Association for Information Systems
AIS Electronic Library (AISeL)

ECIS 2015 Completed Research Papers

ECIS 2015 Proceedings

It's More about the Content than the Users! The Influence of Social Broadcasting on Stock Markets

Janek Benthaus
Goethe University, jbenthaus@wiwi.uni-frankfurt.de

Roman Beck
IT University of Copenhagen, beck@itu.dk

Follow this and additional works at: http://aisel.aisnet.org/ecis2015_cr

IT'S MORE ABOUT THE CONTENT THAN THE USERS! THE INFLUENCE OF SOCIAL BROADCASTING ON STOCK MARKETS

Complete Research

Benthaus, Janek, Goethe University Frankfurt, Frankfurt am Main, Germany, jbenthaus@wiwi.uni-frankfurt.de

Beck, Roman, IT University of Copenhagen, Copenhagen, Denmark, romb@itu.dk

Abstract

Social broadcasting networks facilitate the public exchange of information and contain a large amount of stock-related information. This data is increasingly analyzed by research and practice to predict stock market developments. Insights from social broadcasting networks are used to support the decision-making process of investors and are integrated into automatic trading algorithms to react quickly to broadcasted information. However, a comprehensive understanding about the influence of social broadcasting networks on stock markets is missing. In this study, we address this gap by conceptualizing and empirically testing a model incorporating three dimensions of social broadcasting networks: users, messages, and discussion. We analyze 1.84 million stock-related Twitter messages concerning the S&P 100 companies between January and April 2014 and corresponding intraday stock market data from NYSE and NASDAQ. Our research model is constructed applying factor analyses and tested using a fixed effects panel analysis. The results 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. Specifically, the influence of user mentions, financial sentiment, discussion reach, and discussion volume has the largest impact and should carefully be considered by investors making trading decisions.

Keywords: Social Broadcasting Dimensions, Stock Market Activity, Panel Analysis, Decision Support.

1 Introduction

Social broadcasting networks, such as Twitter, empower a large number of users to act as content providers delivering information to the public (Shi et al., 2014). Users share messages containing their opinions and thoughts about certain topics and reveal how they are connected to each other. Thus, the user activities provide a continuous stream of valuable information that can be analyzed in real-time and at low cost (Dellarocas, 2003). In this regard, research on social media has found that information shared on social broadcasting networks can be used to predict election results (Tumasjan et al., 2010), box office revenues (Rui et al., 2013), and even disease outbreaks (Signorini et al., 2011) and can also have an impact on economic outcomes (Granovetter, 2005). Social broadcasting messages regarding the public's interest in financial topics can be deployed to enhance the understanding of stock market participants' behavior (Preis et al., 2013). The predictive power of user-generated content on stock markets has firstly been shown by research on online discussion boards (Antweiler and Frank, 2004). With the rise of social media, research also investigated whether information derived from social

broadcasting can be used to predict developments at stock markets (e.g., Bollen et al., 2011; Oh and Sheng, 2011; Sprenger et al., 2014b).

The impact of social broadcasting on stock markets has been shown in several events when broadcasted information led to significant price changes at the markets. In April 2013, a faked message sent through the Associated Press Twitter account caused a significant drop of the S&P 500 index, erasing $136 billion (Bloomberg, 2013). Corrections were published immediately and the stock markets recovered within five minutes. This event revealed how fast information from Twitter is integrated in investment decisions (Fortune, 2013). Moreover, Tweets can also have sustainable impact on stock prices. In 2013, a well-known investor, revealed on Twitter information about his investment in Apple and his opinion about the company's future stock price (Yahoo!Finance, 2013). Broadcasting this information led to an increase in Apple's firm value of about $12.5 billion that day. In conclusion, both examples show that social broadcasting significantly affects market participants in their trading decisions and the information is quickly integrated into decision making.

Existing research on social broadcasting and stock markets analyzed how broadcasted information can be used to predict stock prices (Bollen et al., 2011; Sprenger et al., 2014b; Zhang et al., 2011; Zheludev et al., 2014). Data from social broadcasting networks reveals information about the participating users, the shared messages, and the interaction among the community. However, existing studies focus mainly on the predictive power of sentiments and trading signals like buy- and sell-signals on stock prices and stock indices (e.g., Nann et al., 2013; Sprenger et al., 2014b). Analyzing the interplay of different constructs in a single research model on stock markets is missing. Hence, a comprehensive understanding of how social broadcasting networks impact stock markets is missing. Moreover, existing research uses data on a daily level (e.g., Sul et al., 2014) missing to account for the fast reaction of stock markets on new information with narrower timeframes.

Accordingly, we will address both research gaps by analyzing the influence of the dimensions of social broadcasting networks on stock markets using intraday data. We develop a conceptual model incorporating a large set of constructs in a single research model and empirically test it by analyzing social broadcasting data about the individual stocks of the S&P 100 companies. Thereby, we want to understand how the users, their messages, and the discussion about the stocks have an impact on the trading activity of the companies' stocks. Based on a dataset of more than 1.84 million stock-related Twitter messages and hourly stock market data, collected between January and April 2014, we examine which dimensions of social broadcasting influence the trading activity at stock markets. Hence, we are interested in *how the dimensions users, messages, and discussion of social broadcasting are influencing trading activity on stock markets*. With this study, we contribute to the literature on social media by developing a conceptual model that allows a comprehensive analysis of social networks rather than focusing on single aspects. Furthermore, our study enhances our understanding of how social broadcasting and stock markets are connected.

The remainder is structured as follows. Section two provides the theoretical background of social broadcasting and theory of financial markets. In section three, we develop our conceptual model and hypothesize the relation of the social broadcasting dimensions users, messages, and discussion and stock market trading activity. Afterwards, in section four, we conduct an empirical study to test our model with data from Twitter, NYSE and NASDAQ by applying a fixed effects panel analysis. The paper ends with a conclusion as well as limitations and opportunities for future research.

2   Background on Stock Markets and Social Broadcasting

Stock markets play a central role in the economies to exchange stocks and thereby, determine stock prices. The determined prices reflect the evaluations of the market participants regarding the companies' success in the future (Shiller, 2003). The price development of financial instruments at stock markets, is described by the famous finance theory of the efficient market hypothesis stated by

Fama (1970). The efficient market hypothesis explains that stock markets are efficient in the way that all available and relevant information about the prospects of companies are fully reflected in their stock prices (Shiller, 2003). With respect to the availability of complete information, market participants act as rational individuals and cannot forecast price changes as additional information do not improve the decision-making. This implies that market participants cannot gain systematically higher returns than the average market return. Based on this assumption on the efficiency of stock markets, the random walk hypothesis states that stock price developments cannot be predicted and follow a random walk (Fama, 1965). However, the efficient market hypothesis is rather strict in its assumptions and is empirically and theoretically criticized (Shiller, 2003). In this context, theory of behavioral finance accounts for behavioral biases of individuals which influence their decision-making in situations characterized by uncertainty (Lo, 2004). Therefore, the preferences and the behavior of market participants such as overconfidence, overreaction, risk aversion, and herd behavior are taken into account by behavioral finance theory to explain stock market developments. As a consequence, market participants do not always act rational which results in phases were markets are inefficient (Shiller, 2003). The irrational behavior of market participants makes market developments predictable as stock markets do not always follow a random walk (Lo, 2004). Thus, on inefficient markets arbitrage opportunities exists which makes gathering and processing of additional information valuable to outperform the market (Grossman and Stiglitz, 1980). Hence, analyzing stock-related information can help to draw conclusions on stock market developments and to seize investment opportunities.

