The Effect of Social Media on Market Liquidity

Research-in-Progress

Tianyou Hu
University of Auckland
Address
Owen G Glenn Building, 12 Grafton Road, Auckland, New Zealand
t.hu@auckland.ac.nz

Arvind Tripathi
University of Auckland
Address
Owen G Glenn Building, 12 Grafton Road, Auckland, New Zealand
a.tripathi@auckland.ac.nz

Abstract

Recent studies have investigated the effect of investors' opinions, or peer-based advice, available on social media platforms or stock discussion forums on the financial market. The researchers have been focusing on stock earnings surprises, abnormal returns and trading volume. But the jury is still out on when and how the discussion on social media impacts financial markets, especially market liquidity. For example, are social media platforms heterogeneous in the value they offer to investors? Further, we do not know if peer-based advice contains information that will be reflected by the market liquidity. We aim to shed light on these concerns by evaluating whether sentiments from messages on online discussion forums could influence the short sales activities in the Australian market. We employ textual analysis techniques to extract sentiments from these platforms. Practical implications for the markets have been discussed.

Keywords: User-Generated Content, Market Liquidity, Text Mining, Wisdom of Crowds

Introduction

Social media platforms and online discussion forums have become popular places to share and learn about products, services, political or international events and even financial markets. It is interesting to note that users in these online communities and social media platforms are willing to trust the investment advice/opinions coming from unknown fellow community members. The social media platforms are becoming more important in comparison with traditional information sources such as news media because the information or messages on social media platforms are updated in real-time by millions of contributors and these messages spread very quickly, providing first-hand information to investors before other sources. Instead of only focusing on experts' advice, peer-based opinions from social media have started to play a bigger role in financial markets (Chen et al., 2014). However, it remains an open debate how the activities from varied social media platforms could influence the market.

There has been conflicting research outcomes showing quite different results. For instance, Chen et al. (2014) find that opinions revealed on an equity research website strongly predict future stock returns and earnings surprises. On the other hand, Kim and Kim (2014) find no evidence that investor sentiment forecasts further stock returns either at the aggregate or the individual firm level. As such, we aim to elucidate this matter by examining whether messages posted in different online social media platforms and online investor discussion forums have an impact on the stock market. Moreover, we use a dataset of three different extensively used social media platforms containing all companies in the S&P/ASX50 index. This dataset allows us to conduct a comprehensive research across social media channels with different characteristics.

Prior work on investor discussion forums has been focusing on different aspects of the market movement. For example, Das and Chen (2007) show that larger number of messages posting in Yahoo Finance Message Boards will lead to higher volatility and lower stock price. Sabherwal et al. (2011) demonstrate that more positive sentiments on social media will be followed by subsequent higher stock returns. Leung and Ton (2015) argue that larger number of board messages and more positive message sentiments will lead to higher stock returns. Instead of focusing on stock returns, volatility or volumes, we will focus on the stock market short sales, which is a proxy of market liquidity (Diether et al. 2006). A commonly stated benefit of liquidity is that it allows the rapid exit from a stock when the share price falls (Kemp 2014).

We propose that the sentiments disclosed on investor discussion forums will impact the short sales. To understand the sentiments and the views dispersed through investor discussion forum, we first adopt a predefined dictionary from the literature (Loughran and Mcdonald 2011), which suggests that the percentage of negative words in an article affects the tone of the report and eventually affects the market. We will test the robustness of our results with the inclusion of control variables mirroring the overall market movement (S&P/ASX 50 index), the sentiments from public news services (the average percentage of negative words in the Google Finance News and DJNS news articles). A few studies have analyzed the sentiments disclosed on investor discussion forums using machine-learning classifiers, such as Naïve Bayes (NB) and support vector machine (SVM). This stream will be covered in the papers following this one.

We are collecting data from three different extensively used social media platforms (Hot Copper, Facebook, and Twitter) focusing on companies in the Australian Market. This dataset allows us to conduct a comprehensive research across different social media platforms with different characteristics. This study investigates the effect of sentiments disclosed on Hot Copper on market liquidity. To understand the tone from Twitter (TW), there are another two possible ways besides the machine-learning classifiers: Bollen, Mao, & Zeng (2011) use OpinionFinder to measure positive and negative mood as well as utilizing Google-Profile of Mood States (GPOMS) to measure mood in terms of 6 dimensions (e.g. happy) and Davies and Ghahramani (2011) demonstrate a new language-independent model for analysis of short, social-network statuses (posts). For Facebook (FB), and Twitter (TW) we will first use machine-learning classifiers, and later we aim to develop sophisticated textual analysis tools for better performance.

