissemination of price sensitive information is more difficult to document. Bagnoli and Lipman (1988) address the problem of market manipulation through the announcement of a takeover bid. They develop a model where the manipulator takes a position, announces a takeover bid, and then unwinds his position. In their attempt to elucidate market manipulation from the 1600s to the internet technology frauds of today, Leinweber and Madhavan (2001) show that, through technology, information can be quickly disseminated to influence investor perception on particular stocks. Stock spam emails and posting on internet message boards are two major avenues manipulators employ to affect investor sentiment because of their global reach and high degree of accessibility to investors. In particular, significant reactions of trading volume and returns have been shown to respond to spam campaigns (Bohme and Holz, 2006). Hanke and Hauser (2008) then extends the works of Leinweber and Madhavan (2001) and Bohme and Holz (2006) to provide corroborating evidence that stock spam e-mails significantly and positively impact volatility and intraday spread. Similar evidence is noted by Hu et al. (2009), who found that pump and dump email campaigns show a statistically significant decline in stock price from the peak spam day to the following day.

Evidence of online impact of message board posting on stock activity is well documented in prior literature (Antweiler and Frank (2004); Das and Sisk (2005); Aggarwal and Wu (2006); Cook and Lu (2009)). Antweiler and Frank (2004) use computational linguistics methods to analyze the effect of message sentiment (buy, sell or hold) in forums to determine whether the information content of posts influence the stock market. They find that stock messages help predict volatility and disagreement between messages to be associated with increased trading volume. However, they find no correlation of message sentiments with the direction of returns. By improving sample selection and removing noise caused by program generated sentiments, Cook and Lu (2009) find that the bullishness of board messages positively and significantly predict abnormal stock return up to 2 days ahead. When taking poster's credibility into account, they also find that the board messages' predictive power over stock returns becomes much stronger in terms of both economic magnitude and significance.

These studies demonstrate stock markets are likely to be sensitive to Internet message board activities. Moreover, many anecdotal evidences of market manipulation in Internet stock message boards have been reported in the financial literature and in the media (Leinweber and Madhavan, 2001; Harris, 2002). However, the phenomenon has not been investigated through a rigorous approach. To the best of our knowledge, this work provides the first quantitative analysis of market manipulation in Internet stock message boards.

3. Hypotheses development

This research employs moderated postings unique to the Australian HotCopper online message forum. Improving upon prior literature such as Das and Sisk (2005), Chemmanur and Yan (2009) and Brown and Cliff (2004), HotCopper moderated postings allow us to proxy for manipulative behaviour directly. This is a unique opportunity available to us.

Stock ramping through forums occurs when a ramper recommends a stock and highlights its huge potential to rise in share price in the very near future. The ramper buys a large quantity of a stock, and posts information to influence the masses into buying the stock with the aim to drive up the share price. By implying that readers of these posts are in possession of privileged information - such as insider knowledge of an impending takeover offer - rampers seek to persuade the gullible into purchasing a particular stock. If a significant enough number of easily-led individuals invest in the touted stock, a ramper can “ramp up” the share price so that he or she could sell their shares at a quick profit. Ramping can be both up or down.

Stock message board administrators use the moderation process to restrict the manipulations of stock prices through ramping posts as a service to Internet users. However, the wide accessibility of internet technology and the high speed at which investors absorb news may diminish the effect of moderation. Therefore we want to test the following hypothesis:

Hypothesis 1: Manual moderation of stock message boards, as a monitoring strategy, does not protect users against manipulation.

The direct measure of ramping through HotCopper moderated posts allow us to investigate the impact of ramping on stock market activity. It has been shown in the

literature that well advertised stocks are related to higher stock returns (Das and Sisk, 2005; Chemmanur and Yan, 2009; Brown and Cliff, 2004). Based on the investor attention hypothesis proposed in Chemmanur and Yan (2009), which implies that high levels of stock advertising is associated with larger stock returns, stocks heavily promoted to investors via ramping postings may also generate higher returns. One possible explanation proposed by Barber and Odean (2008) may be investors' need to select stocks out of the many available the ones that attract their attention will be the ones most likely to be included in their investment portfolio.