Market participants can gather information from several sources to monitor stock developments and stock evaluations of others to improve their trading decisions. Besides traditional resources like financial data, company announcements or newspapers (Fang and Peress, 2009), internet platforms on which users share stock-related information (e.g., Twitter, StockTwits) present a promising source to support trading decisions. In this regard, online message boards such as Yahoo! Finance enable users to discuss their expectations about certain stocks (Antweiler and Frank, 2004). Today, social media presents an additional channel to discuss stock-related information (Sprenger et al., 2014b) where users are empowered to engage in social interaction with each other and to act as content providers (Shi et al., 2014). Especially, social broadcasting networks gained large attention in practice for sharing stock-related information (Greenfield, 2014). Individuals use social broadcasting networks to share their opinions and thoughts about certain topics (Dellarocas, 2003) which others integrate into their decision-making process (Rindova et al., 2005). A social broadcasting network is defined as a technology that fosters the diffusion of information in a way that weak ties are exploited to overcome social distance (Granovetter, 1973; Shi et al., 2014). Weak ties exist in such networks based on the unidirectional form of how relationships can be established to receive broadcasted information (Shi et al., 2014). Unidirectional relationships are established to receive the message from other users without permission of the broadcasting user. Hence, information shared on such platforms can reach an enormous group of people (Kaplan and Haenlein, 2010). In essence, key features of social broadcasting are the speed and reach of the information provision as well as the large number of contributors (Rui and Whinston, 2012b; Shi et al., 2014).

In order to analyze social broadcasting, research makes use of different features of social broadcasting networks. They analyze information about the users (e.g., Rui and Whinston, 2012b; Shi et al., 2014), the content of the messages (e.g., Liu et al., 2012; Rui et al., 2013) and to what extent messages are exchanged (e.g., Rui et al., 2010; Weng et al., 2010). Specifically, the users are investigated based on their social structure considering how many in- and out-going connections they have in the network (Lu et al., 2014) or based on the profile information of their accounts (Rui and Whinston, 2012b). The messages are analyzed based on the sentiment of contributed information (Rui et al., 2013) and the included information such as links or hashtags (Liu et al., 2012) as well as how often a message was redistributed (Suh et al., 2010; Goh et al., 2013). How topics are discussed is considered using data about the volume of message, the amount of users who received the message and whether users

participate in the discussion by contributing messages or sharing existing messages (Rui et al., 2010; Sprenger et al., 2014b). Aggregating these constructs shows that there are three superior dimensions of social broadcasting that can be considered by research: user, message, and discussion dimension. The user dimension describes the characteristics of the users providing stock-related messages, the message dimension takes into account which kind of information is provided by the messages and the discussion dimension represents the extent to which information regarding companies stocks is shared.

A growing body of literature examines the relation between social broadcasting and stock markets. Studies analyze how data form social broadcasting is related to stock indices (Bollen et al., 2011; Zhang et al., 2011), individual stocks (Oh and Sheng, 2011; Sprenger et al., 2014b), and exchange rates (Papaioannou et al., 2013). A reason to focus in the beginning on stock indices as aggregate measure was due to the available amount of social broadcasting data. The predictive power of social broadcasting data on stock markets depends on the sufficient volume of shared messages. Due to an increase in social broadcasting messages over the last years, recent research found that sentiment in social broadcasting networks can even predict the development of individual stocks instead of aggregated stock indices (Sprenger et al., 2014b; Zheludev et al., 2014). The volume of social broadcasting messages depends on the on hand on the interest of users in the stock (Mao et al., 2013) and on the other hand on the available access level to the data. Most research focused on a daily level due to an insufficient volume of stock-related messages or limits that restrict the full access to the broadcasted messages (e.g., Zhang et al., 2011; Sul et al., 2014). Based on the integration of social media data into the trading systems, market participants nowadays react quickly to broadcasted information. To analyze relations of social broadcasting on individual stocks on an hourly basis a full access to Twitter is required (Zheludev et al., 2014). Following this, there are calls for research on single stocks using narrower timeframes to account for the fast reaction of stock markets to new information (Oh and Sheng, 2011; Sprenger et al., 2014b). We will answer the calls by using hourly intraday data with an unrestricted Twitter access investigating single stock developments.

Regarding the methods used to analyze the influence of social broadcasting on stock markets, studies apply sentiment analyses (Zhang et al., 2011), data mining techniques to classify tweets based on categories (Sprenger et al., 2014a), and social network analyses (Mao et al., 2013; Sprenger et al., 2014b). Although, sentiment analysis is largely applied in this research area, we found only one study conducting a context specific sentiment analysis (Li et al., 2014). The authors adopt the wordlist from Loughran and McDonald (2011) as they showed that almost three-quarters of the negative words in the Harvard emotion dictionary should not considered negative when analyzing text with a financial context. Following Li et al. (2014) we will conduct a domain specific sentiment analysis in our empirical study. Furthermore, analysis of the social structure of a network shows that the number of relationships is an indicator for the impact of the users on stock markets as well as the number of user mentions in other users' messages (Sul et al., 2014; Sprenger et al., 2014b). In conclusion, existing studies include constructs from one to two of the three dimensions of social broadcasting and only recent research from Sprenger et al. (2014b) applied constructs addressing all three dimensions, but not in a single research model. Thus, a comprehensive analysis considering the interplay of the three dimensions of social broadcasting on stock markets is missing. Hence, we address this research gap by developing a research model including constructs from all dimensions in a single research approach.

3 Conceptual Model: Impact of Stock-related Social Broadcasting on Stock Market Activity

Reflecting existing literature on social media and finance, our proposed research model is depicted in Figure 1. Based on our preceding discussion, we argue that the impact of social broadcasting networks on trading activity at stock markets can be analyzed based on three dimensions: user, message, and discussion. Thereby, we aggregate a comprehensive set of constructs based on prior research to evaluate the impact of a social broadcasting network on stock markets.

