Twitter, Investor Sentiment and Capital Markets: What Do We Know?

Heba Ali[1]

[1] Faculty of Management Technology, German University in Cairo, Cairo, Egypt

Correspondence: Heba Ali, Faculty of Management Technology, German University in Cairo, New Cairo City 11835, Cairo, Egypt. Tel: 20-2-2758-9990-8. E-mail: Heba.abbas-ali@guc.edu.eg

Abstract

Nowadays, the social media play a central role not only in "de-asymmetrizing" the information between firms and investors but also in influencing the emotional response to this information. The social media have provided firms with the opportunity to construct their image and stimulate significant attention and positive emotional responses (i.e. celebrity firm). Investors also become no longer passive participants; they can now communicate, re-tweet, comment, mention, react to information and express their sentiment/views. Theoretically, this should exert a positive impact on information diffusion and so the market efficiency. However, as the social media also significantly influence the public mood and emotional response to any new information/news, several behavioral explanations contradicting with the concept of market efficiency (e.g. investor sentiment and herding behavior) become more reasonable. The study aims at providing a literature review and synthesis of research on the impact of social media sentiment in the context of capital markets, scrutinizing the theoretical understanding of this impact, underlining the methodological challenges related to extracting the sentiment, and reviewing the main empirical findings on the impact in the context of Twitter and StockTwits, which will enable researchers to evaluate and classify existing studies, obtain useful insight into the theoretical understanding of the impact of social media sentiment, hence spurring further theoretical and empirical research.

Keywords: social media, Twitter and StockTwits, investor sentiment, market efficiency, data analysis

1. Introduction

Nowadays, the social media play a central role not only in "de-asymmetrizing" information between firms and investors (i.e. reducing the information asymmetry between firms and investors) but also in influencing – and sometimes leading – the emotional response to this information (Note 1). Via being a relatively objective and authoritative source of information, the social media have provided firms with the opportunity to construct their image and stimulate significant attention and positive emotional responses (i.e. celebrity firm), and the greater the extent the firm can exploit this opportunity, the more valuable this "intangible asset" is to the firm (e.g. Deephouse, 2000; Pollock & Rindova, 2003; Rindova, Pollock, & Hayward, 2006). Investors also become no longer passive participants; they can now communicate, re-tweet, comment, mention, react to information and express their sentiment/views, and the greater the extent to which a firm is subject to public attention and media coverage, the more the firm will be affected by investors' emotions, sentiment and behaviors.

Theoretically, this should exert a positive impact on information diffusion and so the market efficiency (i.e. information is expected to spread more rapidly and so it will not take that long to be instantly and fully reflected into prices, which makes any arbitrage opportunities not feasible). However, as the social media also significantly influence the public mood and emotional response to any new information/news, several behavioral explanations contradicting with the concept of market efficiency (e.g. investor sentiment and herding behavior) become more reasonable. Methodologically, these behavioral factors are empirically unobservable and so alternative measures are usually proposed to serve as proxies for these factors, such as Baker and Wurgler's (2006) investor sentiment index, but fortunately a closer look at these factors can be taken now after the social media have created an opportunity to access more detailed data on the investors' perception, interactions, sentiment and behavior collectively at the society level (Bukoniva, 2016). This is why a fast growing research has been triggered to scrutinize the impact of information/sentiment extracted from social media platforms in different contexts and its implications for the notion of behavioral finance versus market efficiency and investor rationality.

Currently, there are several micro-blogging platforms that allow users (e.g. firms and investors) to disclose, share, comment and react to information, such as Facebook and Seeking Alpha. Yet, Twitter - since its conception in 2006- represents one of the most popular micro-blogging platforms, being currently the 13th most popular website globally (Baik., Choi, & Kim, 2016; Alexa, 2018). This is due to its different unique features that created an interactive channel for fast and easy dissemination of information. For instance, Twitter's messages are short and concise (even after the word limit was increased in 2017 from 140 firm up to 280 characters). Users can also use the dollar symbol ($) before a firm's stock ticker in its tweets/search to indicate that those tweets contain financial information. Moreover, to facilitate more information diffusion in the context of the financial market, StockTwits platform was launched in 2008 to run on top of the Twitter platform and be specifically dedicated to the financial domain. In StockTwits, users can also classify their own tweets and label their sentiment as bullish or bearish, which can be utilized to automatically obtain an overall sentiment indicator.

Empirically, the majority of the studies that examined the impact of Twitter and StockTwits in the context of finance were mainly focusing on examining the predictive power of extracted sentiment/information in forecasting the stock market performance (e.g. returns, trading volume and volatility). Bollen, Mao, and Zeng (2011) were the first to utilize Twitter sentiment to predict stock market performance, exhibiting supportive empirical evidence. Following them, numerous studies have examined the usefulness of Twitter sentiment in different contexts, such as IPOs returns, earnings announcements and commodity markets. Overall, the findings stand in consistency with the predication power of Twitter sentiment in these several perspectives. Researchers- in these studies- typically rely on two main approaches. One approach extracts and analyzes sentiment/emotional content from tweets that are posted by users (e.g. investors, CEOs and firms) and then examines the impact of this sentiment on selected financial variables (e.g. Rao & Srivastava, 2014; Nisar & Yeung, 2018). The second approach considers the firm-specific characteristics on social media platforms as a proxy for the level of firm activity, popularity and influence, such as the number of firm posts/tweets, and then investigates how this information will affect selected financial variables (e.g. Cole, Daigle, & Van Ness, 2015; Lundmark, Oh, & Verhaal., 2016). Only the first approach can provide insights into the sentimental or emotional content of the investors' options/messages, yet it is an uneasy task to extract this sentiment.

The purpose of this study is threefold. The first is to critically review the theoretical explanations that are asserted to interpret the impact of the information, sentiment and emotional content extracted from social media in the context of capital markets. Secondly, the study also underlines the methodological considerations and difficulties when extracting and studying the sentiment/emotional content from social media platforms. Thirdly, the study offers a detailed literature survey and synthesis of the main studies on this impact in the context of Twitter, StockTwit and capital markets. By doing so, the paper will enable researchers to understand this impact and its implications for theory (i.e. market efficiency vs. behavioral finance hypotheses), take into account the possible challenges when extracting sentiment and to evaluate and classify existing studies, which is expected to spur further theoretical and empirical research on the topic.

The paper now proceeds as follows. Section 2 critically reviews the theoretical interpretations for the role of the social media sentiment/information in the context of capital markets. Section 3 summarizes the different approaches of investor sentiment extraction/analysis and discusses the related methodological difficulties. The following section summarizes and compares the main empirical studies. The final section concludes and offers areas of further research.

2. Theoretical Background

In efficient markets, stock prices should fully and instantly incorporate all relevant information as soon as it becomes available and so the movements in stock prices are random and cannot be predicted (Fama, 1970, 1991; Dockery & Kavussanos, 1996; Dimson & Mussavian, 1998; Dragota & Oprea, 2014). Theoretically, the social media should exert a positive impact on the market efficiency as information is expected to spread more rapidly for stocks subject to media coverage/investor attention than for stocks that are relatively less covered by the media, which implies that it will not take that long for information to be fully reflected into the prices, and so any arbitrage opportunities to take advantage of this information will not be feasible as suggested by the market efficiency hypothesis. However, the social media also play a central role in influencing – and sometimes leading – the public mood and emotional response to any new information/news communicated about/by the firm to the market. Here, several behavioral explanations - such as investor sentiment and herding behavior - become more relevant and widely adopted, which explicitly contradicts with the concept of market efficiency.