The rest of the paper is structured as follows: We first provide a brief and relevant literature followed by our data collection strategy and details about data and sources. Later we discuss our methodology and report preliminary results, discussion and plan for future research.

Literature Review

Social media platforms provide valuable information to users and have been dubbed as the sources of collective wisdom. In a message board-related study, Antweiler and Frank (2004) used more than 1.5 million messages posted on Yahoo! Finance and Raging Bull, including 45 companies to demonstrate that stock messages have predicting power over market returns, trading volume, and volatility. They computed a “Bullishness index” based on computational linguistics method, which is extensively used in the literature. Furthermore, A recent study focusing on a discussion forum collected more than 2.5 million messages from HotCopper (www.hotcopper.com) related to over 2000 companies and showed that “the number of board messages and message sentiments significantly and positively relate to the contemporaneous return of underperforming small capitalization stocks with high market growth potential; posting activity is positively associated with trading volume for small stocks and negatively associated with bid-ask spreads for small and large stocks in the short term” (Leung and Ton 2015). In another study about the stock message board, Das and Chen (2007) collected data from Yahoo’s message board and reported that tech-sector postings were related to stock index values, volumes, and volatility. Sabherwal et al. (2011) showed the possibility of using online investor discussion forums to manipulate the small-cap stocks and demonstrated that online sentiments could significantly predict the small-cap stock’s trading volume in the following day after the publication of the posts. These studies argue that the market information affects the message board not vice versa. In contrast to these studies on the stock message board, Tumarkin and Whitelaw (2001) argued the stock price performance affects the subsequent investor sentiments. Kim and Kim (2014) have reinforced this finding by reporting the absence of any link between message board activity and stock returns, and volume. Although researchers have investigated the effect of investor discussion forums on the stock market in many ways such as stock price, abnormal returns, etc., no published work has dealt with the effect of the message board on the financial market liquidity.

Besides stock message boards, many other social media platforms have been proven to have an effect on the subsequent market movement. Equity research platforms pay more attention to expert reviews. Chen et al. (2014) collected 97,070 single ticker articles from Seeking Alpha (an equity research platform). They found that the views revealed from Seeking Alpha articles are confirmed by the subsequent stock returns and earnings surprises. Nevertheless, the focus is still not on market liquidity. The topics from stock message boards and equity research platforms mainly focus on the financial market, and the users on these two platforms are mainly investors and industry experts. However, there are other social media platforms, which don't focus on financial markets, but have been proven to influence the company's equity value. Review websites provide platforms on which users could share their experience of the products from Australian Stock Exchange (ASX) listed companies. The users on these platforms are customers of the companies, who do not have to be shareholders. Tirunillai and Tellis (2012) collected 347,628 reviews from Amazon.com, Epinions.com, and Yahoo! Shopping. They demonstrate that volume of chatter has a strong positive effect on the abnormal return and trading volume. Interestingly, Negative User-Generated Content (UGC) also has a positive effect on trading volume.

For Facebook, Chung et al. (2014) argue that there exists a significant relationship between firm's social media effort in Facebook and the firm's market performance, which is captured by abnormal return. In the following research, instead of focusing on firm's effort on Facebook, we will get the tone of users' messages and examine the effect of the tone of users' messages on the market liquidity. Twitter is the famous micro-blogging service website, which differs from Facebook and other online social media in many ways. Bollen, Mao, and Zeng (2011) collected 9,853,498 tweets and found that mood states derived from Twitter are correlated to the value of Dow Jones Industrial Average (DJIA) over time. In the future research, we will examine the relationship between the sentiments in Twitter and the stock market liquidity.

Textual analysis, especially sentiment analysis, is a domain dependent problem. An expression that has a clear sentiment in one domain may be ambiguous in other domains. This issue is particularly strong in the financial context analysis, as there are specialized concepts and limited use of effective words. Most textual analysis papers show vaguely how a document is parsed and often use a software package in which the driving forces behind the results are not clear. It will be challenging to replicate most of existing textual analysis in various studies. The nature of the English language, where the context of the sentence helps define the meaning of a specific word, also adds to the imprecision of parsing mechanics. To understand the UGC contributed by the social media users, there are two streams of studies.