3.1   User Dimension

The user dimension reflects the expertise and the popularity of the contributing users in the social broadcasting network. The expertise of a user is an important factor for information exchange (Jones et al., 1997) and has been proposed to impact the ability to exchange information in a meaningful manner (Wasko and Faraj, 2005). Hence, a higher expertise has a positive effect on the quality of exchanged information (Constant et al., 1996) and users with a higher expertise provide more valuable information in a social broadcasting network (Nevo et al., 2012). This is reflected by the fact that messages of users with higher expertise are redistributed by other users to a greater extent (Liu et al., 2012). Additionally, Larson et al. (1996) found that individuals with higher expertise exchange more information including new content that has not been shared before and expertise increases the chance that individuals exchange information at all (Wasko and Faraj, 2005). Hence, users with a higher level of expertise are responsible for the majority of exchanged information in a network (Wasko et al., 2009). Thus, stock-related information from users with a higher expertise should have a greater impact on stock market activity.
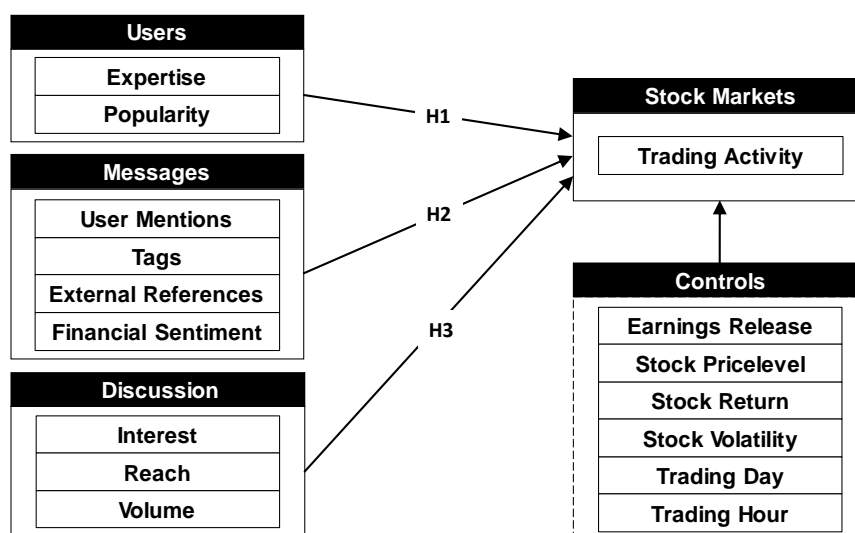

Figure 1.   Research model for analyzing the impact of social broadcasting on stock markets.

Additional to the users' expertise, Sul et al. (2014) found that the impact of shared information from more popular users on the stock market is greater. The popularity of an individual can signal reliability to other individuals within a network (Jones et al., 1997). Individuals are concerned about their popularity in a network (Wasko et al., 2009) and will contribute valuable information to maintain or enhance their popularity (Wasko and Faraj, 2005). In a social broadcasting network the number of connections of a user represents the user's popularity (Zhang et al., 2011). Additionally, individuals with a higher number of connections exchange information with a greater number of individuals in a network (Ahuja et al., 2003). Furthermore, it is proposed that these individuals are more involved in a network which has a positive influence on the sharing of information (Wasko and Faraj, 2005). Hence, research has found that the number of connections reflects how influential a user is within a social broadcasting network (Cha et al., 2010). Consequently, stock-related information provided by such users should have a greater influence on trading activity. Given these assumptions, information provided by users that are more popular and having a higher expertise should have greater influence on others. Hence, we assert that such user characteristics positively influence trading activity. This leads to the following hypothesis:

H1:   Users of a social broadcasting network which are more popular and have a higher expertise have a positive influence on trading activity.

3.2 Message Dimension

Investors need valuable information to support the typically complex and risky investment decisions (Grossman and Stiglitz, 1980). The message dimension covers information about the content of the messages shared on social broadcasting networks. Users have the opportunity to tag keywords, add links to external reference, or mention other users in their messages (Suh et al., 2010). Users providing valuable information are proposed to be mentioned more often in messages on social broadcasting networks (Sprenger et al., 2014b). Accordingly, Cha et al. (2010) found that such users have greater influence in a social broadcasting network. By mentioning other users in their messages, users can connect their content with how the mentioned users are perceived in the network (Mohamed et al., 1999). Therefore, user mentions are used to increase the perceived validity of the shared information. To organize the discussion around a certain topic users can add tags in their messages (Kane et al., 2014). Tags makes it easier to follow a discussion or to monitor keywords that are of interest as tags signal commonality of messages (Celli and Rossi, 2012). The amount of information within social broadcasting messages is limited due to length restrictions of the messages by including external references directing to further information the messages can be enriched with further information. Research found that external references to valuable information help to increase proliferation of stock-related information within the network (Bakshy et al., 2011) which in turn is supposed to support the trading decisions of market participants. In essence, messages with a higher information richness are found to support individuals in their decision making (Goh et al., 2013).

In addition to information that can be directly extracted from a social broadcasting message, it is possible to process the content of a message by using language processing techniques (Hu and Liu, 2012). Especially, sentiment analysis is applied to analyze the opinions and moods of users in social media (Liu, 2010). It is important to consider the sentiment in social broadcasting as messages containing emotional content are shared more frequently and at a greater speed (Stieglitz and Dang-Xuan, 2013). Detecting the sentiment of news, ad-hoc notifications or social media content is largely used in financial research (Bollen et al., 2011; Geva and Zahavi, 2014; Loughran and McDonald, 2011) to predict stock prices. Stock exchanges such as NYSE and NASDAQ integrate information about stock sentiments in their trading systems to support investment decisions. Furthermore, the sentiment of messages can be used as indicator for buy-, hold-, and sell-signals (Zhang et al., 2011). Hence, stock-related messages containing positive sentiments should result in a positive mood about the stock and an increased trading activity. In turn, uncertainty expressed by negative sentiments results in more cautious investors and the higher likelihood that they are not investing (Zhang et al., 2011). Thus, messages containing positive emotions should lead to an increase in trading activity. Based on these explanations, we argue that messages shared in social broadcasting networks with a richer set of information and which are expressed more positively should help investors to make trading decisions. This results in the following hypothesis:

H2: *Stock-related social broadcasting messages with a richer set of information and which are expressed positively have a positive influence on trading activity.*

3.3 Discussion Dimension

The third dimension of our conceptual model comprises how the stocks are discussed in terms of interest in the information, reach of the information and the amount of shared information in the network. Social broadcasting networks like Twitter allow users to redistribute messages composed by other users (Suh et al., 2010). Thus, users can show their interest in the information by redistributing them. The interest of individuals regarding companies' stocks is found to influence trading activity (Preis et al., 2013). Although the efficient market hypothesis would neglect the impact of redistributed and hence already reflected information on the trading activity at stock markets, behavioral finance shows that herd behavior can impact stock prices when irrational investors act collectively (Lo, 2004).

Hence, users give the discussion about a stock more attention by showing interest in stock-related information in the form of redistributing messages, which we propose affects stock markets.