By relaxing the conventional notion of market efficiency and investor rationality, behavioral finance appears to provide more convincing explanations to several questions remained unanswered for long time in the literature,

such as IPOs anomalies (Ritter & Welch, 2002; Ali, 2017) and timing of mergers and acquisitions (e.g. Shleifer & Vishny, 2003; Bi & Gregory, 2011). By doing so, factors such as investor psychology, herding behavior, peers influence, opportunistic managerial decisions, stock mis-valuation and cognitive bias become important determinants that behavioral finance theory takes into account to explain the behavior of market agents and stock performance (Shiller, 2003; Nofsinger, 2005; Bakera & Wurgler, 2007). Methodologically, these factors are empirically unobservable (i.e. it is not possible to directly measure these forms of irrationality). This is why alternative measures serving as proxies for these factors, such as the investor sentiment indices developed by Baker and Wurgler (2006) and Lee, Shleifer and Thaler (1991) have been proposed and used in the literature.

Fortunately, a closer look at these factors can be taken now after the social media have created an opportunity to access more detailed data on the investors' perception, interactions and behavior collectively at the society level (Bukoniva, 2016). These data are extracted from social media platforms that have enabled the markets participants (e.g. firms, investors and analysts) to disclose, share, comment, correct, react to information and express their views/recommendations about it. In addition, when the markets are not efficient, there will be feasible and persistent arbitrage opportunities due to prevailing market frictions (Mitchell, Pulvino, & Stafford, 2002; Herschberg, 2012). Therefore, in seeking to outperform the market, investors will find it is valuable to collect and process information - that might be still unincorporated in prices – to identify and exploit these opportunities. Such that, examining the role of stock-related information extracted from social media in identifying such opportunities and then testing its predictive power can help to shed another light on the notion of market efficiency (Benthaus & Roman, 2015).

These reasons have triggered a fast growing research scrutinizing the impact of information/sentiment extracted from social media in different contexts related to capital market and its implication for the notion of behavioral finance versus market efficiency and investor rationality. Theoretically, most of the explanations asserted to interpret the link between the social media and stock performance measures (e.g. prices and trading volume) attribute that impact to behavioral factors - such as investor sentiment and emotional contagion in the social media (i.e. explanations based on the interrelationship among sentiment/emotions, information diffusion in a social media and stock performance).

One of the main behavioral factors in the social media literature is investor sentiment. There are enormous studies that analyze the sentimental content (i.e. mood, tone, valence) extracted from social media and then examine its impact on selected financial variables. Conceptually, investor sentiment has several definitions and implications in the literature (Note 2). These implications are mainly related to the deviations in price from fundamental value due to irrational factors including optimism and pessimism as in Baker and Wurgler (2006) and Brown and Cliff (2004), under-reaction and over-reaction as in Barberis, Shleifer, and Vishny (1998) and noise trader as in DeLong, Shleifer, Summers, and Waldmann (1990). When reviewing the research conducted in the context of social media, sentiment is found simply defined as the public mood or emotion that spreads through social interactions. Empirically, the impact of social media sentiment has been widely supported as it will be displayed later.

Related to the concept of investor sentiment is the notion of emotional contagion and herding behavior, yet it differs from investor sentiment in the sense that it needs to be proven to practically provide evidence that some sort of transfer of emotions or herding behavior actually occurred. The idea of emotional contagion refers to a transfer of emotions among individuals and so influencing their mood/behavior. Being easily able to view the posts, rated news, and attached comments after the rise of social media; it will not be surprising to see investors more subject to the effect of emotional contagion. Empirically, several studies – e.g. Coviello et al. (2014); and Kramer, Guillory, and Hancock (2014) – support the contagious spread of emotions among internet users via text-based communication. Nofer and Hinz (2015) attempted to examine this effect in financial setting using Twitter mood and its effect on stock returns. Although the findings exhibit supportive evidence of the influences of emotional contagion investors' behavior, the authors state "we cannot actually prove emotional contagion at this point in time. We can only assume the spread of mood states among Twitter users" (p. 239). In other words, it will be less difficult to implicitly deduce emotional contagion rather than directly measure or prove it.

A very related concept to emotional contagion - and sometime used exchangeably- is herding behavior. The effect of herding - defined as convergence of behavior brought about by direct or indirect social interactions without influence of any central coordination – occurs when the perceptions/behavior of investors are influenced by each other (e.g. Hirshleifer & Teoh, 2009; Michael & Otterbacher, 2014). Empirically, the herding behavior has been widely supported in the setting of capital markets (Note 3). In regard with the social media, Muchnik, Aral and Taylor (2013) and Trinkle, Crossler, and Bélanger (2015) find consistent empirical finding showing that participants' perceptions and reactions to the news are influenced by majority opinions delivered via social

media comments.

In related context, "word of mouth" effect is another perspective of the relationship between information dissemination in social media and stock market reaction. The word-of-moth effect accelerates the diffusion of information across investors that are influenced by each other about specific event/news. Shiller (2000: p.155) assets that "word-of-mouth transmission of ideas appears to be an important contributor to day-to-day or hour-to-hour stock market fluctuations". Empirically, several studies exhibit strong positive co-movement in the trading decisions among investors that are more likely to be in direct contact. For example, Hong, Kubik and Stein (2005) and Ivkovich and Weisbenner (2007) find supportive impact of word-of-mouth communication in the contexts of mutual funds trades in the same city and households' stock purchases in the same neighborhood respectively. However, Lundmark et al. (2016) and Zhang, Ma, and Wang (2017) both find unsupportive empirical evidence of the impact of   the word-of mouth on the IPOs initial returns and stock returns,.

Related to the word-of-mouth effect is the hypothesized impact of geographic proximity on the investor information/sentiment within clustered groups. Local investors - via word of mouth effect - are likely to be influenced by locally perceived values and information (Shiller, 2000). Using the same logic, local users in the same social media network are expected to have information advantage due to better access to information about local firms and stronger social interactions within this clustered network (Takhteyev, Gruzd, & Wellman, 2012; Baik et al., 2016). Giannini, Grant, and Hodge (2013) and Baik et al. (2016) both exhibit consistent findings on the impact of the geographic proximity of Twitter users on stock returns

In addition to the above-mentioned behavioral explanations, the literature has also introduced other interpretations that are consistent with the concept of market efficiency and related to the role of social media in information diffusion and how this can impact the market reaction to any new information/news - without necessarily assuming the market is not efficient - via channels/factors affecting the information asymmetry and quality. The social media play a central role in "de-asymmetrizing" information between the different agents in the market even if the information communicated to the market is not new (Alexander & Gentry, 2014). In addition, social media allow for two-way communication: among individuals (i.e. individuals are now connected and can discuss their views mutually via social media platforms and forums) and between individuals and firms (i.e. individuals can now disclose, share, comment, correct, react to information and express their views/recommendations about it), which allow firms not only to communicate information to investors but also allow investors to communicate their perceptions to the firm (Cade, 2018). Related to this is the idea of social bond that causes individuals to place higher trust in the other individual or group due to being personally interacting with the source of the information via the social media and so feeling interpersonally connected with each other (e.g. Shapiro, Sheppard, & Cheraskin, 1992; Haslam & Ellemers, 2005; Elliott, Grant, & Hodge, 2018). As a result, the market reaction to the firm news is expected to relate to the extent to which this social bond is strong (i.e. less negative reaction following non-pleasing news should be expected when more trustful relationships are developed). Elliott et al. (2018) find that investors develop a stronger social bond and higher trust when a CEO communicates firm news via a personal Twitter account.