Furthermore, extreme news announcements have been shown to produce high trade volume (Tetlock, 2007). Similarly, ramping posts, as a type of extreme news announcements, may potentially have the same effect. Both stock spam e-mails and messages in online message forums rely on the internet technology to increase its reach to investors (Leinweber and Madhavan, 2001). Drawing upon prior works (Bohme and Holz, 2006; Hanke and Hauser, 2008; Hu et al., 2009) on the impact of stock spam e-mails on trade volatility and intraday spread, we hypothesize that manipulators may employ online message forum ramping to influence investor sentiments in a similar fashion. If ramping posts does indeed have an impact on investor sentiment prior to moderation, we may form the following hypothesis to test the relation between moderated ramping posts and stock market response:

Hypothesis 2: Ramping in internet stock message boards affects stock price volatility, volume and returns.

Finally, we attempt to describe the characteristics of rampers' target firm stocks. Since manipulation is more effective when there is greater uncertainty about the value of a firm, we expect that stocks with a higher return volatility would attract more ramping attempts (Eom et al., 2009). Firm size has also been shown to be significantly positively associated with the level of disclosure (Alsaeed, 2006). This can be explained by the fact that larger publicly listed firms receive higher analyst coverage. On the other hand, smaller firms receive less coverage and thus are more susceptible to price movements driven by private information release. To that end, we expect smaller firms to be more prone to manipulation through ramping. Hanke and Hauser (2008) report that liquidity is a significant characteristic of companies targeted by spammers. The lower the liquidity, the larger the price impact. In addition, the impact of spamming on trade volume is markedly higher for

low-turnover stocks. The reaction of high-turnover stocks to spamming is generally lower because it takes a greater number of participants to move prices. To similar effect, we expect rampers to target low-turnover stocks (low liquidity). Therefore, in order to investigate the characteristics of firms targeted by rampers, we form the following hypothesis.

Hypothesis 3: Stocks targeted for manipulation have higher return volatility, lower market capitalization, lower price level, and lower turnover.

4. Data and methodology

4.1. Data

The data for this study comes from HotCopper, Australia's largest internet stock message board. The time period runs from January 2008 to December 2008 inclusive. Messages were downloaded from the HotCopper website using software written by the authors, and we restricted our collection of data to only include messages that have a ticker symbol field representing an ASX listed company. Each message contains the following fields: date, time, author, ticker symbol and content. In total, the data set contains 1,146,223 messages for 1,825 firms listed on the Australian Securities Exchange (ASX).

An interesting feature of the HotCopper forum is that the moderated posts are labeled with the main reason of their moderation. In accordance with Australian law regarding all information available in the securities market, HotCopper users are expected to comply with a set of strict usage guidelines.³ Messages that do not comply may be moderated by a volunteer or by an administrator.⁴ Such messages are not removed from the forum, but their content is replaced by a message, which contains the following information: moderation time, moderation type, and a comment specifying the reasons for moderation. The full set of moderation types with their respective percentage composition are listed in Table 1. Each type is a simple phrase that describes the primary reason for moderation. Some examples of textual comments are provided below:

³ User Posting Guidelines, <http://www.HotCopper.com.au/postingguidelines/>.

⁴ List of moderators, http://www.HotCopper.com.au/forum_modList.asp.

- All you seem to be doing is plaguing threads with your down-ramping on non-held stocks. Find something else to do with your time instead of wasting others. Snide remarks on stocks are not considered helpful or useful posts.
- Rumors aren't required thanks. They usually lead to misinformation on a public forum. No more like this, thanks.
- This post is being moderated because of the unsubstantiated information particularly "Every deal done falls through". This just isn't true. Please refrain from flaming other posters too.

In addition to the HotCopper data, we obtained the daily closing prices, high and low intraday prices and volume data for all companies listed on the ASX in 2008 from Reuters. We excluded all firms that had fewer than 50 posts during 2008 from our data set. Our final sample size consists of 1,083,913 posts for 938 firms. Note that results presented in Table 1 are based on the final sample size.

Table 1: Moderation types found on HotCopper in 2008

| Moderation cause | Total | Total |
|-------------------|--------|-------|
| Other | 6,379 | 45.1% |
| Flaming | 3,583 | 25.3% |
| Ramping | 1,519 | 10.7% |
| Profanity | 1,207 | 8.5% |
| Spamming | 389 | 2.8% |
| Defamatory | 351 | 2.5% |
| Unknown Reason | 159 | 1.1% |
| Advertising | 152 | 1.1% |
| Unlicensed Advice | 72 | 0.5% |
| Duplicate | 73 | 0.5% |
| Insider Trading | 73 | 0.5% |
| Copyright | 63 | 0.4% |
| Sexist | 43 | 0.3% |
| Racist | 43 | 0.3% |
| Blasphemous | 33 | 0.2% |
| Total | 14,139 | |

4.2. Methodology

We apply an event study methodology to analyze the market impact of ramping in

internet stock message boards. We define a trade event for a certain stock as a day on which the market is open for trading in that stock. We further define a posting event for a certain stock as a day on which at least one message has been posted in the message board discussing this stock. Similarly, we define a moderation event for a certain stock as a day on which at least one message discussing this stock has been moderated. Finally, we define a ramping event for a certain stock as a day on which at least one message has been moderated because of an attempt to ramp this stock.