Discussions of stock-related information can reach a large number of users on social broadcasting networks (Cha et al., 2010). The reach of a message depends upon how many users have connected oneself with a content provider to receive his or her messages (Rui and Whinston, 2012a). Additionally, redistributed messages extend the reach of information by increasing the number of users retrieving the message (Shi et al., 2014). The influence of users on other users increases with the number of connections a user has to others in a network (Rui et al., 2013). Hence, stock-related messages sent by users with a higher number of connections are proposed to affect trading behavior of others. Furthermore, in social broadcasting networks, the volume of exchanged information has been found to have an impact on economic outcomes based on word-of-mouth (Rui et al., 2013). Antweiler and Frank (2004) analyzed discussions on messaging boards about stocks and found that the amount of messages can predict the trading volume at stock markets. A positive correlation of the volume of social broadcasting messages and trading volume has been confirmed by Sprenger et al. (2014b). Moreover, Tirunillai and Tellis (2012) found that the volume of messages is even suitable to predict future stock returns. Thus, we include the volume of messages as a factor of a stock-related discussion in our model. Given this argumentation, the discussion of stock-related information on social broadcasting networks should affect others trading decisions, which leads to the final hypothesis:

H3:   An increased discussion of stock-related information in terms of interest, reach, and volume in social broadcasting networks has a positive influence on trading activity.

4   Empirical Study

In this section, we analyze how social broadcasting of stock-related information is related to stock market activity based on our conceptual model. Analyzing social broadcasting activities require coping with a large amount of data (Greenfield, 2014). Volume as well as velocity of social data is rapidly increasing and the fast reaction of stock markets to new events requires close to real-time data analysis. These circumstances have brought manual approaches for data analysis to their limits and procedures for automated analysis are needed. In this respect, the knowledge discovery process described by Fayyad et al. (1996) proposes a way how to handle large amount of data. We adapt this process, originally developed to handle analyses in database environments, to our social media analytics purpose. The resulting research approach consists of four consecutive steps depicted in Figure 2, which will be discussed in the subsequent subsections.
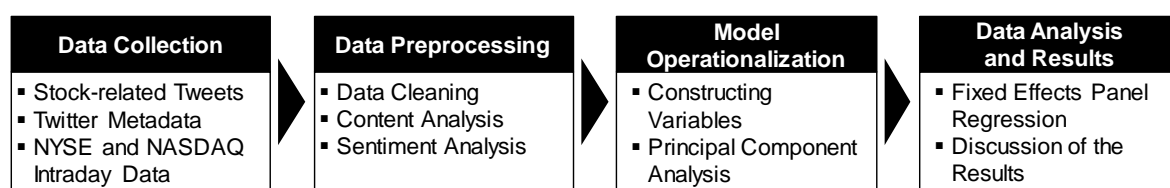
Data Collection	Data Preprocessing	Model Operationalization	Data Analysis and Results
▪ Stock-related Tweets ▪ Twitter Metadata ▪ NYSE and NASDAQ Intraday Data	▪ Data Cleaning ▪ Content Analysis ▪ Sentiment Analysis	▪ Constructing Variables ▪ Principal Component Analysis	▪ Fixed Effects Panel Regression ▪ Discussion of the Results

Figure 2.   Research approach to analyze the relation of social broadcasting and stock markets.

4.1   Data Collection

The process of Fayyad et al. (1996) starts by choosing the data source needed for the analysis. During the last years a variety of social broadcasting networks emerged from which Tumblr, Twitter, and Weibo are among the most known ones today. For research purposes, Twitter has gained a lot of attention because of its large user base contributing 500 million messages each day (Twitter, 2015). Moreover, the amount of stock-related messages is increasing heavily in the past years (Greenfield, 2014). To analyze the relation between the social broadcasting activity and financial markets we focus

on messages related to companies' stocks. In line with existing literature, we concentrate on well-known large companies whose stocks are extensively traded at the sock markets and are also intensively discussed on social broadcasting networks (Geva and Zahavi, 2014; Zheludev et al., 2014). Due to this, we make sure to have enough and reliable data for our subsequent panel analysis. Hence, we follow Sprenger et al. (2014b) and analyze messages mentioning the stocks of the companies listed in the S&P 100 stock market index. The S&P 100 includes the 100 major, blue-chip U.S. stocks of companies from different industry sectors. The S&P 100 index is rebalanced every quarter and we selected the list of stocks based on the constitution as of March 31st 2014. The index includes the stocks of companies such as Apple, Exxon Mobil, General Electric, Google, Starbucks and Wal-Mart.

For our analysis, we collected the Twitter data via a certified analytics company that provides us with full access to the Twitter "Firehose". Thereby, we can overcome the issue of limited data provision that encumbered intraday analysis which prior research were faced with (e.g., Sprenger et al., 2014b). Collecting data from Twitter confronts us with the challenge to select appropriate keywords, which are used to extract the relevant messages. To get a high percentage of messages that actually contain information and opinions about stocks, we followed Nann et al. (2013), Sprenger et al. (2014b), and Sul et al. (2014) and made use of Twitter's cashtag-feature. Cashtags are composed using the stock symbols (uniquely identifier of stocks) with a dollar-sign placed in front (e.g., $GOOG for Google or $KO for Coca Cola) to refer to a company's stock. Such notation makes it easier for users to monitor the broadcasted information about stocks (Kane et al., 2014) and reduces the amount of unrelated information when trying to follow the discussion about a companies' stock. Therefore, we used the cashtags of the S&P 100 companies as keywords for data extraction. We were able to collect 1.84 million Twitter messages shared between January 1st and April 30th 2014. The obtained dataset of Twitter data not only contains the content of a message, but also meta-information about the message contributors (e.g., the number of followers). The corresponding stock market data was gathered from the NYSE and the NASDAQ, which are the issuing market of the companies' stocks and the most relevant markets in terms of trading volume. Both stock exchanges have trading session from 9:30am to 4:00pm. We collected hourly intraday data for the same period as the Twitter data resulting in seven observations per stock and per day for 82 trading days. The stock data includes information about hourly opening, high, low, and closing prices as well as the number of trades and the traded volume.
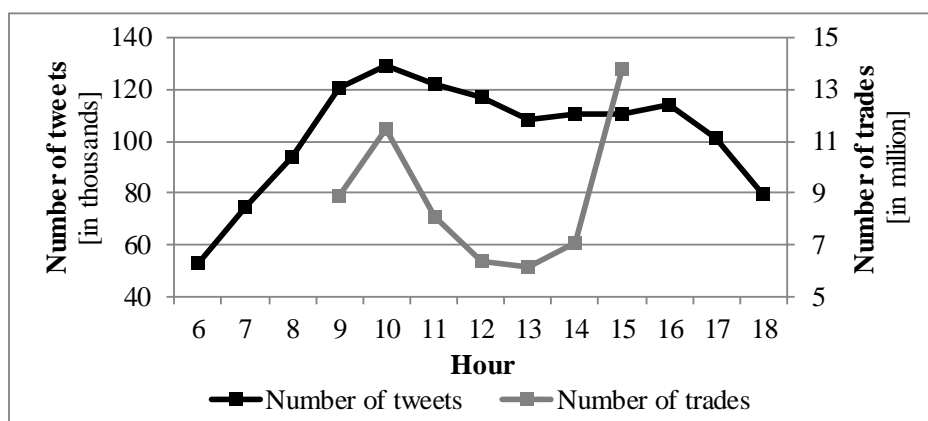

Figure 3.   Average number of tweets and trades per hour.