3. Methodological Considerations

There is ongoing enormous research on the usefulness of information and sentiment that can be extracted from social media platforms in predicating and explaining a wide range of financial series and indicators, such as stock returns, market reactions to earnings announcement and exchange rates. These studies can be classified- according to the analysis approach used - into two main categories (i) studies that extract and analyze sentiment/emotional content from tweets that are posted by users (e.g. investors, CEOs and firms) and then examine the impact of this sentiment on selected financial variable(s) and (ii) studies that consider the firm-specific characteristics on social media platforms as a proxy for the level of firm activity, popularity and influence, such as the number of firm posts/tweets, the number of followers, number of mentions, and number of re-tweets, and then investigate how this information will affect selected financial variable(s).

In the first approach, to extract sentiment from text, sophisticated textual and sentiment-mining algorithms are usually required to classify the information sentiment conveyed in the posts into different emotional groups (e.g. happiness, anger and pessimism) or into different emotional valence (i.e. positive vs. negative). Generally, there are two main methods used for sentiment extraction from text (Note 4). Firstly, lexicon-based method, where sentiment is identified based on words included in a pre-defined lexicon or keywords, and then computing algorithms (e.g. Python's Pattern) using these lexicons for text mining and then utilizing the natural processing language (NPL) to extract sentiment from this mined text (Metaxas, Mustafaraj, & Gayo-Avello, 2011). These lexicons are mostly provided by independent domain (e.g. SentiWordNet and PsychSignal). However,

lexicon-based methods are usually generic and unsupervised and so may lack effectiveness in computing sentiment in the financial context (Oliveira, Cortez, & Areal, 2016 and 2017). Although there are other financial lexicons developed by Loughran and McDonald (2011) and Oliveira et al. (2016), only limited number of studies used it (Oliveira et al., 2017). Also, these lexicon-based methods might become unable to adapt to the dynamic changes in real-time data of public sentiment (Smailović, 2015). The second method is machine learning (ML), where language model classifiers (e.g. Naive Bayes or Support Vector Machines) are applied to extract sentiment based on supervised learning (i.e. a labeled/ annotated training data set are required to build this classification model).Yet, the manual annotation and cross-validation of this data set that are required to obtain reliable labeled tweets is highly expensive and time-consuming (Ranco, Aleksovski, Caldarelli, Grcar, & Mozetic, 2015; Oliveira et al., 2017).

Consequently, the first analysis approach – in addition to requiring sophisticated textual and sentiment-mining algorithms - has several other drawbacks that make it uneasy task. First, extracting sentiment requires considerable efforts spent in creating filters and special programs that use stock symbol, company name, and other keywords to identify firm-related information (Ruiz, Hristidis, Castillo, Gionis, & Jaimes, 2012). As a result, information collected might suffer from interfering noise and biasness due to ambiguous identifiers and criteria, and so it should be carefully handled; especially some messages/Tweets might be spam or fake (Liu, Wu, Li, & Li, 2015; Wlodarczak, Ally, & Soar, 2015). Second, retrieving such information requires managing a large volume of data, which discourage studies to consider long time spans and/or include large samples, which may impact the generalization of results. However, although the second approach that emphasizes on the firm-characteristics on social media do not suffer from these drawbacks as collecting and analyzing the required data/information in this approach is generally straightforward task, yet it does not give insights into the sentimental or emotional content of the investors' options/messages.

4. Main Empirical Findings

The majority of the studies that examined Twitter and StockTwits in the context of finance are mainly focusing on examining the predictive power of Twitter sentiment in forecasting the stock market performance (e.g. returns, trading volume and volatility) for stocks/indices. Bollen et al. (2011) were the first study to utilize Twitter – based on analyzing 9,853,498 tweets posted by approximately 2.7 million users during 2008 - to predict stock market performance, finding that public mood significantly improved the prediction of Dow Jones average index (DJIA). After that, numerous studies have examined the usefulness of Twitter sentiment in different contexts, such as IPOs returns, earnings announcements and commodity markets (oil and gold) and foreign exchange. Overall, the findings exhibit supportive evidence on the predication power of Twitter sentiments in these several perspectives. A summary of the main empirical findings is presented in appendix A.

4.1 Twitter Sentiment/Information and Stock Market Performance

In the context of stock market performance, Sul, Dennis, and Yuan (2014), based on a sample of Twitter messages posted about the constituent firms of S&P 500 index from March 2011 to February 2012, analyze whether the posts contain an overall positive or negative emotional sentiment and how this sentiment affects the stock returns. To classify tweets according to their emotional sentiment, the authors use the word analysis strategy based on Harvard-IV dictionary. The results show that the positive or negative emotion of tweets about a specific firm is significantly related to the firm's stock returns. In specific, they find that emotional sentiment extracted from users with many followers is associated with same-day abnormal returns, which is consistent with their hypothesis that "private information that is diffused faster will be more quickly incorporated into market prices, and will have higher association with same-day returns and lower future-return predictability"(p. 814).

In another study, Rao and Srivastava (2014) investigate the relationship between sentiment and a number of financial performance indicators including volatility, trading volume and stock prices for more than 4 million tweets between June 2010 and July 2011 posted about the two indices of DJIA, NASDAQ-100 in addition to 11 other big-cap technological stocks. In their sentiment analysis, the authors use sentiment algorithm based on Naive Bayesian classification method to classify tweets into positive or negative, and then total number of positive and negative tweets is aggregated for each day. The authors exhibit results supporting the predication power of Twitter sentiments in explaining the price movements of stocks/indices.

In a comprehensive study on the impact of information extracted from Twitter on stock markets, Benthaus and Beck (2015) develop a model incorporating three dimensions of social broadcasting networks: users (e.g. user popularity), messages (e.g. number of mentions and financial sentiment), and discussion (e.g. volume of exchanged messages). Based on analyzing 1.84 million tweets collected for the S&P 100 companies in 2014, the findings show that the influence of social broadcasting on stock markets is driven by the message and discussion

dimensions whereas the user dimension has no significant influence.

In a similar study, Ranco et al. (2015) examine the relationship between Twitter activity and sentiment and stock returns for a sample of the 30 constituent firms of the Dow Jones Industrial Average (DJIA) index. The sample used consists of 1.5 million tweets collected about the sample firms during the period (2013- 2014). To extract Twitter sentiment, the study uses a labeled set consisting of over 100,000 of tweets - labeled by 10 financial experts with three sentiment labels: negative, neutral or positive to build a Support Vector Machine (SVM) classification model which discriminates between negative, neutral and positive tweets - that was applied to the complete set of over 1.5 million tweets. The obtained findings exhibit an impact of the Twitter sentiment on abnormal returns during the peaks of Twitter volume. Consistently, Cole et al. (2015) study the link among corporate Twitter activity, stock returns and trading activity. The authors find that both the number of daily tweets and the number of months a firm tweets is positively correlated with stock excess returns and trading turnover for a data sample of 215 firms listed in the S&P 500 during the period (2010-2011).