The performance of a stock is measured in terms of the dependent variables defined by Hanke and Hauser (2008) and also used by Hu et al. (2009), that is, stock return, volatility, intraday volatility and turnover. Their definitions of these variables are shown in Figure 1.

Figure 1: Dependent variables defined by Hanke and Hauser (2008)

Stock Return:

$$RET_{i,t} = \ln S_{i,t} - \ln S_{i,t-1} \quad (1)$$

where $S_{i,t}$ is the closing price of stock i on day t .

Volatility:

$$SIG_{i,t} = \sqrt{\frac{(RET_{i,t} - \overline{RET}_i)^2}{\sigma_i^2}} \quad (2)$$

where \overline{RET}_i is the average $RET_{i,t}$ over the sample period; and σ_i^2 is the average variance of $RET_{i,t}$ over the sample period.

Intraday volatility:

$$SPR_{i,t} = \frac{\varsigma_{i,t}}{\overline{\varsigma}_i} \quad (3)$$

where $\overline{\varsigma}_i$ is the average intraday price range stock i over the sample period; and $\varsigma_{i,t}$ is the intraday price range of stock i on day t given by:

$$\varsigma_{i,t} = \ln(\text{intraday high}) - \ln(\text{intraday low})$$

Turnover:

$$VOL_{i,t} = \ln \left[1 + \frac{to_{i,t}}{\overline{to}_i} \right] \quad (4)$$

where $to_{i,t}$ is the turnover of stock i on day t and \overline{to}_i is the average turnover of stock i over the sample period.

As discussed in Section 4.1, HotCopper posts may be moderated due to several reasons, but each moderated post can only be labeled with one of the moderation types given in Table 1. This restriction and the close similarity between some of the moderation types may result in the ramping posts being labeled differently. For example, a post which could be labeled as both ‘Unlicensed Advice’ and ‘Ramping’ may only be labeled as ‘Unlicensed Advice’. Posts that are moderated due to multiple reasons are often labeled as ‘Other’, which explains the high percentage of moderation under ‘Other’ (45.1%) as given

in Table 1.

We found a significant number of ambiguous moderated posts in which the moderators assign one label but mention several causes for moderation in the comment field. Since our focus is on investigating the impact of ramping, it is important that we incorporate all posts that could potentially be seen as ramping. As a result, we include posts that mention the keyword ‘ramp’ or ‘unlicensed advice’ in the moderator’s comment as an additional explanatory variable in our regression analysis.

Figure 2: Regression Model for the Impact Study

$$Y_{i,t} = \beta_0 + \beta_1 Vol_{i,t} + \beta_2 Market_t + \beta_3 Comment_{i,t} + \beta_4 Info_{i,t} + \sum_{m \in L} \beta_m Label_{i,m,t} + \epsilon_{i,t} \quad (5)$$

where:

$Y_{i,t}$ - the dependent variable that takes the values of $RET_{i,t}$, $SIG_{i,t}$, $SPR_{i,t}$ or $VOL_{i,t}$

$Vol_{i,t}$ - traded volume of stock i on day t

$Market_t$ - price of S&P/ASX200 on day t - used as a control variable for the market effect

$Comment_{i,t}$ - number of moderated posts regarding stock i on day t that mention ‘ramp’ or ‘unlicensed advice’ in the moderator’s comment

$Info_{i,t}$ - a dummy variable that indicates the release of a company announcement for stock i on day t

L - the set of moderation types given by

{Advertising, Blasphemous, Copyright, Defamatory, Duplicate, Flaming, Insider Trading, Other, Profanity, Racist, Ramping, Sexist, Spamming, Unknown Reason, Unlicensed Advice}

$Label_{i,m,t}$ - number of posts regarding stock i that were moderated with the label $m \in L$ on day t

and β_0 and $\epsilon_{i,t}$ are the intercept and a random error term respectively.

The regression equation for this impact study is given in Figure 2. The market and volume are used as added control variables. We also included in our model a dummy variable that takes the value 1 if a company announcement for stock i was released on day t and 0 otherwise. We include this in our model in order to control for the impact of company announcements on the performance of stocks.