Figure 3 depicts he hourly average number of tweets and the corresponding number of trades. The trading activity increases after the stock markets open, drops during noon and reaches its maximum in the end of the trading session. The amount of tweets follows this curve, but without an equally significant decrease during the trading session. The tweet volume slightly increases after the trading session which suggests the users summarize the trading day before the tweet volume begins to fall.

4.2 Data Preparation

Although we gathered the data using the cashtags as keywords to only extract stock-related messages, it is necessary to clean the dataset by excluding spam (Nann et al., 2013). Hence, we screened the Twitter data to refine the final dataset for further analysis. We used the meta-information of the tweets to develop two spam-rules to identify users posting irrelevant content. First, we excluded tweets from users who have no followers receiving their tweets but have posted more than 20 tweets. The reason for excluding them is that the information of these tweets is not received by other users and therefore should have no impact on trading decisions of others. Second, tweets from users having less than 20 followers, a total posting count of more than 1,000 messages and follow less than 10 users are excluded. These accounts are often used to provide dubious content. Additionally, dubious advertisers or spammers often use cashtags of well-known companies to promote their investment strategies (e.g., investing in penny stocks). Such behavior is also found in newsletters or on websites to provide deceptive content (Siering, 2013). Hence, we excluded tweets of 68 users and tweets with hashtags like "#pennystock" or "#stock #stocks #stockaction". Such messages always include links to direct readers to doubtful websites rather than providing meaningful information on the S&P 100 stocks. Furthermore, we found that the stock symbols regarding three companies (i.e., ALL, LOW, and WAG) are largely used for other purposes than to talk about these stocks. Therefore, the noise regarding these companies' stock is very high and we decided to exclude the three companies from the final dataset. Next, we found that tweets containing more than one cashtag make it difficult to draw conclusions about the mentioned stocks. Hence, we decided to focus on messages that include only one cashtag to reduce ambiguity (Sul et al., 2014). Performing the aforementioned data cleaning steps, our final data set comprises 956,317 stock-related messages regarding 97 companies.

Next, we analyzed the content and the metadata of the remaining tweets in our dataset. Thereby, we determined whether tweets contain links, hashtags, or mentions of other Twitter users. We further evaluated the mood of the messages to determine whether a message contains emotional content. We used the publicly available tool SentiStrength, developed by Thelwal et al. (2010) to perform a sentiment analysis on short informal text such as Twitter messages. SentiStrength is designed to cope with the characteristics of microblogging messages like accounting for abbreviations due to short text size and emoticons. SentiStrength comes along with a list of words signaling emotions based on the general inquirer word list (General Inquirer, 2014). However, words can have very different meanings depending on the domain in which the sentiment analysis is applied (Thelwall et al., 2012). Therefore, Loughran and McDonald (2011) developed a special word list for sentiment analysis in the financial context to acknowledge domain specific meanings during the classification of messages. Hence, to address this issue and to increase the reliability of the result of our sentiment analysis we used this word list consisting of more than 2,600 words signaling emotions.

4.3 Model Operationalization

The next step was to construct the variables to examine our hypotheses. Whenever possible, we refer to already applied measures from prior empirical studies in the area of social broadcasting. Afterwards, we discussed the selected measures with a panel of two researchers, regarding both their validity and relevance. To analyze the effect of social broadcasting on trading activity on stock markets, we used as dependent variable the number of trades regarding a stock. We choose the number of trades over the traded volume because a large trading volume reflects not necessarily increased interest in stock as it can be the result of large orders from institutional investors. Hence the number of trades is more robust against large orders from institutional investors as every trade has equal weight independent of its size. To avoid the presence of endogeneity we used the number of trades of the following hour (t+1) to have zero overlap with the Twitter data. The independent variables included in our model are the three proposed dimensions of social broadcasting. Each of the three dimensions consists of several sub-constructs based on their thematic similarities explained in Section 3. We

aggregated the sub-constructs by applying factor analyses. Therefore, we used a principal component analysis with orthogonal varimax rotation in Stata to generate one factor for each characteristic (Creswell, 2013). Orthogonal factors are used to generate variables without inter-correlated components. The sub-constructs of the three dimensions are measured as follows.

To assess the user dimension, we measured the sub-construct expertise by using a regular expression to extract the usernames mentioned from each message and counted how often users are mentioned on average within the period of analysis. The second sub-construct of the user dimension, popularity, is calculated based on the average number of followers of each user who broadcasts stock-related messages. The sub-constructs of the message dimension reflect the messages' information richness. We measured the sub-constructs user mentions, tags, and external references by calculating how often these features appear on average within the messages. The sub-construct, financial sentiment, is computed based on the results of our sentiment analysis by calculating the average sentiment score of the messages. The discussion dimension's sub-constructs are based on the activity within the social broadcasting network. First, we determined the share of messages that provide new information. Hence, to calculate interest we divided the number of retweets by the total number of messages regarding a stock. The message reach reflects how many users within the social broadcasting network receive the stock-information. Therefore, we used, in difference to popularity, the total number of followers of all information contributors. As the constructs popularity and reach make both use of the follower feature, we checked for collinearity of the constructs, but found no evidence for collinearity. The volume is computed by dividing the amount of messages concerning a stock by the number of all messages sent in an hour regarding all stocks of the final dataset. Therefore, we get the amount of messages in relation to a baseline and account for the increasing usage of Twitter to share stock-related information. Finally, we applied three factor analyses to separately aggregate the sub-constructs to each of the three dimensions. The factor loadings are presented in Table 1.

Factor Loadings				
Dimension	**Sub-Construct**	**Users**	**Messages**	**Discussion**
User	Expertise	0.595	-	-
	Popularity	0.593	-	-
Message	Mentions	-	0.508	-
	Tags	-	-0.185	-
	External References	-	-0.533	-
	Financial Sentiment	-	0.421	-
Discussion	Interest	-	-	0.327
	Reach	-	-	0.467
	Volume	-	-	0.484
Method of extraction: principle component analysis, Rotation: orthogonal varimax rotation				

Table 1.   Results of the factor analysis.

To account for additional factors influencing stock market developments, we included a robust set of control variables identified in previous research. Earnings announcements have been found to increase the volume of Twitter messages significantly as interest in these stocks increases (Mao et al., 2013). Especially unexpected earning results can have a large impact on the trading activity (Bamber, 1986). Hence, we controlled for days at which earnings announcements are released using a dummy variable. Additionally, stock performance measures including price volatility, stock return, and price level are used as control variables regarding the hour in which the tweets are sent and in which the stocks are traded. We also control for each trading hour and weekday using indicator variables.