In the study of causal relationship between social media sentiment and stock markets, Souza and Aste (2015) utilize a sophisticated approach based on natural language processing (NPL) supplied by PsychSignal.com to extract sentiment from 1,767,997 Tweets for the Dow Jones Industrial Average (DJIA) constituent firm list over the period (2012-2014). The findings are supportive of the existence of nonlinear causal relationship between Twitter sentiment and stock returns. Risius, Akolk, and Beck (2015) use sentiment lexicon - based on emotion theory – to classify social media information into seven different emotions (affection, happiness, satisfaction, fear, anger, depression, and contempt), and then examine the connection between these differential emotions and stock price movements. Based on the analysis of approximately 5.5 million tweets about selected 33 stocks listed in S&P 100 index, the results show that when comparing positive and negative sentiment, only the average negative emotionality strength has a significant impact on the stock returns. Consistently, Nisar and Yeung (2018) investigate the relationship between politics-related sentiment extracted from Twitter using lexicon-based sentiment extraction method and FTSE 100 movements during a 6-days window including before, during and after the local elections in 2016. The findings exhibit overall supportive evidence of the usefulness of sentiment extracted from Twitter for forecasting market movements.

In the context of stock-return comovement, Liu et al. (2015) examine the impact of social media metrics on stock return comovement for a sample of 293 US listed firms. By identifying official Twitter accounts for these firms, the authors show that firms with official Twitter accounts have much higher comovement than firms without such accounts. The authors also use firm-specific microblogging metrics (e.g. the number of followers) to cluster those firms into homogeneous groups and find that firms in the same clustered groups have higher comovements than those in the same industry, which indicates that social media can produce better results compared to other traditional industry classifications. Similarly, Sprenger and Welpe (2011) propose a new approach in which the degree to which pairs of companies are associated with each other in an online stock forum (i.e. relatedness) is claimed to be related to the comovement of their stocks. To test this claim, the authors make use of a sample of 439,960 tweet messages attached to constituent firms listed in S&P 500 index for a 6-month period between January and June 2010. The findings reveal that the degree to which companies are mentioned jointly can explain the comovement of these stocks returns and help identify a firm's strategic peers.

Rather than considering the sentiment content of social media information, Sun, Lachanski, and Fabozzi (2016) focus on the usefulness of textual information extracted from social media to predict the stock returns. In a different approach, the authors – based on a sample of 45 million messages posted on StockTwits platform during the period (2011-2015) - firstly create a dictionary of terms via examining the top words of highest frequency for each year. Then, these words are combined with a sample 420 stocks listed in listed in the S&P 500 index. When taking in the consideration the information extracted from StockTwits, the findings show that the used model outperforms the other models in predicting the stock market.

Focusing on the technology sector, Corea (2016) examines whether Twitter can proxy the investors' sentiment and so predict stock market movement. Utilizing a sample consisting of 166,000 tweets covering three major technology companies in the US (Apple, Google and Facebook) over a two-month period in 2014, the findings provide supportive evidence that Twitter sentiment is more powerful in predicting stock price trend while the tweets volume has a strong impact on forecasting both price and trend. Consistently, Corea and Cervellati (2015) examine the usefulness of Twitter data about major technologies companies (Apple, Google, and Facebook) to construct a new index-tracking variable to predict the stock market movement. Using an algorithm provided by DataSift to assign a sentiment score to tweets according to its degree of positivity/negativity, the authors construct a sentiment index-tracking (SIT) variable and examine its predication power of the Nasdaq-100 index. The finding are overall supportive as the models augmented with the newly created variable - with respect to the

benchmark- are able to increase the explanatory power and provide a better prediction of the Nasdaq price and volume. Similarly, Dey (2014) examines the association between Twitter sentiment and stock price of Apple Corporation, finding a significant correlation between Twitter sentiment and daily returns.

In a different perspective, Azar and Lo (2016) include a macro-economic variable in their analysis of the relationship between social media sentiment and stock returns. A natural language-processing (NLP) algorithm is used to classify and score the emotional content of all tweets that cite the Federal Reserve between 2007 and 2014. The findings assert that tweets during the Federal Open Market Committee (FOMC) meetings contain information that can be used to predict stock market returns (proxied by Center for Research in Security Prices (CRSP) value-weighted index) and to build benchmark-outperforming portfolios.

In their study, Liew and Budavari (2017) argue that a new company-specific social media factor should be added to the traditional asset pricing models. Analyzing the sentiment tone (i.e. bullish vs. bearish) using daily stock tweets extracted from StockTwits from January 2013 to November 2015 for a sample of 15 stocks listed in the US, the authors provide significant evidence that the daily bullish percentage helps to explain abnormal returns even when taking in consideration the traditional factors used in asset pricing models. In a related study, Cade (2018) examines – using an experimental study - how firm-investor communications on social media (proxied by a Twitter user criticism and whether and how management responds) affect investors' perceptions about the firm. The findings suggest that the criticism has an influence on investors' perceptions and this effect is increasing in the number of times the criticism gets retweeted.

Karagozoglu and Fabozzi (2017) measure a social media–based market volatility extracted from crowds' opinions on individual stocks and test the economic magnitude of this information in predicting stock market volatility. Using PsychSignal's natural language processing (NLP) algorithm and unique dataset consisting of firehose raw tweets from both StockTwits and Twitter over the period (2012- 2016), each message is classified as bullish or bearish. Then, to construct a measure of social media-based market volatility sentiment, a second algorithm analyzes the various dimensions of the social media sentiment identifying time periods of anomalous behavior in the data. The resulting social anomaly score (SAS) is a unique variable used to proxy for volatility. Using the SAS data for the S&P 500 index constituent list, the authors construct various weighted social anomaly scores for the index (iSASs) and examine its prediction power of stock market volatility. The findings exhibit that information contained in the volatility sentiment covered in the social media contains useful information about future stock market volatility.

On the other side, other studies exhibit unsupportive evidence on the impact of Twitter sentiment on the stock returns. For example, Brown (2012), analyzing 1.9 million Twitter messages collected for a sample of 92 firms listed in S&P 500 index over the period from November 2011 to August 2012, compares the sentiment information extracted from Twitter to existing financial market sentiment measures, such as sentiment survey, and examines its predictive power of stock market performance proxied by the closing price of S&P 500 index. In his sentiment analysis, the author uses Naïve Bayesian Classification algorithm to extract sentiment from tweets based on a manually- created raining dataset. The findings reveal that although Twitter sentiment is found to be strongly positively correlated with the existing sentiment survey methods, yet a weak statistical correlation is found between the Twitter Sentiment and the closing price of S&P 500 index. Nofer and Hinz (2015) examine the link between Twitter sentiment and stock returns, using a sample consisting of approximately 100 million tweets released in Germany during the time period (2011-2013). However, the study exhibits an insignificant relationship between aggregate Twitter mood and the stock market. Similarly, Wang et al. (2015) investigate the quality and impact of information content of StockTwits based on analyzing the messages posted by 86,497 users in the USA during the time period (2009-2014). The findings exhibit a low correlation between the information sentiment and aggregate stock performance.