As outlined in section 3, particular types of firms are more likely to be targeted for manipulation than others. We use the regression model shown in Figure 3 in order to investigate the characteristics of the firms targeted by rampers. In this regression, the dependent variable is the probability for a post to be moderated because of ramping, which we take as a proxy for the probability for the stock of being ramped.

Figure 3: Regression Model for determinants of Target Firms

$$Y_i = \beta_0 + \beta_1 Vol_i + \beta_2 MarketCap_i + \beta_3 Price_i + \beta_4 SIG_i + \epsilon_i \quad (6)$$

where the explanatory variables are:

Vol_i - average traded volume of stock i during the sample period.

$Price_i$ - average closing price of stock i during the sample period.

$MarketCap_i$ - market capitalization of stock i at the end of the sample period.

SIG_i - average volatility of stock i during the sample period calculated as described in Figure 1

and β_0 and ϵ_i are the intercept and a random error term respectively.

The dependent variable is the proportion of ramping posts, and is calculated as:

$$Y_i = \frac{Ramp_i}{Mod_i} \quad (7)$$

where

$Ramp_i$ - the number of ramping posts of stock i during the sample period.

Mod_i - the number of moderated posts of stock i during the sample period.

5. Results and Discussion

5.1. Descriptive statistics

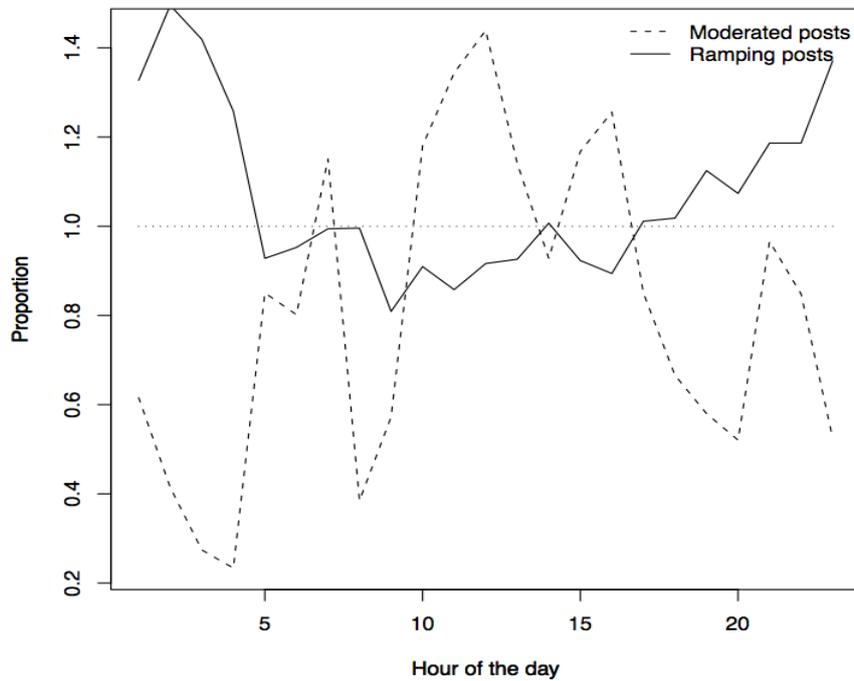
Moderation activity: We define moderation delay as the difference between publication time and moderation time for a particular post on an internet message board. In 2008, for the 938 stocks included in the sample, the mean moderation delay was 7.95 hours ($\sigma = 15.93$). In addition, 40.78 per cent of moderated posts were moderated after two hours. This is sufficient time for ramping on forums to be effective, which supports Hypothesis 1.

We found that in general, inappropriate posts tend to be published outside of normal trading hours. This is demonstrated by the graph shown in Figure 4, which represents the proportion of moderated posts and with respect to the hour of day. What is most interesting here is the daily pattern of posts, which are moderated because of ramping. The graph identifies quite clearly that ramping is more active during trading hours, which is quite different to the general pattern for moderated posts. We believe that the reason for this is that ramping messages created outside of normal trading hours are less likely to be effective since they are more likely to be moderated before market opening.

This figure plots the proportions of moderated posts and ramping posts divided by the

proportion of posts with respect to the hour of the day.

Figure 4: Ramping versus Moderation Activities



This figure plots the proportions of moderated posts and ramping posts divided by the proportion of posts with respect to the hour of the day.