4.4 Panel Analysis

The final dataset contains information about the stocks of 97 companies at 574 points in time. Hence, our dataset consists of a time series dimension and cross-sectional dimension. To address both dimensions in one regression model we calculate our model by using a panel analysis (Hamilton, 1994). Additional to the factors that we include as control variables, a panel analysis accounts for factors that are unobserved (Wooldridge, 2009). In this regard, a panel analysis accounts for effects that are company specific and do not change in the period of analysis like industry sector or company size. To account for effects that change over time but affect all companies like national policies or federal regulations we include the trading day and the trading hour as indicator variables. Therefore, a panel analysis accounts for individual heterogeneity (Wooldridge, 2011). To test our hypotheses we combined the hourly intraday data of each stock with the social broadcasting data about the stock of the preceding hour. There are missing values in our dataset because not every company's stock is mentioned in every hour. Thus, regarding the 97 stocks in our sample, we have 42,504 observations with a minimum of 275 observations and a maximum of 574 observations per stock.

DV: Trading Activity$_{t+1}$				
Variables	**Coef.**	**t**	**P**	**SE**
Constant	606.14	26.34	0.0	23.02
Users$_t$	-4.68	-1.30	0.19	3.59
Messages$_t$	9.26	2.41	0.02**	3.84
Discussion$_t$	74.16	13.84	0.0***	5.36
Price Level$_t$	-1.35	-1.88	0.06*	0.72
Price Level$_{t+1}$	0.14	0.2	0.841	0.72
Price Volatility$_t$	25793.38	33.72	0.0***	764.95
Price Volatility$_{t+1}$	-25154.69	-32.69	0.0***	769.47
Stock Return$_t$	-9669.86	-11.38	0.0***	850
Stock Return$_{t+1}$	-10255.97	-12.21	0.0***	839.83
Earnings Releases	21.58	0.54	0.59	40.06
$F_{20,42387} = 639.31$, $p < 0.01$***, R^2-within = 0.23				
Number of Observations = 42,504, Number of Groups = 97				
P-Value: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ (two-tailed significance test)				
DV = Dependent Variable, Coef. = Regression Coefficient, SE = Standard error, p = P-Value, R^2-within = percentage of explained variance in DV through changes within predictors				
The indicator variables trading day and trading hour are not displayed in the table for the sake of clarity.				

Table 1.   Results of the fixed effects panel analysis.

First, we checked for multicollineartiy using a collinearity analysis in Stata. Variance inflation factor are below 5 and the mean of the variance inflation factors is 1.6 (Belsley, 1991). Thus, there is no evidence for multicollinearity found. Additionally, we performed a Wooldridge test for autocorrelation to account for serial correlation in the idiosyncratic errors (Wooldridge, 2011). The result of the test implies that we have to reject the null hypothesis that there is no first order autocorrelation in our data. Therefore, to account for autocorrelation, we estimate our model using a fixed-effects panel analysis with a first order autocorrelation disturbance (Baltagi and Wu, 1999). The results of our panel analysis are presented in Table 2. They confirm a significant influence of the message dimension ($p < 0.05$) as well as the discussion dimension ($p < 0.01$) of social broadcasting on stock market trading activity of the following hour. The coefficients of message and discussion dimensions are both positive indicating the positive influence of both dimensions on trading activity. However, the user dimension is not

found to be significant ($p > 0.1$) and its coefficient is negative. Hence, we can confirm that our results support hypothesis 2 and hypothesis 3. The included control variables are almost all highly significant ($p < 0.01$) with exception of the variable earning releases and the price level at t+1, which are above the threshold of $p < 0.1$. The R^2-within of our panel analysis is 0.23 and shows the explained variance regarding each stock. The F-test of the panel analysis, whether all coefficients in the model are different from zero, is confirmed by a high significance level of the model ($p < 0.01$).

4.5   Discussion of the Results

Based on the results of our empirical analysis, we found that the message and discussion dimension significantly influence the succeeding hour's trading activity. Hence, the information richness of a social broadcasting message is an important predictor for trading activity on stock markets. Broadcasting a richer set of information supports investors in making their trading decisions leading to an increased number of trades. The factor loadings of the message dimension show that the sub-constructs mentions and financial sentiment positively influence trading activity. However, messages containing tags and external references have a negative influence on trading activity. An explanation for this could be that although tagging is a major feature of Twitter, the use of cashtags limits the benefit of additional tags to follow the discussion of stocks. Moreover, adding tags in addition to cashtags reduces the space available for information since Twitter messages are limited to 140 characters. According to the efficient market hypothesis, stock prices should fully reflect all available information (Fama, 1970). External references in tweets contain often links to professional websites like newspaper which content should already been processed. Thus, we argue that the negative impact of external references might be due to the fact that tweets containing external references represent a summary of already processed information from websites. Moreover, tweets containing external references are sometimes composed in a vague way to increase the likelihood that the user will follow the reference to present doubtful investment advice such as advertising penny stocks. Regarding the discussion dimension, the factor analysis reveals that volume and the reach of the messages are important factors to draw conclusions about future trading activity. The interest regarding stocks expressed by the share of retweets has a lower factor loading and thus a lower influence on trading activity. This can be explained by the timespan in which new messages are retweeted. Hence, retweets may occur when new information is already processed at the stock markets and will not affect trading activity to the same extent as the other sub-constructs of the discussion dimension.

Contradictory to what we expected is that the user dimension of social broadcasting has no significant influence on trading activity and even has a negative coefficient. Although, social media highly focuses on empowering the users to engage in relationship building and content provision, the sub-constructs of the user dimension are not influential on stock market trading activity. An explanation for this result may be based on how relationships are built in social broadcasting networks. On Twitter, it is possible to engage in unidirectional relationships and information sending and receiving are the main purpose of the network. In contrast to that, the essential element of social networking platforms is the creation of profiles that may include personal or work-related information and the revealing of individual, bidirectional relationships to other users. Thus, building and revealing relationships are more important on social networking platforms than on social broadcasting networks. Furthermore, even users with rather few relationships can provide meaningful information that can be integrated in decision-making. Hence, it is more important what is written than who is sending the information.

5   Conclusion

The aim of our study was to analyze how the user, message, and discussion dimensions of social broadcasting influence trading activity on stock markets. Therefore, we aggregated constructs from prior research to the superior dimensions of social broadcasting to conduct a comprehensive analysis.

5.1   Implications for Theory and Practice

The primary theoretical contribution of our work is the developed conceptual model to investigate the influential dimensions of social broadcasting on trading activity at stock markets. Specifically, we describe three dimensions (user, message, and discussion) of a social broadcasting network and the underlying sub-constructs by building on existing literature on social media and financial markets. Therefore, we show how to investigate activity in social broadcasting networks as a whole rather than concentrating on single aspects like followers or message sentiment. This leads to the finding that information about who is broadcasting information is less important than the broadcasted content itself in a financial context. The developed model is not limited to the application in the financial sector and can help to understand the impact of social media platforms on economic outcomes in general. Furthermore, we contribute to the still emerging literature on social broadcasting and financial markets with our research model and the intraday analysis. Based on the empirical study we provide a panel analysis with preceding factor analyses to aggregate the sub-construct to the proposed dimensions.