4.2 Twitter Sentiment/Information and IPOs Returns

The usefulness of Twitter sentiment has been also investigated in the context of IPOs. Liew and Wang (2016) were the first to investigate the link between that Twitter sentiments and the IPOs' short-term returns. The study uses a sample consisting of 325 US IPOs launched in the period (2013-2014), utilizing natural language processing (NPL) algorithm provided by iSENTIUM LLC in the sentiment analysis of tweets. The findings exhibit that a direct contemporaneous relationship between IPOs' Twitter sentiment and returns on the first trading day. Also, Twitter sentiment pre-IPO can be used as a predictor of its short-term returns.

In a similar study, Lundmark et al. (2016) test if social media serve as a mechanism for conferring legitimacy in the market for initial public offerings (IPOs) by analyzing the impact Twitter's use had on the IPO performance of 272 US firms during the period (2009-2011). Different from the most-commonly used method of extracting

Twitter sentiment as a proxy for the impact of Twitter, the authors focus on considering other factors related the degree of firm activity and popularity in Twitter, such as Twitter membership and number of followers and retweets (i.e. firm specific matrices). The findings show that firms that have a Twitter account and the extent to which they utilize Twitter are associated with systematically higher levels of IPO first-day returns. Moreover, they find that firm specific matrices (i.e. number of followers and retweets) impact the level of IPO first-day returns.

4.3 Twitter Sentiment/Information and Earning Announcement

In a different context, other studies have also examined the market reaction to the earnings announcement and how it can be influenced by the sentiment/information extracted from Twitter and StockTwits. For instance, Jung, Naughton, Tahoun, and Wang (2014) study the determinants and effects of quarterly earnings disclosure on social media (Twitter and Facebook), the earnings announcement-related tweets are examined for a sample of firms included in the S&P 1500 index from 2010 to early 2013. They find that firms with lager number of followers are more reluctant to use Twitter for financial disclosure as using social media in earnings announcement is inversely related to the number of followers. Yet, firms tend to engage more in earnings news announcements when the news is positive. Also, the market reaction for firms with a consistent rather than ad-hoc disclosure policy on social media is found stronger

In consistency, Bartov, Faurel, and Mohanram (2017) examine the predicative power of individual opinions conveyed in the tweets before the firm's earnings announcement in forecasting the firm's earnings and the market reaction to these earnings. The authors use four alternative methods to extract aggregate information from individual tweets for a sample consisting of 998,495 tweets, covering 34,040 firm-quarters from 3,662 firms listed in Russell 3000 index from 2009 to 2012. They find that the aggregate opinion extracted individual tweets predicts a firm's coming quarterly earnings and announcement returns. These results are robust to the type of the information disseminated in the tweets (i.e. Tweets conveying original information, or disseminating existing information).

Jung, Naughton, Tahoun and Wang (2017) focus on the firms tweeting activity around quarterly earnings announcements to examine if firms are strategic "selective" in their dissemination of news. The authors measure the extent of dissemination using the number of tweets that a firm sends in a given quarter about the same earnings announcements. Based on the analysis of earnings-related tweets collected for the S&P 1500 index constituent list during the period (2010-2013), the findings - standing in line with strategic dissemination behavior - reveal that firms are less likely to disseminate bad news. Furthermore, firms tend to send fewer earnings announcement tweets and "rehash" tweets when the news is bad. Also, Yang, Liu, and Zhou (2016) test if corporate governance impacts the degree of financial disclosure/dissemination on a social media platform via analyzing 2014 earnings-related disclosures on Twitter for a sample of UK FTSE 350 firms. As a measure of information dissemination on Twitter, the authors use the number of earnings-related tweets. The overall findings suggest that firms with better governance mechanisms are more proactive in adopting social media as a complementary corporate disclosure channel.

In another study by Baik et al. (2016), the link between Twitter tone and earnings announcement returns and stock returns has also been examined. The authors also examine how this link is affected by the geographic proximity of Twitter users to companies (i.e. whether local users have better access to information about local firms than remote users and so exert more influence). The authors use lexicon-based sentiment extraction method - based on Loughran and McDonald's (2011) word list - to construct the tone of tweets for a sample of 552,012 messages on Twitter about US firms 2011/2012. They find that the tone in tweets by local Twitter users can predict future stock returns and subsequent earnings announcement returns, especially for firms with high information asymmetry. These findings stand in consistency with Giannini's at al. (2013) study on the impact of the geographic proximity of Twitter users on stock returns. Using a dataset of 216,266 messages posted on StockTwits platform for 1,819 US firms during the period (2009-2011), the results assert that local Twitter users are more informed of future stock returns than their nonlocal peers.

4.4 Twitter Sentiment/Information and Exchange Rates

The power of investor sentiment extracted from Twitter in predicating the exchange rate has also drawn the attention of several researchers. Papaioannou, Russo, Papaioannou, and Siettos (2013) analyze 20,250 tweets related to the EUR/USD exchange rate over the period (2010-2011) to extract information about the types of orders that each Twitter user had made, and the target-price of each of these orders. The findings provide supportive evidence of the role of this information in predicting the very short (intra-daily) foreign exchange rate. In a similar study, Janetsko (2014) examines to which extent the Twitter sentiment trends can forecast the daily

EUR/USD rates using Twitter posts from January 2013 to September 2013. The author follows a five-steps approach - based on "outperforming" concepts mining - in sentiment extraction from tweets, and uses this information as a predicator, finding that the sentiment based model consistently outperforms the random walk model.

Consistently, Ozturk and Ciftci (2014) investigate the relationship between Twitter content categorized as sentiments, such as Buy, Sell and Neutral, with USD/TRY currency daily movements during the year of 2013. The results reveal a relationship between the number of tweets and the change in USD/TRY exchange rate. In a similar study by Plakandaras, Papadimitrioua, Gogasa, and Diamantarasb (2015), the investor sentiment is proxied by the sentiment expressed in the messages posted on StockTwits about exchange rates (USD/EUR, USD/JPY, USD/GBP and AUD/USD) during the year 2013 based on the tweet writer's own assessment of the tweet sentiment. Using various forecasting econometric models, the authors find that social media sentiment can be another predictor of future directional movement of the exchange rates added to the traditional ones proposed by economic theory.

5. Conclusion and Future Recommendations

This paper aims at critically reviewing the theoretical interpretations of the impact of sentiment and information extracted from social media in the context of capital markets, and identifying the associated methodological difficulties with extracting this sentiment/emotional content. A synthesis of related studies (classified according to the context of study) is then conducted. Theoretically, the literature has introduced contradicting interpretations for this impact, some of which support the idea of behavioral finance (e.g. Investor sentiment and herding behavior) while the other explanations stand in line with the concept of market efficiency (e.g. social bond). Empirically, the findings are overall consistent with the behavioral explanations, exhibiting supportive evidence of the predictive power of investor sentiment in the context of capital markets.