Dependent variables: Table 2 presents descriptive statistics for our sample. In 2008, for the 938 stocks included in the sample, we have 238,252 trade events, 100,695 posting events, 4,870 moderation events and 768 ramping events. The number of ramping events reported is only a lower bound to the actual number of ramping events. With the increasingly high volume of messages that are posted on internet stock forums, it is likely that some of the inappropriate posts will not be discovered in the manual moderation process. However, these posts are not taken into account in our analysis. For example, the following unmoderated posts could be considered as ramping:

- I heard a rumor that the plant is still being repaired and may be operating again later today.
- I am focusing on oil stocks now, not affected by all this, and also I heard a rumor even General Motors now admit our oil has run out on what we can produce in one day and now in decline, which is good news to the oil stocks.

Table 2 shows that ramping events have the highest mean for returns, volatility, intra-day volatility and volume which supports Hypothesis 2. Ramping events have the lowest

mean price of the four types of events.

Table 2: Descriptive Statistics

| | Nb. events | Price | RET | SIG | SPR | VOL |
|-----------------------------------|------------|--------|---------|--------|--------|--------|
| Panel A: Trade event | 238252 | | | | | |
| Mean | | 1.9820 | -0.0052 | 0.0785 | 0.3747 | 1 |
| Median | | 0.24 | 0 | 0.3747 | 1 | 1.4717 |
| St.dev. | | 7.7096 | 0.0785 | 0.7776 | 0.7116 | 0.518 |
| Panel B: Posting events | 100695 | | | | | |
| Mean | | 2.3754 | -0.0022 | 0.0878 | 0.4817 | 1.1994 |
| Median | | 0.265 | 0 | 0.4817 | 1.1994 | 1.6403 |
| St.dev. | | 9.7972 | 0.0878 | 0.8489 | 0.8886 | 0.6476 |
| Panel C: Moderation events | 4870 | | | | | |
| Mean | | 2.7410 | -0.0053 | 0.1151 | 0.5474 | 1.3636 |
| Median | | 0.245 | 0 | 0.5474 | 1.3636 | 2.6274 |
| St.dev. | | 9.1831 | 0.1151 | 1.019 | 0.9844 | 0.7723 |
| Panel D: Ramping events | 786 | | | | | |
| Mean | | 1.6275 | 0.0027 | 0.1395 | 0.6446 | 1.8063 |
| Median | | 0.19 | 0 | 0.6446 | 1.8063 | 5.5706 |
| St.dev. | | 5.7589 | 0.1395 | 1.2379 | 1.1167 | 0.9636 |

This table presents the mean, median and standard deviation for the dependent variables namely, Return (RET), Volatility (SIG), Intraday volatility (SPR) and Turnover (VOL). Panel A presents the statistics for all the firms across all the days for January 2008 through December 2008. Panels B, C and D present the same statistics only for days with at least one post, days with at least one moderated post and days with at least one ramping post respectively.

It indicates that the “pump and dump” interest is the highest for small-priced firms. Interestingly, the mean returns for panel A, B and C are negative. A possible explanation could be the beginning of the Global Financial Crisis that hit the market in the last quarter of 2008. However, the mean return for ramping events is positive, which shows that ramping co-occurs with price increase.

In sub-section 5.3 we further investigate the impact of ramping on the four dependent variables taking into account the firm characteristics. This table presents the mean, median and standard deviation for the dependent variables namely, Return (RET), Volatility (SIG), Intraday volatility (SPR) and Turnover (VOL). Panel A presents the statistics for all the firms across all the days for January 2008 through December 2008. Panels B, C and D present the same statistics only for days with at least one post, days with at least one moderated post and days with at least one ramping post respectively.

Impact of manipulation in IMS: Table 3 shows the results of the regression analysis for the model given in Figure 2. The results show that the independent variables Volume, Market, the release of company Announcements and Ramping are significantly correlated to return, volatility, intra-day volatility and volume. The correlation coefficients are

positive for the four variables, but the confidence is stronger for volatility, intra-day volatility and volume. Insider Trading is also significantly correlated to volume, intra-day volatility and return.