The first stream is the use of predefined dictionaries to analyze document tone. There are four dictionaries that are widely used in the literature: Henry (Henry 2008), Harvard's General Inquirer (or Harvard's Psychological Dictionary), Diction, and the L&M (Loughran and Mcdonald 2011). The first three dictionaries all have their short comings. For instance, the Henry word list only has limited coverage of words; Harvard's Dictionary is widely used in psychology and sociology fields, and three-fourth of the words identified as negative are not considered negative in the financial contexts (Loughran and Mcdonald 2011); Diction Dictionary has 35 subcategories and has to merge three subcategories to create the positive word list. The L&M lists have become predominant in more recent studies (Kearney and Liu 2014). In line with Loughran and McDonald (2014), this word list has two main advantage over the other dictionaries, which are widely used in the accounting and finance literature. First, unlike the Henry (2008), it is a relatively comprehensive word list. Roughly, no widely used negative or positive words are missing. Second, the Loughran and McDonald (L&M) word lists were built with a financial incentive in mind.

Specifically, the L&M negative and positive word lists have been used to measure the sentiment of the financial documents. Feldman et al. (2009) implemented the L&M word lists to analyze the effect of sentiment from company's annual report on the market and found that positive tone will lead to higher stock market returns. Dougal et al. (2012) examined the Wall Street Journal's (WSJ) columns and found that more pessimistic column tone are linked with more negative market returns in the following day (Tetlock, 2007). Garcia (2013) use the L&M dictionary to demonstrate that the tone of two financial columns in the New York Times are proved to have an influence on the stock returns. Solomon et al.

(2014) use the L&M dictionary to argue that more positive report of the fund will be followed by more cash flow into the mutual fund quarterly. Chen et al. (2014) used the negative word list of the L&M dictionary to show that the sentiments contained in SA(SeekingAlpha.com) articles and SA commentary shed light on future stock abnormal returns and even subsequent earning surprises. In this research, we will also utilize the L&M dictionary to get the percentage of negative words, which is a proxy of the sentiments from social media platforms.

Besides the wide usage of predefined dictionaries, machine-learning algorithms are also becoming popular for analyzing financial articles or comments. The most extensively used machine-learning algorithms to analyze the financial data are Naïve Bayes (NB), and Support Vector Machine (SVM). Machine-learning algorithms are trained on the training set. We could then apply the "knowledged" algorithms to the remaining sample dataset or out-of-sample dataset and get a sentiment for each of the messages. The Naïve Bayes algorithm, one of the oldest, is called "naive" because it assumes the words are independent of each other, even though it is quite unlikely. It has been used to classify messages (Leung and Ton 2015) as bullish, neutral or bearish. SVM is another extensively used classifier in the financial context. Malo et al. (2013) showed that substantial performance could be improved by combining different underlying pattern analysis algorithms. Tirunillai and Tellis (2012) compared the accuracy and sensitivity of NB and SVM. The overall results show that SVM performs better than NB in the products review context. Zhang et al. (2012) conducted a comprehensive research on eight widely applied text classifiers to stock message board data. They found that NB performs better than SVM in the out-of-sample test. A recent study (Hu and Tripathi, 2015) has compared the performance of L&M dictionary and machine-learning classifiers (NB and SVM) and confirmed the superior performance of SVM in a financial context. Future research will focus on how to combine predefined dictionaries and machine-learning classifiers to get better performance for textual analysis for financial data.