In the future, further work could build on examining the impact of sentiment/information extracted from Twitter in specific (and from social media in general) in different contexts. For instance, the market reaction to several corporate events - such as seasoned equity offerings and mergers and acquisitions - represents new angle for future research. The impact of social media sentiment can be explored in the context of bond markets. Also, the impact of social media information/sentiment in the context of firms' downside risk (i.e. how investors respond to down movement in the market) is still unexplored. The expected empirical findings in these contexts will have important implications for firms to test if (and how) they can take advantage of social media sentiment/information in times of greater asymmetric information, ambiguity and market declines. Moreover, as the majority of the explanations of the role of the information/sentiment extracted from social media attribute this role to investor sentiment, future research should exert more efforts to develop more direct measures/models to proxy for other behavioral factors – rather than investor sentiment - such as emotional contagion and herding behavior in the context of social media.

References

Alexander, R., & Gentry, J. (2014). Using social media to report financial results. *Business Horizons, 57*(2), 161-167. https://doi.org/10.1016/j.bushor.2013.10.009

Ali, H. (2017). Behavioral Timing, Valuation and Post-issue Performance of UK Initial Public Offerings. *Journal of Behavioral Finance, 18*(10), 152-166. https://doi.org/10.1080/15427560.2017.1308938

Azar, P., & Lo, A. (2016). The wisdom of Twitter crowds: Predicting stock market reactions to FOMC meetings via Twitter feeds. *Journal of Portfolio Management, 42*(5), 123-134. https://doi.org/10.2139/ssrn.2756815

Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of returns. *Journal of Finance, 61*(4), 1645-1680. https://doi.org/10.3386/w10449

Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives*, *21*(2), 129-152. https://doi.org/10.2139/ssrn.973860

Baik, B., Cao, Q., Choi, S., & Kim, J. (2016). *Local Twitter Activity and Stock Returns.* Working paper. https://doi.org/10.2139/ssrn.2783670

Barberis, N., Shleifer, A., & Vishny, R. (1998). A Model of Investor Sentiment. *Journal of Financial Economics*, *49*(3), 307-343. https://doi.org/10.1007/978-3-8349-6170-9_3

Bartov, E., Faurel, L., & Mohanram, P. (2017). Can Twitter Help Predict Firm-Level Earnings and Stock Returns? *The Accounting Review,* In-Press. https://doi.org/10.2139/ssrn.2631421

Benthaus, J., & Beck, R. (2015). It's More about the Content than the Users! The Influence of Social

Broadcasting on Stock Markets. The *ECIS 2015 Completed Research Papers.* Retrieved from http://aisel.aisnet.org/ecis2015_cr/17/

Bi, X., & Gregory, A. (2011). Stock market driven acquisitions versus the Q theory of takeovers: UK evidence. *Journal of Business Finance and Accounting, 38*(5-6), 628-656. https://doi.org/10.1111/j.1468-5957.2011.02234.x

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science, 2*(1), 1-8. https://doi.org/10.1109/mc.2011.323

Brown, E. (2012). Bulls, Bear and Birds? Studying the Correlation between Twitter Sentiment and the S&P500. Proceedings of *the Thirty Third International Conference on Information Systems* (pp. 1-14). Orlando, Florida.

Brown, G., & Michael, T. (2004). Investor Sentiment and the Near-Term Stock Market. *Journal of Empirical Finance, 11*(1), 1-27. https://doi.org/10.2139/ssrn.1927801

Bukovina, J. (2016). Social media and capital markets. An overview. *Procedia- Social and Behavioral Sciences*, *220*, 70-78. https://doi.org/10.1016/j.sbspro.2016.05.470

Cade, N. (2018). *Corporate social media: How two-way disclosure channels influence investors*. Working paper. https://doi.org/10.2139/ssrn.2619249

Cole, B., Daigle, J., & Van Ness, B. (2015). Do Tweets Matter for Shareholders? An Empirical Analysis. *Journal of Accounting and Finance, 15*(3), 39-51.

Corea, F., & Cervellati, E. (2015). The Power of Micro-Blogging: How to Use Twitter for Predicting the Stock Market. *Eurasian Journal of Economics and Finance, 3*(4), 1-7. https://doi.org/10.15604/ejef.2015.03.04.001

Coviello, L., Sohn, Y., Kramer, A., … Marlow, C. (2014). Detecting emotional contagion in massive social networks. *PLoS ONE, 9*(3), 1-6. https://doi.org/10.1371/journal.pone.0090315

Daniel, K., Grinblatt, M., Titman, S., & Wermers, R. (1997). Measuring Mutual Fund Performance with Characteristic-Based Benchmarks. *Journal of Finance, 52*(3), 1035-1058. https://doi.org/10.2307/2329515

De Long, J., Shleifer, A., Summers, L., & Waldmann, R. (1990). Noise Trader Risk in Financial Markets. *Journal of Political Economy, 98*(4), 703-738. https://doi.org/10.1086/261703

Deephouse, D. (2000). Media reputation as a strategic resource: An integration of mass communication and resource-based theories. *Journal of Management, 26*(6), 1091-1112. https://doi.org/10.1016/s0149-2063(00)00075-1

Dey, S. (2014). Stock Market Prediction Using Twitter Mood. *International Journal of Scientific & Engineering Research, 5*(5), 44-47.

Dimson, E. and Mussavian, M. (1998). A Brief History of Market Efficiency and Stock Market Anomalies. *European Financial Management*, 4 (1), 91–103.

Dockery, E., & Kavussanos, M. (1996).Testing the efficient market hypothesis using panel data, with application to the Athens stock market. *Applied Economics Letters, 3*, 121-3. https://doi.org/10.1080/135048596356834

Dragota, V., & Oprea, D. (2014). Informational Efficiency Tests on the Romanian Stock Market: A Review of the Literature. *The Review of Finance and Banking, 6*(1), 15-28.

Elliott, W., Grant, S., & Hodge, F. (2018). *Investor Reaction to $Firm or #CEO Use of Twitter for Negative Disclosures.* Working paper. https://doi.org/10.2139/ssrn.2612444

Fama, E. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, *25*(2), 383-417. https://doi.org/10.2307/2325486

Fama, F. E. (1991). Efficient capital market II. *Journal of Finance, 46*(5), 1575-1617. https://doi.org/10.2307/2328565

Gans, H. (1979). *Deciding What's News*. New York, NY: Pantheon Books.

Giannini, R., Irvine, P., & Shu, T. (2014). *Do local investors know more? A direct examination of individual investors? Information set*. https://doi.org/10.2139/ssrn.1866267

Haslam, S., & Ellemers, N. (2005). Social Identity in Industrial and Organizational Psychology: Concepts, Controversies and Contributions. In G. Hodgkinsonand, & J. Ford (Eds.), *International Review of Industrial*

and Organizational Psychology. Chichester, UK: John Wiley & Sons, Ltd. https://doi.org/10.1002/0470029307.ch2

Herschberg, M. (2012). Limits to Arbitrage: An introduction to Behavioral Finance and a Literature Review. *Palermo Business Review, 7*, 7-21.