The results of our study have several implications for practice to better understand social broadcasting activity in relation to trading activity on stock markets. Investors are interested in using insights from social broadcasting as decision support. Based on our comprehensive research model, we provide evidence that it is more valuable to focus in general on the content of messages rather than on the characteristics of the authors. Therefore, the informative content of broadcasted messages should be implemented in decision support systems for trading decisions. Additionally, information about how the discussion about stocks evolves on social broadcasting platforms should be integrated into decision-making. Furthermore, we illustrate in our empirical study how to extract social broadcasting data and how to process it subsequently to measure the dimensions by the proposed sub-constructs. Thereby, it is important to apply intelligent filters to separate valuable information from spam to reduce the risk of undervaluing important information due to information overflow.

5.2   Limitations and Future Research

Financial markets are exposed to a variety of influential factors that lead to volatility at stock markets. By using a panel analysis, we can neglect the influence of unobserved factors that affect the stocks of all companies equally (like changes in monetary policy or economic outlook). However, there are further individual factors that have an influence on the trading activity on stocks that remain unobserved (e.g., ad-hoc notifications of a company). Furthermore, although we used a systematic way to select the companies for our empirical study based on well-known stock index, it results in a selection-bias of large companies. Companies of medium and small size might change the results in some way. Therefore, future research should examine whether the results hold true when the dataset contains companies of different sizes. However, the amount of stock-related messages decreases significantly the smaller the size of a company gets. Another limitation is based on how we collected the data. We used the cashtag notation of Twitter to collect only stock-related messages. However, there are messages not containing cashtags which affects trading decisions. Future research, should evaluate the integration of such message which requires large effort to filter the relevant information.

Besides the future research directions emerging based on the limitations of our work, we think that it is especially of interest to understand the impact of individual messages and users on the development in stock markets. There are plenty of examples of single tweets that caused immediate reactions on stock markets such as the fake tweet of Associated Press that have not been investigated yet. Although, we used hourly intraday stock market data, we believe that our data would not be granular enough to address this research problem as stock markets adapt rapidly within minutes to new information. Moreover, it could be of interest to compare how our results change by group comparison based on the industry sector of the companies. This could shed more light on the results as we assume that public interest in social broadcasting networks may be sector dependent.

References

Ahuja, M. K., Galletta, D. F. and Carley, K. M. (2003). Individual Centrality and Performance in Virtual R&D Groups: An Empirical Study. *Management Science* 49 (1), 21-38.

Antweiler, W. and Frank, M. Z. (2004). Is all that Talk just Noise? The Information Content of Internet Stock Message Boards. *The Journal of Finance* 59 (3), 1259-1294.

Bakshy, E., Hofman, J. M., Mason, W. A. and Watts, D. J. (2011). Everyone's an Influencer: Quantifying Influence on Twitter. In: *Proceedings of the 4th ACM International Conference on Web Search and Data Mining.* Hong Kong: China.

Baltagi, B. H. and Wu, P. X. (1999). Unequally spaced Panel Data Regressions with AR (1) Disturbances. *Econometric Theory* 15 (6), 814-823.

Bamber, L. S. (1986). The Information Content of Annual Earnings Releases: A Trading Volume Approach. *Journal of Accounting Research* 24 (1), 40-56.

Belsley, D.A. (1991). *Conditioning Diagnostics. Collinearity and weak Data in Regression*. New York, NY: Wiley.

Bloomberg (2013). Fake Post erasing $136 Billion shows Markets need Humans. Available online at http://www.bloomberg.com/news/2013-04-23/fake-report-erasing-136-billion-shows-market-s-fragility.html, checked on 10/16/2014.

Bollen, J., Mao, H. and Zeng, X. (2011). Twitter Mood predicts the Stock Market. *Journal of Computational Science* 2 (1), 1-8.

Celli, F. and Rossi, L. (2012). The Role of Emotional Stability in Twitter Conversations. In: *Proceedings of the Workshop on Semantic Analysis in Social Media.* Avignon: France, 10-17.

Cha, M., Haddadi, H., Benevenuto, F. and Gummadi, P. K. (2010). Measuring User Influence in Twitter: The Million Follower Fallacy. In: *Proceedings of the 4th International AAAI Conference on Weblogs and Social Media.* Washington: United States, 10-17.

Constant, D., Sproull, L. and Kiesler, S. (1996). The Kindness of Strangers: The Usefulness of Electronic Weak Ties for Technical Advice. *Organization Science* 7 (2), 119-135.

Creswell, J.W. (2013). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. 4th Edition. Thousand Oaks, CA: Sage Publications.

Dellarocas, C. (2003). The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms. *Management Science* 49 (10), 1407-1424.

Fama, E. F. (1965). Random Walks in Stock Market Prices. *Financial Analysts Journal* 21 (5), 55-59.

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance* 25 (2), 383-417.

Fang, L. and Peress, J. (2009). Media Coverage and the Cross-section of Stock Returns. *The Journal of Finance* 64 (5), 2023-2052.

Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases. *AI Magazine* 17 (3), 37-54.

Fortune (2013). Tweet Retreat: Did High-Frequency Reading crash the Market? - Fortune. Available online at http://fortune.com/2013/04/25/tweet-retreat-did-high-frequency-reading-crash-the-market/, checked on 3/21/2015.

General Inquirer (2014). Inquirer Home Page. Available online at http://www.wjh.harvard.edu/~inquirer, checked on 3/23/2015.

Geva, T. and Zahavi, J. (2014). Empirical Evaluation of an Automated Intraday Stock Recommendation System incorporating both Market Data and Textual News. *Decision Support Systems* 57 (January), 212-223.

Goh, K.-Y., Heng, C.-S. and Lin, Z. (2013). Social media brand community and consumer behavior: Quantifying the relative impact of user- and marketer-generated content. *Information Systems Research* 24 (1), 88-107.

Granovetter, M. (1973). The Strength of Weak Ties. *American Journal of Sociology* 78 (6), 1360-1380.

Granovetter, M. (2005). The Impact of Social Structure on Economic Outcomes. *The Journal of Economic Perspectives* 19 (1), 33-50.

Greenfield, D. (2014). Social Media in Financial Markets: The Coming of Age. Available online at http://gnip.com/pages/social-media-and-the-markets-the-coming-of-age-whitepaper, checked on 3/23/2015.

Grossman, S. J. and Stiglitz, J. E. (1980). On the Impossibility of Informationally Efficient Markets. *The American Economic Review* 70 (3), 393-408.

Hamilton, J.D. (1994). *Time Series Analysis*. Princeton, NJ: Princeton University Press.

Hu, X. and Liu, H. (2012). Text Analytics in Social Media. In Aggarwal, C. C. and Zhain, C. (Eds.). In: *Mining Text Data*. New York, NY: Springer, 385-414.

Jones, C., Hesterly, W. S. and Borgatti, S. P. (1997). A General Theory of Network Governance: Exchange Conditions and Social Mechanisms. *The Academy of Management Review* 22 (4), 911-945.

Kane, G. C., Alavi, M., Labianca, G. and Borgatti, S. P. (2014). What's Different about Social Media Networks? A Framework and Research Agenda. *Management Information Systems Quarterly* 38 (1), 275-304.