Table 3: Impact Study Results

| Variable | RET | | | SIG | | | SPR | | | VOL | | |
|--------------------|---------|---------|-----|---------|----------|-----|---------|----------|-----|---------|----------|-----|
| | Est. | t-stat | | Est. | t-stat | | Est. | t-stat | | Est. | t-stat | |
| Intercept | 0.0166 | 15.2390 | *** | 6.9634 | 86.5380 | *** | 4.6550 | 29.8520 | *** | -3.3960 | -64.8660 | *** |
| Volume | 0.0009 | 17.6740 | *** | 0.0550 | 170.8620 | *** | 0.0992 | 157.6600 | *** | | | |
| Market | 0.0044 | 22.1930 | *** | -0.8165 | -85.6260 | *** | -0.5583 | -30.2240 | *** | 0.4563 | 74.1110 | *** |
| Ramping | 0.0131 | 2.5530 | * | 0.1428 | 4.2970 | *** | 0.2816 | 4.2890 | *** | 0.2560 | 11.5920 | *** |
| Spamming | -0.0163 | -2.1080 | * | -0.0460 | -0.8840 | | -0.1540 | -1.4860 | | 0.0004 | 0.0120 | |
| Insider Trading | 0.0129 | 0.8020 | | 0.1129 | 0.9590 | | 1.2281 | 4.8220 | *** | 0.3740 | 4.7730 | *** |
| Other | 0.0047 | 2.0800 | * | -0.0233 | -1.6500 | | -0.0968 | -3.4970 | *** | 0.0979 | 10.4490 | *** |
| Unknown Reason | 0.0033 | 0.2660 | | 0.0306 | 0.3930 | | 0.0187 | 0.1250 | | 0.0745 | 1.4370 | |
| Unlicensed Advice | 0.0047 | 0.2450 | | -0.0913 | -0.7510 | | -0.2059 | -0.8770 | | 0.0652 | 0.8050 | |
| Flaming | -0.0061 | -2.1050 | * | -0.0039 | -0.2130 | | 0.0065 | 0.1810 | | 0.0699 | 5.7930 | *** |
| Defamatory | -0.0175 | -2.1110 | * | 0.0360 | 0.6720 | | -0.1362 | -1.3150 | | 0.0016 | 0.0460 | |
| Duplicate | 0.0211 | 0.9990 | | -0.0148 | -0.1240 | | -0.1753 | -0.7450 | | 0.0413 | 0.5200 | |
| Advertising | 0.0102 | 0.8230 | | -0.0202 | -0.2410 | | 0.0408 | 0.2520 | | 0.2176 | 3.9640 | *** |
| Profanity | -0.0061 | -1.3240 | | 0.0389 | 1.3000 | | 0.1073 | 1.8130 | | 0.2227 | 11.2000 | *** |
| Copyright | -0.0118 | -0.6230 | | 0.0070 | 0.0540 | | -0.1801 | -0.7090 | | 0.0197 | 0.2300 | |
| Ramping in comment | -0.0004 | -0.0630 | | 0.0813 | 2.0880 | * | 0.2430 | 3.1460 | ** | 0.0472 | 1.8220 | |
| Company info | 0.0080 | 13.2160 | *** | 0.0736 | 19.1700 | *** | 0.1368 | 18.4040 | *** | 0.1641 | 65.0160 | *** |
| R^2 | | 0.010 | | | 0.1266 | | | 0.1044 | | | 0.0441 | |
| n | | 104375 | | | 233585 | | | 228075 | | | 238236 | |

This table reports the results of the regression analysis with four dependent variables: Return(RET), Volatility (SIG), Intraday volatility (SPR) and Turnover(VOL) in turn. $RET_{i,t} = \ln S_{i,t} - \ln S_{i,t-1}$ where $S_{i,t}$ is the closing price of stock i on day t . $SIG_{i,t} = \sqrt{\frac{(RET_{i,t} - \overline{RET}_i)^2}{\sigma_i^2}}$ where \overline{RET}_i is the average $RET_{i,t}$ over the sample period and σ_i^2 is the average variance of stock i over the sample period. $SPR_{i,t} = \frac{\varsigma_{i,t}}{\overline{\varsigma}_i}$ where $\varsigma_{i,t}$ is the intraday price range of stock i on day t given by $\varsigma_{i,t} = \ln(\text{intraday high}) - \ln(\text{intraday low})$ and $\overline{\varsigma}_i$ is the average intraday price range stock i over the sample period. $VOL_{i,t} = \ln \left[1 + \frac{to_{i,t}}{\overline{to}_i} \right]$ where $to_{i,t}$ is the turnover of stock i on day t and \overline{to}_i is the average turnover of stock i over the sample period. The explanatory variables are as follows : $Volume_{i,t}$ is the traded volume of stock i on day t . $Market_t$ is the price of S&P/ASX200 on day t . Ramping, Spamming, Insider Trading, Other, Unknown Reason, Unlicensed Advice, Flaming, Defamatory, Duplicate, Advertising, Profanity and Copyright indicate the number of posts moderated under the respective category for the respective stock on a given day. $Ramping\ in\ comment_{i,t}$ is the number of moderated posts regarding stock i on day t that mention ‘ramping’ in the moderator’s comment. $Company\ Info_{i,t}$ is a dummy variable that indicates the release of a company announcement for stock i on day t . All data are for January 2008 through December 2008. All the explanatory variables are log transformed except for $Info_{i,t}$.