Hirshleifer, D., & Teoh, S. (2009). Thought and behavior contagion in capital markets. In T. Hensand, & K. Schenk (Eds.), *Handbook of financial markets: Dynamics and evolution* (pp. 1-56). North-Holland: Elsevier. https://doi.org/10.1016/b978-012374258-2.50005-1

Hong, H., Kubik, J., & Stein, J. (2005). Thy Neighbor's Portfolio: Word−of−Mouth Effects in the Holdings and Trades of Money Managers. *Journal of Finance, 60*(6), 2801-24. https://doi.org/10.2139/ssrn.331140

Ivković, Z., & Weisbenner, S. (2007). Information diffusion effects in individual investors' common stock purchases: Covet thy neighbors' investment choices. *Review of Financial Studies, 20*(4), 1327-1357. https://doi.org/10.3386/w13201

Janetzko, D. (2014). *Using Twitter to model the eur/usd exchange rate.* Working paper. https://doi.org/10.1007/s11066-014-9087-y

Jung, M., Naughton, J., Tahoun, A., & Wang, C. (2014). *Corporate Use of Social Media*. Working paper. Retrieved from https://www.rhsmith.umd.edu/files/JungMichael.pdf

Jung, M., Naughton, J., Tahoun, A., & Wang, C. (2017). Do firms strategically disseminate? Evidence from corporate use of social media. *The Accounting Review*, forthcoming. https://doi.org/10.2308/accr-51906

Karagozoglu, A., & Fabozzi, F. (2017*).* Volatility Wisdom of Social Media Crowds. *Journal of Portfolio Management, 43*(2), 136-151. https://doi.org/10.3905/jpm.2017.2017.1.065

Kolchyna, O., Souza, T., Treleaven, P., & Aste, T. (2015). *Twitter sentiment analysis: Lexicon method, machine learning method and their combination*. Working paper. Retrieved from https://arxiv.org/abs/1507.00955

Kramer, A., Guillory, J., & Hancock, J. (2014). Experimental Evidence of Massive-Scale Emotional Contagion through Social Networks. *Proceedings of the National Academy of Sciences, 111*(24), 8788-8790.

Lee, C., Shleifer, A., & Thaler, R. (1991). Investor sentiment and the closed-end fund puzzle. *Journal of Finance*, *46*(1), 75-109. https://doi.org/10.3386/w3465

Liew, J., & Budavari, T. (2017). The 'Sixth' Factor - Social Media Factor Derived Directly from Tweet Sentiments. *Journal of Portfolio Management*, forthcoming. https://doi.org/10.3905/jpm.2017.43.3.102

Liew, J., & Wang, G. (2016). Twitter Sentiment and IPO Performance: A Cross-Sectional Examination. *Journal of Portfolio Management, 42*(4), 129-135. https://doi.org/10.2139/ssrn.2567295

Liu, L., Wu, J., Li, P., & Li, Q. (2015). A social-media-based approach to predicting stock comovement. *Expert Systems with Applications, 2*(8), 3893-3901. https://doi.org/10.1016/j.eswa.2014.12.049

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *Journal of Finance, 66*(1), 35-65. https://doi.org/10.1111/j.1540-6261.2010.01625.x

Lundmark, L., Oh, C., & Verhaal, J. (2016). A little Birdie told me: Social media, organizational legitimacy, and underpricing in initial public offerings. *Information Systems Frontiers, 19*(6), 1407-1422. https://doi.org/10.1007/s10796-016-9654-x

Metaxas, P. T., Mustafaraj, E., & Gayo-Avello, D. (2011). How (Not) to predict elections. *Proceedings of PASSAT/SocialCom 2011, 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom)*, 165-171. https://doi.org/10.1109/passat/socialcom.2011.98

McQuail, D. (1985). The sociology of mass communications. *Annual Review of Sociology, 11*, 93-111. https://doi.org/10.1146/annurev.soc.11.1.93

Michael, L., & Otterbacher, J. (2014). Write like I write: Herding in the language of online reviews. In *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media* (pp. 356-365).

Mitchell, M., Pulvino, T., & Stafford, E. (2002). Limited arbitrage in equity markets. *Journal of Finance, 5*(2), 551-584. https://doi.org/10.2139/ssrn.267998

Muchnik, L., Aral, S., & Taylor, S. (2013). Social influence bias: A randomized experiment. *Science, 341*(6146), 647-651. https://doi.org/10.1126/science.1240466

Nisar, T., & Yeung, M. (2018). Twitter as a Tool for Forecasting Stock Market Movements: A Short-window Event Study. *The Journal of Finance and Data Science,* In Press. https://doi.org/10.1016/j.jfds.2017.11.002

Nofer, M., & Hinz, O. (2015). Using Twitter to predict the stock market: Where is the mood effect? *Business & Information Systems Engineering, 57*(4), 229-242.

Nofsinger, J. (2005). Social mood and financial economics. *Journal of Behavioral Finance, 6*(3), 144-160. https://doi.org/10.1207/s15427579jpfm0603_4

Oliveira, N., Cortez, P., & Areal, N. (2016). Stock market sentiment lexicon acquisition using microblogging data and statistical measures. *Decision Support System, 85*, 62-73. https://doi.org/10.1016/j.dss.2016.02.013

Oliveira, N., Cortez, P., & Areal, N. (2017). The impact of microblogging data for stock market prediction: Using Twitter to predict returns, volatility, trading volume and survey sentiment indices. *Expert System with Application, 73*, 125-144. https://doi.org/10.1016/j.eswa.2016.12.036

Ozturk, S., & Ciftci K. (2014). A Sentiment Analysis of Twitter Content as a Predictor of Exchange Rate Movements. *Review of Economic Analysis, 6*(2), 132-140. https://doi.org/10.13140/RG.2.1.1022.9201

Papaioannou, P., Russo, L., Papaioannou, G., & Siettos, C. (2013). Can social microblogging be used to forecast intraday exchange rates? *NETNOMICS: Economic Research and Electronic Networking, 14*(1-2), 47-68. https://doi.org/10.1007/s11066-013-9079-3

Peterson, R. (1979). Revitalizing the Culture Concept. *Annual Review of Sociology, 5*, 137-166. https://doi.org/10.1146/annurev.so.05.080179.001033

Plakandarasa, V., Papadimitrioua, T., Gogasa, P., & Diamantarasb, K. (2015). Market sentiment and exchange rate directional forecasting. *Algorithmic Finance, 4*(1-2), 69-79.

Pollock, T., & Rindova, V. (2003). Media legitimation effects in the market for initial public offerings. *Academy of Management Journal, 46*(5), 631-642. https://doi.org/10.5465/30040654

Ranco, G., Aleksovski, D., Caldarelli, G., Grcar, M., & Mozetic, I. (2015). The Effects of Twitter Sentiment on Stock Price Returns. *PLoS ONE, 10*(9), 1-21. https://doi.org/10.1371/journal.pone.0138441

Rao, T., & Srivastava, S. (2013). *Modeling Movements in Oil, Gold, Forex and Market Indices Using Search Volume Index and Twitter Sentiments*. Working paper. Retrieved from https://arxiv.org/abs/1212.1037

Rindova, V., Pollock, T., & Hayward, M. (2006). Celebrity firms: The social construction of market popularity. *Academy of Management Review, 31*(1), 50-71. https://doi.org/10.5465/amr.2006.19379624

Risius, M., Akolk, F., & Beck, R. (2015). *Differential Emotions and the Stock Market-The Case of Company-Specific Trading*. Paper presented at the European Conference on Information Systems. Retrieved from https://aisel.aisnet.org/ecis2015_cr/147/

Ritter, J., & Welch, I. (2002). A review of IPO activity, pricing, and allocations. *Journal of Finance, 57*(4), 1795-1828. https://doi.org/10.3386/w8805