This research examines the effect of user-generated content (UGC) in an investor discussion forum on the stock market short sales. Knowing that the short sale is a proxy of market liquidity (Diether et al. 2006), we propose that investor discussion forum has an effect on market liquidity. Short sales account for a big proportion of the whole stock market. Diether et al. (2006) show that short sales account for 24% of NYSE and 31% of Nasdaq share volume. Thus, it's worthy to study the short sales and test the effect of social media on short sales. People join short selling mainly to either speculate or hedge. Speculators are looking for fluctuations in the market, and could increase the trading volume and add market liquidity. Hedging insures your stock position against risk. Short selling is beneficial to the market for the following reasons: 1. Adding liquidity to share transactions. The spread between bid and ask could be reduced with additional buying and selling. 2. Short selling will lead to lower cost to execute a trade, and overpriced securities are driven down. 3. The overall efficiency of the markets is increased with quickening price adjustment. 4. Detecting financial fraud. Besides these benefits, (Diether et al. 2006) argue that short sellers both trade on short-term deviations of price from fundamentals and trade as voluntary liquidity providers. Thus, short selling could be proxy for market liquidity. Short selling is also important in predicting subsequent abnormal return. Diether et al. (2006) show that short sellers correctly predict future negative abnormal return following positive returns. Significant positive returns are generated using a trading strategy based on daily short-selling activity. Aitken et al. (1998) focus on a market where short trades are transparent shortly after the time of execution. They find a significantly negative abnormal return time following short sales. In addition, short sale constraints have an effect on the stock price efficiency. Saffi and Sigurdsson (2011) argue that stocks with higher short sale constraints (lending supply) have lower price efficiency. To sum up, a short sale is a crucial part of the stock market, which will help investors monitor market liquidity, subsequent abnormal return, and even stock price efficiency.

Data

To answer our research questions, we have collected data from HotCopper (HC) discussion forum, Google Finance (GF), market index data from Yahoo! Finance, and market short sale volume from ASXonline. We have developed two agents (web crawlers) to collect data from HC and GF. The remaining data was downloaded from university library database or the website directly. We have finished the first iteration of data collection spanning from January 2014 to March 2015. We have collected data for 46 stocks included in the S&P/ASX 50 index. Four firms underwent identity changes during the sample period and have been removed from the dataset (Tirunillai and Tellis, 2012).

To examine the sentiment of the UGC from social media, we downloaded and stored the messages from HotCopper (http://hotcopper.com.au/) in a database format. HotCopper is the biggest stock discussion forum in Australia. It has 250,000 register members and more than 200,000 unique visitors every month and a huge 21 million monthly page views. On HotCopper, users share their opinions about different stocks. Downloaded messages are dated and timed to the minute with a unique message id and author name as well as respective company stock ticker codes. Users are required to disclose their position on a stock when they post a comment. Therefore, we are able to get authors' profile, and their disclosure (Not Held or Held) of stock holdings. Authors could also disclose the sentiments of their post (Sell, Hold, and Buy) for the discussed stock. The data set is too large for manual interpretation. We employ the L&M dictionary to analyze the sentiments of the message posts.

Message board activity and market response in the news media could be endogenous. Sabherwal et al. (2011) argue that news media coverage has an influence on message board activity. As a consequence, we have controlled the sentiment from news media. News articles from Google Finance are downloaded and kept in a database. Downloaded articles are also dated and timed with a company ticker. The subject, author and content for each article are saved accordingly with the percentage of negative words calculated using the L&M dictionary.

The dataset for HC was cleaned before classification. First, we shifted the message to the most recent trading date if the message was posted on a non-trading date. Second, we aggregated the percentage of negative words for the same stock on the same date and calculated the average accordingly. Overall, our dataset contains 36088 posts and 46 stocks listed on the ASX during the period from January 2014 to March 2015. The dataset for Google Finance was also cleaned and we have a total of 65923 news articles for the same 46 stocks in the same period. Short sales data was downloaded from ASXOnline (https://www.asxonline.com).

We have also developed two agents to collect data from Facebook and Twitter using their APIs. From Facebook, We can collect all the direct messages, posts, comments, likes, and users' profile. From Twitter, we can collect tweets, followers and following of each user, and users' profiles. The data from Facebook and Twitter is not included in our analysis in this research-in-progress (RIP) paper due to page limits. In this study, we focus on the effect of message posts on a discussion forum (Hot Copper) on short sales.

Methodology

We demonstrate our main analysis by the following general model:

$$ShortCap_{i,t+2} = \alpha + \beta_1 NegHC_{i,t} + \beta_2 NegNews_{i,t} + \beta_3 \log(StockIndex_{i,t}) + X\delta + \varepsilon \qquad \text{- (1)}$$

$$ShortCap_{i,t} = Short\ Sales_{i,t} \,/\, Issued\ Capital_{i,t} \qquad \text{-(2)}$$

The dependent variable $ShortCap_{i,t}$ is the percentage of issued capital reported as short sold, where i gives the index for the stocks and t represent the day on which the post appears on the HotCopper or the following trading day if the message is posted on a non-trading day.