We previously introduced the dummy variable ‘Ramping in comment’ which corresponds to moderated events where at least one moderation comment contain the word “ramp” or “unlicensed advice”. The variable is also correlated to the two volatility measures, and volume to a lower extent. For volume we find a positive correlation with Ramping, Insider Trading, Other, Advertising, Flaming and Profanity. Other is a general category that is used for posts moderated because of miscellaneous or mixed reasons (and so can include Ramping, Flaming and Profanity).

Our results are limited by the fact that we only have closing prices and do not currently have access to intra-day prices. This results in an inability to distinguish between different moderated posts, that is, we cannot identify the impact of a specific ramping post on the market. Indeed, a ramping posting cannot be measured independently of other moderated posts (such as profanity), which occur on the same day. Hence, we cannot accurately identify whether a specific ramping post has been effective in its cause.

Table 4: Determinants of Target Firms

| | (1) | (2) | (3) | (4) | (5) |
|-------------|-----------------------|----------------------|------------------------|------------------------|-----------------------|
| Intercept | -0.0992 * (-2.489) | 0.1239 ** (2.980) | 0.0704 *** (11.450) | 0.0058 *** (0.453) | -0.0906 (-1.341) |
| Volume | 0.0123 *** (4.025) | | | | 0.0186 *** (3.545) |
| Market cap | | -0.0036 (-1.506) | | | -0.0059 (-1.243) |
| Price | | | -0.0188 ** (-2.856) | | -0.0157 (-1.261) |
| Volatility | | | | 1.0479 *** (-2.306) | 0.4135 (1.251) |
| \bar{R}^2 | 0.017 | 0.0029 | 0.0086 | 0.02203 | 0.0501 |
| n | 936 | 758 | 935 | 896 | 750 |

This table presents the results of the regression analysis described in Figure 3. The dependent variable is the proportion of ramping posts given by $\frac{Ramp_i}{Mod_i}$ where $Ramp_i$ is the number of ramping posts of stock i during the sampling period and Mod_i is the number of moderated posts of stock i during the sampling period. $Volume_i$ is the average turnover for stock i during the sampling period, $Market\ cap_i$ is the market capitalization of stock i measured at the end of 2008, $Price_i$ is the average closing price for stock i during the sampling period and $Volatility_i$ is the average volatility of stock i during the sampling period.

Determinants of target firms: In this section, we present our findings in investigating the characteristics of firms, which are the most common targets for ramping in forums. Since some of the independent variables are significantly correlated, we implement both simple and multiple regressions in order to minimize the multicollinearity problem. In the simple regressions (regressions (1) through (4)), stocks with higher turnover, lower price level and higher volatility tend to receive a higher proportion of ramping posts. Market capitalization is also negatively correlated to the proportion of ramping posts but with a low significant value. Regression (5), which controls for multicollinearity, confirms that firms with higher turnover, lower price level, lower market capitalization and higher volatility receive a higher proportion of ramping posts. However, the correlation is highly significant only for volume.

6. Conclusions and Directions for Future Work

To the best of our knowledge, this work provides the first quantitative analysis of market manipulation of Internet stock message boards. Employing a unique body of moderated messages, we are able to identify messages that have been moderated as manipulative due to ramping and examined its relationship with market response. We find that manual moderation of internet stock message boards does not effectively protect users against manipulation. This is consistent with our expectation that rampers in forums are difficult to identify even for moderators. The possible reasons could be due to the fact that ramping

can be a long progressive process, by which a ramper inconspicuously disseminates his ideas but attempts to hide the motivation to ramp. Our experiment also shows that ramping is positively and significantly related to market returns, volatility and volume. In addition, our results demonstrate that stocks with higher turnover, lower price level, lower market capitalization and higher volatility are more common targets of ramping.

Our proxy for ramping is not perfect. Future works may include the development of techniques to detect well-concealed smart manipulators amongst posts not yet reviewed by moderators. Computational linguistics techniques could be employed to detect ramping and to refine our proxy for ramping. As discussed in section 5, the number of posts moderated due to ramping is only a lower bound to the actual number of ramping posts. To identify ramping posts that have not been labeled by moderators, we need to examine the content of all unmoderated posts. We plan to build a text classifier to automatically identify posts with inappropriate content. Applying this classifier to unmoderated posts will enable us to discover posts that should have been moderated due to ramping and subsequently to refine our proxy for ramping. The following two sample unmoderated posts shows that an improvement can be made to detect the poster's intention to ramp.