Ruiz, E., Hristidis, V., Castillo, C., Gionis, A., & Jaimes, A. (2012). Correlating financial time series with microblogging activity. In proceedings of *the fifth ACM international conference on web search and data mining* (pp. 513-522). Washington, USA. https://doi.org/10.1145/2124295.2124358

Schmitz, P., Glaser, M., & Weber, M. (2007). *Individual investor sentiment and stock returns – What do we learn from warrant traders?* Working paper, University of Mannheim, Mannheim. https://doi.org/10.2139/ssrn.676844

Shapiro, D., Sheppard, B., & Cheraskin, L. (1992). Business on a handshake. *Negotiation Journal, 8*(4), 365-377. https://doi.org/10.1007/bf01000396

Shleifer, A., & Vishny, R. (2003). Stock Market Driven Acquisitions. *Journal of Financial Economics, 70*(3), 295-311. https://doi.org/10.3386/w8439

Shiller, R. (2000). *Irrational Exuberance*. Princeton, New Jersy: Princeton University Press. https://doi.org/10.1515/9781400865536

Shiller, R. (2003). From Efficient Market Theory to Behavioral Finance. *Journal of Economic Perspectives*, *17*(1), 83-104. https://doi.org/10.1257/089533003321164967

Souza, T., & Aste, T. (2015). *A nonlinear impact: Evidences of causal effects of social media on market prices*. Working paper. Retrieved fromhttps://arxiv.org/abs/1601.04535

Sprenger, T., & Welpe, I. (2011). Tweets and peers: Defining industry groups and strategic peers based on investor perceptions of stocks on Twitter. *Algorithmic Finance, 1*(1), 57-76.

Sul, H. K., Dennis, A. R., & Yuan, L. (2014). Trading on Twitter: The Financial Information Content of Emotion in Social Media. *Proceeding of the 47th Hawaii International Conference on System Sciences* (pp. 806-815). Hawaii: United States. https://doi.org/10.1109/hicss.2014.107

Sun, A., Lachanski, M., & Fabozzi, F. (2016). Trade the tweet: Social media text mining and sparse matrix factorization for stock market prediction. *International Review of Financial Analysis, 48*, 272-281. https://doi.org/10.1016/j.irfa.2016.10.009

Takhteyev, Y., Gruzd, A., & Wellman, B. (2012). Geography of Twitter networks. *Social Networks, 34*(1), 73-81. https://doi.org/10.1016/j.socnet.2011.05.006

Trinkle, B., Crossler, R., & Bélanger, F. (2015). Voluntary disclosures via social media and the role of comments. *Journal of Information Systems, 29*(3), 101-121.

Wang, G., Wang, T., Wang, B., … Sambasivan, D. (2014). *Crowds on wall street: Extracting value from social investing platforms.* Working paper. https://doi.org/10.1145/2675133.2675144

Welch, I. (2000). Herding among security analysis. *Journal of Financial Economics, 58*(3), 369-396. https://doi.org/10.1016/s0304-405x(00)00076-3

Wlodarczak, P., Ally, M., & Soar, J. (2015). *Opinion Mining in Social Big Data.* Working paper. https://doi.org/10.2139/ssrn.2565426

Yang, J., Liu, S., & Zhou, D. (2016). *Voluntary Financial Disclosure on Social Media: Does Corporate Governance Matter*? Working paper. https://doi.org/10.2139/ssrn.2836570

Zhang, X., Ma, L., & Wang, G. (2017). Investigating consumer word-of-mouth behavior in a Chinese context. *Total Quality Management & Business Excellence, 28*(1), 1-15.

Notes

Note 1. Earlier theoretical roots can be found in the context of sociology of mass communications (Gans, 1979; McQuail, 1985; and Peterson, 1979).

Note 2. E.g., Brown and Cliff (2004, p. 2) define investor sentiment as "the 'expectations of market participants relative to a norm: a bullish (bearish) investor expects returns to be above (below) average, whatever average may be". Baker and Wurgler (2006, p 1649) define it as "optimism or pessimism about stocks in general". Schmitz, Glaser, Weber (2006, p. 4) define it as the expectation of investors regarding the price of one or more financial assets that is not based on fundamental information.

Note 3. See Grinblatt et al. (1995), Welch (2000), and Hirshleifer and Teoh (2009) for detailed explanation and literature reviews of the herding behavior in the context of finance.

Note 4. See Kolchyna et al. (2015) and Oliveira et al. (2016) for detailed explanation of the different approaches of sentiment analysis.

Appendix A. Summary of the main empirical findings

Tested Variable	Source	Sample	Time Period	Empirical Findings
Stock Performance (Return, Risk, Trading Volume)	Bollen et al. (2011)	DJIA index	Feb.- Dec. 2008	**Supportive**
	Sul et al. (2014)	S&P 500 index	2011- 2013	
	Rao and Srivastava (2014)	DJIA, NASDAQ-100 Indices and 11 tech. stocks	2010-2011	
	Dey (2014)	One US stock of Apple	Jan. - April 2014	
	Ranco et al. (2015)	DJIA index	2013- 2014	
	Cole et al. (2015)	215 US stocks	2010-2011	
	Souza and Aste (2015	DJIA index	2012-2014	
	Liu et al. (2015)	293 US stocks	July-Dec. 2013	
	Corea and Cervellati (2015)	Three US stocks of Apple, Google and Facebook	Sep.- Nov. 2014	
	Sprenger and Welpe (2011)	S&P 500 index	Jan. -June 2010	
	Sun et al. (2016)	420 US Stocks	2011- 2015	
	Corea (2016)	Three US stocks of Apple, Google and Facebook	Sep. -Nov.2014	
	Azar and Lo (2016)	CRSP index	2007 – 2014	
	Liew and Budavari (2017)	15 US stocks	2013-2015	
	Karagozoglu and Fabozzi (2017)	the S&P 500 index	2012- 2016	
	Nisar and Yeung (2018)	FTSE 100 index	2018	
Stock Performance (Return, Risk, Trading Volume)	Benthaus and Beck (2015)	S&P 100 index	Jan- April 2014	**Mixed**
	Risius et al. (2015)	33 US stocks	Jan-March 2014	
	Nofer and Hinz (2015)	The DAX index	2011-2013	
	Wang et al. (2015)	9300 US firms	2009-2014	**Non-Supportive**
	Brown (2012)	S&P 500 index	2011-2012	
IPOs' Short-Term Returns	Liew and Wang (2016)	325 US IPOs	2013-2014	
	Lundmark et al. (2016)	272 US IPOs	2009-2011	
The Market Reaction to Earnings Announcement	Giannini et al. (2013)	1,819 US firms	2009-2011	**Supportive**
	Jung et al. (2014)	S&P 1500 index constituents	2010-2013	
	Baik et al. (2016)	646 US Stocks	2011-2012	
	Yang et al. (2016)	UK FTSE 350	2014	
	Bartov et al. (2017)	Russell 3000 index	2009- 2012	
	Jung et al. (2017)	S&P 1500 index constituent	2010-2013	
The exchange Rate	Papaioannou et al. (2013)	EUR/USD	2010-2011	
	Ozturk and Ciftci (2014)	USD/TRY	2013	
	Janetsko (2014)	EUR/USD	Jan.- Sep 2013	
	Plakandaras et al. (2015)	USD/EUR, USD/JPY, USD/GBP AUD/USD	2013	

Copyrights