Our main independent variables are $NegHC_{i,t}$, and $NegNews_{i,t}$. $NegHC_{i,t}$ is the average fraction of negative words across all messages posted about stock i on day t, and $NegNews_{i,t}$, is the average fraction of negative words across all news posted about stock i on day t in Google Finance. The $NegNews_{i,t}$ will work as a control variable. If there is no message or news on any day for a stock, both $NegHC_{i,t}$, and $NegNews_{i,t}$ are treated as neutral on that day. $StockIndex_{i,t}$ is the daily ASX 50 index value download from Yahoo Finance. The dependent variable is already scaled with issued capital. We find that $Issued\ Capital_{i,t}$ does no show significant relationship with $ShortCap_{i,t}$. Thus, it is not included in the regression model (1). δ is other possible control variables.

In the extant literature, abnormal return is calculated as the difference between raw return minus return on a value-weighted portfolio of firms with similar book-to-market ratio and past returns using the Fama and French (1993) three-factor (plus the Carhart (1997) momentum factor) model. The Fama-French model has been proved to be a good proxy for stock market performance while considering market size, firm size, and other relevant factors (Tirunillai and Tellis, 2012; Luo, Zhang, and Duan, 2013). We then follow the Fama-French model with Carhart momentum:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{i,MKT}(\mathrm{R}_{MKT,t} - R_{f,t}) + \beta_{i,SMB}SMB_t + \beta_{i,\mathrm{HML}}HML_t + \beta_{i,\mathrm{MOM}}MOM_t + e_{i,t} \quad \text{-(3)}$$

$$e \sim N(0,\sigma^2)$$

where $R_{i,t}$ is the return for a stock (firm) *i* on day *t*; $R_{f,t}$ is the risk-free rate; $R_{MKT,t}$ is the average market return; SMB_t is the small-minus-big capitalization factor; HML_t is the high-minus-low book-to-market equity factor; and MOM_t is the momentum factor in the given period. Fama and French attempted to better measure market returns and, through research, found that value stocks outperform growth stocks; similarly, small-cap stocks tend to outperform large-cap stocks. SMB_t and HML_t measure the historic excess returns of small caps over big caps and of value stocks over growth stocks. Recent studies have examined the effect of user generated content on social media on stock earning surprises and abnormal returns (Chen et al. 2014). In this study, we aim to explore the predictive power of UGC on social media on short sales. We know that short sale tells us the liquidity provision, which abnormal return fails to tell us. Diether et al. (2006) show that the abnormal returns are monotonically decreasing with the increase of short sales. Instead of using regression models, they divide the stocks into five groups with different short sales level and compare the abnormal returns accordingly. Future work will investigate the relationship between short sales and abnormal returns using the following model:

$$AR_{i,t+2} = \alpha + \beta_1 ShortSales_{i,t} + \varepsilon \quad \text{-(4)}$$

The dependent variable $AR_{i,t+2}$ is the abnormal return for stock *i* on day *t+2*. $ShortSale_{i,t}$ is the gross short sale for stock *i* on day *t*.

Direct messages or replies on Facebook and Tweets on Twitter are unstructured and short text messages. The L&M dictionary has been proven to perform well for long financial reports. Therefore, for short messages on Twitter and Facebook, we intend to use machine-learning techniques such as NB and SVM to abstract the sentiments. We will use the messages posted on HC as the training set to train the algorithms, and then we will apply the "knowledged" algorithm on the messages collected from Facebook and Twitter. After we get a sentiment classification for each message, following Bullishness index from Antweiler and Frank (2004) can be used:

$$Bullishness_{i,t} = \frac{M_{i,t}^{BUY} - M_{i,t}^{SELL}}{M_{i,t}} \cdot \ln(1 + M_{i,t}) \quad \text{-(5)}$$

where $M_{i,t}^{BUY}$ is the number of messages with sentiment "BUY" for company *i* on day *t*, $M_{i,t}^{SELL}$ is the number of messages with sentiment "SELL" for company *i* on day *t*. The total number of the relevant information is defined as $M_{i,t} = M_{i,t}^{BUY} + M_{i,t}^{SELL}$. Extending our proposed model in equation 1, we can investigate the effect of messages posted on Facebook (FB) and Twitter (TW) as following:

$$ShortCap_{i,t+2} = \alpha + \beta_1 BullishnessFB_{i,t} + \beta_2 BullishnessNews_{i,t} + \beta_3 \log(StockIndex_{i,t}) + X\delta + \varepsilon \quad \text{-(6)}$$

$$ShortCap_{i,t+2} = \alpha + \beta_1 BullishnessTW_{i,t} + \beta_2 BullishnessNews_{i,t} + \beta_3 \log(StockIndex_{i,t}) + X\delta + \varepsilon \quad \text{-(7)}$$