- “I heard a rumor that the plant is still being repaired and may be operating again later today. Does anyone know how to contact the fellow from Atmos that is mentioned in this article?”
- “I’m focusing on oil stocks now, not affected by all this, and also I heard a rumor even General Motors now admit our oil has run out on what we can produce in one day and now in decline, which is good news to the oil stocks.”

Author information: Jean-Yves Delort, the submitting author is a software engineer at Google and was a staff member in the Department of Computing, Macquarie University, Sydney, when this work was conducted. Contact E-mail: jydelort@google.com. Bavani Arunasalam is a staff member at the Capital Markets CRC Limited, Sydney, NSW 2001, Australia. Henry Leung is a staff member at the Faculty of Economics and Business, University of Sydney, NSW 2006, Australia. Maria Milosavljevic is a specialist in advanced analytics working for the Australian Government in Canberra.

References

- Aggarwal, R. K., Wu, G., (2006). Stock market manipulations. *Journal of Business* 79 (4): 1915–1953.
- Allen, F., Litov, L., Mei, J., (2006). Large investors, price manipulation, and limits to arbitrage: An anatomy of market corners. *Review of Finance* 10: 645–693.
- Alsaeed, K., (2006). The association between firm-specific characteristics and disclosure: The case of Saudi Arabia. *Managerial Auditing Journal* 21 (5): 476–496.
- Antweiler, W., Frank, M. Z., (2004). Is all that talk just noise? The information content of internet stock message boards. *The Journal of Finance* 59 (3): 1259–1294.
- Mark Bagnoli, Barton L. Lipman, (1988). Successful Takeovers without Exclusion, *Review of Financial Studies*, Oxford University Press for Society for Financial Studies 1 (1): 89-110.
- 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 (2): 785–818.
- Bohme, R., Holz, T., (2006). The effect of stock spam on financial markets. Brown, G. W., Cliff, M. T., 2004. Investor sentiment and the near-term stock market. *Journal of Empirical Finance* 11: 1–27.
- Chakraborty, A., Ylmaz, B., (2004). Manipulation in market order models. *Journal of Financial Markets* 7: 187–206.
- Cook, D. O., Lu, X., (2009). Noise, information, and rumors: Internet board messages affect stock returns. Working paper.
- Das, S., Sisk, J., (2005). Financial communities. *Journal of Portfolio Management* 31 (4): 112–123.
- Eom, K. S., Lee, E. J., Park, K. S., (2009). Microstructure-Based Manipulation: Strategic Behavior and Performance of Spoofing Traders. SSRN eLibrary.
- Fraser, J., (2007). The mysterious world of stock forums. Compareshares.com.au.
- Gastineau, G. L., Jarrow, R. A., (1991). Large-trader impact and market regulation. *Financial Analysts Journal* 47: 40–72.
- Guo, S., Liu, M. H., Song, W., (2008). Stock splits as a manipulation tool: Evidence from mergers and acquisitions. *Journal of Financial Management* (Winter): 695-712.
- Hanke, M., Hauser, F., February (2008). On the effects of stock spam e-mails. *Journal of Financial Markets* 11 (1): 57–83.
- Harris, L., (2002). *Trading and Exchanges: Market Microstructure for Practitioners*. Oxford University Press, USA.
- Hillion, P., Suominen, M., (2004). The manipulation of closing prices. *Journal of Financial Markets* 7: 351–375.
- Hu, B., McInish, T., Zeng, L., (2009). The can-spam act of 2003 and stock spam emails. *Financial Services Review* 18 (1): 87–104.
- Kumar, P., Seppi, D. J., (1992). Futures manipulation with cash settlement. *The Journal of Finance* 47 (4): 1485–1502.
- Leinweber, D. J., Madhavan, A. N., (2001). Three hundred years of stock market manipulations. *The Journal of Investing* 10 (2): 7–16.
- Merrick, J. J., Naik, N. Y., Yadav, P. K., (2005). Strategic trading behavior and price distortion in a manipulated market: anatomy of a squeeze. *Journal of Financial Economics* 77: 171–218.

- Tetlock, P. C., (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance* 62 (3): 1139–1168.
- Zigrand, J. P., (2006). Endogenous market integration, manipulation and limits to arbitrage. *Journal of Mathematical Economics* 42: 301–314.