$BullishnessFB_{i,t}$ is Bullishness index for FB on day *t* for company *i*. $BullishnessTW_{i,t}$ is Bullishness index for TW on day *t* for company *i*. Percentage of negative words in news media is replaced with $BullishnessNews_{i,t}$ to make the independent variables comparable. $StockIndex_{i,t}$ is the daily ASX 50 index value. δ is other possible control variables. A similar approach can be applied to investigate the effect of messages posted on an investor discussion forum (Hot Copper) as shown in the equation below.

$$ShortCap_{i,t+2} = \alpha + \beta_1 BullishnessHC_{i,t} + \beta_2 BullishnessNews_{i,t} + \beta_3 \log(StockIndex_{i,t}) + X\delta + \varepsilon \quad \text{-(8)}$$

Results and Discussion

The desire to understand whether the sentiments in online discussion board postings contain value-relevant information is thriving with a growing interest in online social media platforms. To shed some light on this issue, we apply a panel regression to test the two-day predictive power of user generated content on an investor discussion forum (Hot Copper). The two-day regression results are reported in Table 1. The two-day time window to predict the effect of social media on financial markets is widely used in the literature (Chen et al., 2014; Leung and Ton, 2015). Table 1 shows the results of our model proposed in equation 1.

Table 1. The Effect of Messages on Hot Copper on ShortCap (Equation 1)

ShortCap	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
NegHC	.0015346	.0004955	3.10	0.002	.0005634	.0025058
NegGF	.0007366	.0002477	2.97	0.003	.000251	.0012222
Log(StockIndex)	.0006652	.0001949	3.41	0.001	.0002832	.0010472
_cons	-.0048898	.0016834	-2.90	0.004	-.0081895	-.0015902

For the two-day regression, we find that the percentage of negative words in HC and GF is positively and significantly related to the subsequent ShortCap (the percentage of issued capital reported as short sold). This finding means that more negative sentiment in the investor discussion forum will contribute to more gross short sale (scaled by issued capital), which is consistent with the phenomenon that investors sell short when they expect a bear market. The average ShortCap is 8.63E-4 across all the stock in S&P/ASX 50. One unit increase in the percentage of negative words, namely 1% increase in NegHC will contribute to a 1.53E-5 increase in ShortCap. This means, on average, ShortCap is increased by 1.8%. One unit increase in the percentage of negative words in Google Finance will increase the ShortCap by 0.74E-5, which is much smaller than the contribution of NegHC. This result tells us that the sentiments from social media have a bigger effect on the subsequent short selling activity than the sentiments from news media. Aitken et al. (1998) argue short sales will lead to a significantly negative abnormal return in the same day, and short sales are almost instantaneous bad news. Our result verifies this finding in a different way by showing that short sales are the confirmation of the negative sentiments.

But how could social media influence the market short sale volume? Several factors might play a role. (1) People conduct short sales when there is a downward pressure on stock price or liquidity pressure. For either case, the price will drop after some time if the market confirms the investors' opinion. In this case, negative sentiments on the social media platforms and investor discussion forums are confirmed by the market (short sale volume goes up). Note that the short sale trend in the market will reinforce the negative sentiments accordingly (Diether et al. 2006). (2) Social media platforms are unique as they allow users to interact with each other and to give publicly visible feedback on the users' view on a stock. As we assume that there is collective wisdom generated from the crowd, this characteristic of social media enables users to correct bad or wrong posts. Thus, a user with a good posts/comments will get a positive feedback and win additional investors. (3) If the some users trade based on opinions on HotCopper, this will have a price impact on the market and converge the market price to what the other users in HotCopper perceive. In this way, the market will follow the sentiments of the HotCopper. (4) Message authors also want to maintain the quality of their posts as they will receive utility from attention and recognition. If they become an opinion leader, they could attract more followers (Chen et al. 2014).

Conclusion and Implications

This research was intended to investigate the predictive power of social media and the relationship between social media and firm's daily short sales. The results demonstrate that the sentiments disclosed on investor discussion forums is significantly related to the subsequent short sales.

This research contributes to the literature in Information Systems, Marketing, and Finance disciplines in many ways. First, we show that the sentiments disclosed on investor discussion forums have an effect on the market liquidity. More negative sentiments will lead to more capital being short sold in a two-day time window. Our results are robust because we control for market index and information provided by the news media. Second, we show the efficacy of L&M dictionary for analyzing messages on investor discussion forums. We are the first to use the L&M dictionary to analyze the messages from investor discussion forums. Third, we propose to measure and compare the effect of different social media platforms on market liquidity. Social media platforms differ in their functionality, users and market reach. Social media platforms also differ from other market channels as they enable direct interaction among users. These interactions, lead to "collective wisdom", which is more valuable than the information provided by traditional or online news media. This research also informs investment managers. Social media allows investors to monitor liquidity and adjust the positions accordingly. Our models propose a way to examine the effect of sentiments on discussion forums and other social media platforms on market liquidity. Future research aims to investigate the relationship between social media platforms and various market investment options.

Reference

Aitken, M. J., Frino, A., McCorry, M. S., and Swan, P. L. 1998. "Short Sales Are Almost Instantaneously Bad News: Evidence from the Australian Stock Exchange," *The Journal of Finance* (53:6), pp. 2205–2223 (doi: 10.1111/0022-1082.00088).

Antweiler, W., and Frank, M. Z. 2004. "Is all that talk just noise? The information content of Internet stock message boards," *Journal of Finance* (59:3), pp. 1259–1294 (doi: 10.1111/j.1540-6261.2004.00662.x).

Bollen, J., Mao, H., and Zeng, X. 2011. "Twitter mood predicts the stock market," *Journal of Computational Science* (2:1), pp. 1–8 (doi: 10.1016/j.jocs.2010.12.007).

Carhart, M. M. 1997. "On Persistence in Mutual Fund Performance," *The Journal of Finance* (52:1), pp. 57–82 (doi: 10.1111/j.1540-6261.1997.tb03808.x).

Chen, H., De, P., Hu, Y. Y. (Jeffrey), and Hwang, B.-H. B.-H. 2014. "Wisdom of Crowds: The Value of Stock Opinions Transmitted Through Social Media," *Rev. Financ. Stud.* (27), p. hhu001– (doi: 10.1093/rfs/hhu001).

Chung, S., Animesh, A., Han, K., and Pinsonneault, A. 2014. "Firm's Social Media Efforts, Consumer Behavior, and Firm Performance," *ICIS 2014 Proceedings* (available at http://aisel.aisnet.org/icis2014/proceedings/SocialMedia/14).

Das, S. R., and Chen, M. Y. 2007. "Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web," *Management Science* (53:9), INFORMS, pp. 1375–1388 (available at http://pubsonline.informs.org/doi/abs/10.1287/mnsc.1070.0704).

Davies, A., and Ghahramani, Z. 2011. "Language-independent Bayesian sentiment mining of Twitter," (available at http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.378.634).

Diether, K. B., Lee, K.-H., and Werner, I. M. 2006. "Short-Sale Strategies and Return Predictability," *Review of Financial Studies* (22:2), pp. 575–607 (doi: 10.1093/rfs/hhn047).

Dougal, C., Engelberg, J., Garcia, D., and Parsons, C. A. 2012. "Journalists and the Stock Market," *Review of Financial Studies* (25:3), pp. 639–679 (doi: 10.1093/rfs/hhr133).

Fama, E. F., and French, K. R. 1993. "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics* (33:1), pp. 3–56 (doi: 10.1016/0304-405X(93)90023-5).

Feldman, R., Govindaraj, S., Livnat, J., and Segal, B. 2009. "Management's tone change, post earnings announcement drift and accruals," *Review of Accounting Studies* (15:4), pp. 915–953 (doi: 10.1007/s11142-009-9111-x).

Garcia, D. 2013. "Sentiment during Recessions," *The Journal of Finance* (68:3), pp. 1267–1300 (doi: 10.1111/jofi.12027).

Henry, E. 2008. "Are Investors Influenced By How Earnings Press Releases Are Written?," *Journal of Business Communication* (45:4), pp. 363–407 (doi: 10.1177/0021943608319388).

Hu, Tianyou, and Tripathi, Arvind 2015. "The Performance Evaluation of Textual Analysis Tools in Financial Markets," Working paper, University of Auckland New Zealand.

Kearney, C., and Liu, S. 2014. "Textual sentiment in finance: A survey of methods and models," *International Review of Financial Analysis* (33), pp. 171–185 (doi: 10.1016/j.irfa.2014.02.006).

Kemp, M. 2014. "Why share liquidity is so important - ASX," *ASX* (available at http://www.asx.com.au/education/investor-update-newsletter/201405-why-share-liquidity-is-so-important.htm; retrieved September 8, 2015).

Kim, S.-H., and Kim, D. 2014. "Investor sentiment from internet message postings and the predictability of stock returns," *Journal of Economic Behavior & Organization* (107), Elsevier B.V., pp. 708–729 (doi: 10.1016/j.jebo.2014.04.015).

Leung, H., and Ton, T. 2015. "The impact of internet stock message boards on cross-sectional returns of small-capitalization stocks," *Journal of Banking & Finance* (55:July 1998), Elsevier B.V., pp. 37–55 (doi: 10.1016/j.jbankfin.2015.01.009).

Loughran, T., and Mcdonald, B. 2011. "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10 - Ks," *The Journal of Finance* (66:1), Blackwell Publishing Inc, pp. 35–65 (doi: 10.1111/j.1540-6261.2010.01625.x).

Loughran, T., and McDonald, B. 2015. "Textual Analysis in Finance and Accounting: A Survey," *SSRN Electronic Journal* (doi: 10.2139/ssrn.2504147).

Luo, X., Zhang, J., and Duan, W. 2013. "Social Media and Firm Equity Value," *Information Systems Research* (24:1), INFORMS, pp. 146–163 (doi: 10.1287/isre.1120.0462).

Malo, P., Sinha, A., Takala, P., Ahlgren, O., and Lappalainen, I. 2013. "Learning the roles of directional expressions and domain concepts in financial news analysis," *Proceedings - IEEE 13th International Conference on Data Mining Workshops, ICDMW 2013*, pp. 945–954 (doi: 10.1109/ICDMW.2013.36).

Sabherwal, S., Sarkar, S. K., and Zhang, Y. 2011. "Do Internet Stock Message Boards Influence Trading? Evidence from Heavily Discussed Stocks with No Fundamental News," *Journal of Business Finance & Accounting* (38:9-10), pp. 1209–1237 (doi: 10.1111/j.1468-5957.2011.02258.x).

Saffi, P. A. C., and Sigurdsson, K. 2011. "Price Efficiency and Short Selling," *Review of Financial Studies* (24:3), pp. 821–852 (doi: 10.1093/rfs/hhq124).

Solomon, D. H., Soltes, E., and Sosyura, D. 2014. “Winners in the spotlight: Media coverage of fund holdings as a driver of flows,” *Journal of Financial Economics* (113:1), pp. 53–72 (doi: 10.1016/j.jfineco.2014.02.009).

Tetlock, P. C. 2007. “Giving content to investor sentiment: The role of media in the stock market,” *Journal of Finance* (62:3), pp. 1139–1168 (doi: 10.1111/j.1540-6261.2007.01232.x).

Tirunillai, S., and Tellis, G. J. 2012. “Does Chatter Really Matter? Dynamics of User-Generated Content and Stock Performance,” *Marketing Science* (31:November 2014), pp. 198–215 (doi: 10.1287/mksc.1110.0682).

Tumarkin, R., and Whitelaw, R. F. 2001. “News or Noise? Internet Postings and Stock Prices,” *Financial Analysts Journal* (57:3), Association for Investment Management and Research, pp. 41–51 (doi: 10.2469/faj.v57.n3.2449).

Zhang, Y., Swanson, P. E., and Prombutr, W. 2012. “MEASURING EFFECTS ON STOCK RETURNS OF SENTIMENT INDEXES CREATED FROM STOCK MESSAGE BOARDS,” *Journal of Financial Research* (35:1), pp. 79–114 (doi: 10.1111/j.1475-6803.2011.01310.x).