Kaplan, A. M. and Haenlein, M. (2010). Users of the World, unite! The Challenges and Opportunities of Social Media. *Business Horizons* 53 (1), 59-68.

Larson, J. R., Christensen, C., Abbott, A. S. and Franz, T. M. (1996). Diagnosing Groups: Charting the Flow of Information in Medical Decision-Making Teams. *Journal of Personality and Social Psychology* 71 (2), 315-330.

Li, Q., Wang, T., Li, P., Liu, L., Gong, Q. and Chen, Y. (2014). The Effect of News and Public Mood on stock Movements. *Information Sciences* 278, 826-840.

Liu, B. (2010). Sentiment Analysis and Subjectivity. In Indurkhya, D. (Ed.). In: *Handbook of Natural Language Processing*, 627-666.

Liu, Z., Liu, L. and Li, H. (2012). Determinants of information retweeting in microblogging. *Internet research* 22 (4), 443-466.

Lo, A. W. (2004). The Adaptive Markets Hypothesis. *The Journal of Portfolio Management* 30 (5), 15-29.

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), 35-65.

Lu, X., Yu, Z., Guo, B. and Zhou, X. (2014). Predicting the Content dissemination Trends by Repost Behavior modeling in Mobile Social Networks. *Journal of Network and Computer Applications* 42, 197-207.

Mao, Y., Wei, W. and Wang, B. (2013). Twitter Volume Spikes: Analysis and Application in Stock Trading. In: *Proceedings of the 7th Workshop on Social Network Mining and Analysis.* Chicago: United States.

Mohamed, A. A., Gardner, W. L. and Paolillo, J. G. (1999). A Taxonomy of Organizational Impression Management Tactics. *Advances in Competitiveness Research* 7 (1), 108-130.

Nann, S., Krauss, J. and Schoder, D. (2013). Predictive Analytics on Public Data. The Case of Stock Markets. In: *Proceedings of the 21st European Conference on Information Systems.* Utrecht: Netherlands.

Nevo, D., Benbasat, I. and Wand, Y. (2012). The Knowledge Demands of Expertise Seekers in two Different Contexts: Knowledge Allocation versus Knowledge Retrieval. *Decision Support Systems* 53 (3), 482-489.

Oh, C. and Sheng, O. (2011). Investigating Predictive Power of Stock Micro Blog Sentiment in Forecasting Future Stock Price Directional Movement. In: *Proceedings of the 32nd International Conference on Information Systems.* Shanghai: China.

Papaioannou, P., Russo, L., Papaioannou, G. and Siettos, C. I. (2013). Can Social Microblogging be used to Forecast Intraday Exchange Rates? *NETNOMICS: Economic Research and Electronic Networking* 14 (1-2), 47-68.

Preis, T., Moat, H. S. and Stanley, H. E. (2013). Quantifying Trading Behavior in Financial Markets using Google Trends. *Scientific Reports* 3.

Rindova, V. P., Williamson, I. O., Petkova, A. P. and Sever, J. M. (2005). Being Good or being Known. An Empirical Examination of the Dimensions, Antecedents, and Consequences of Organizational Reputation. *Academy of Management Journal* 48 (6), 1033-1049.

Rui, H., Liu, Y. and Whinston, A. (2010). Chatter Matters: How Twitter can open the Black Box of Online Word-of-Mouth. In: *Proceedings of the 31st International Conference on Information Systems.* Saint Louis: United States.

Rui, H., Liu, Y. and Whinston, A. (2013). Whose and What Chatter matters? The Effect of Tweets on Movie Sales. *Decision Support Systems* 55 (4), 863-870.

Rui, H. and Whinston, A. (2012a). Designing a Social-Broadcasting-based Business Intelligence System. *ACM Transactions on Management Information Systems* 2 (4), 1-19.

Rui, H. and Whinston, A. (2012b). Information or Attention? An Empirical Study of User Contribution on Twitter. *Information Systems and e-Business Management* 10 (3), 309-324.

Shi, Z., Rui, H. and Whinston, A. B. (2014). Content Sharing in a Social Broadcasting Environment: Evidence from Twitter. *Management Information Systems Quarterly* 38 (1), 123-142.

Shiller, R. J. (2003). From Efficient Markets Theory to Behavioral Finance. *The Journal of Economic Perspectives* 17 (1), 83-104.

Siering, M. (2013). All Pump, no Dump? The Impact of Internet Deception on Stock Markets. In: *Proceedings of the 21st European Conference on Information Systems.* Utrecht: Netherlands.

Signorini, A., Segre, A. M. and Polgreen, P. M. (2011). The Use of Twitter to Track Levels of Disease Activity and Public Concern in the US during the Influenza A H1N1 Pandemic. *PloS ONE* 6 (5), 1-10.

Sprenger, T. O., Sandner, P. G., Tumasjan, A. and Welpe, I. M. (2014a). News or Noise? Using Twitter to Identify and Understand Company-specific News Flow. *Journal of Business Finance & Accounting* 41 (7-8), 791-830.

Sprenger, T. O., Tumasjan, A., Sandner, P. G. and Welpe, I. M. (2014b). Tweets and Trades: The Information Content of Stock Microblogs. *European Financial Management* 20 (5), 926-957.

Stieglitz, S. and Dang-Xuan, L. (2013). Emotions and Information Diffusion in Social Media – Sentiment of Microblogs and Sharing Behavior. *Journal of Management Information Systems* 29 (4), 217-248.

Suh, B., Hong, L., Pirolli, P. and Chi, E. H. (2010). Want to be Retweeted? Large Scale Analytics on Factors impacting Retweet in Twitter Network. In: *Proceedings of the 2nd International Conference on Social Computing.* Minneapolis: United States, 177-184.

Sul, H. K., Dennis, A. R. and Yuan, L. (2014). Trading on Twitter: The Financial Information Content of Emotion in Social Media. In: *Proceeding of the 47th Hawaii International Conference on System Sciences.* Hawaii: United States.

Thelwall, M., Buckley, K. and Paltoglou, G. (2012). Sentiment Strength Detection for the Social Web. *Journal of the American Society for Information Science and Technology* 63 (1), 163-173.

Thelwall, M., Buckley, K., Paltoglou, G., Cai, D. and Kappas, A. (2010). Sentiment Strength Detection in Short Informal Text. *Journal of the American Society for Information Science and Technology* 61 (12), 2544-2558.

Tirunillai, S. and Tellis, G. J. (2012). Does Chatter really Matter? Dynamics of User-generated Content and Stock Performance. *Marketing Science* 31 (2), 198-215.

Tumasjan, A., Sprenger, T. O., Sandner, P. G. and Welpe, I. M. (2010). Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. In: *Proceedings of the 4th International AAAI Conference on Weblogs and Social Media.* Washington: United States, 178-185.

Twitter (2015). About Twitter. Available online at https://about.twitter.com/company, checked on 3/18/2015.

Wasko, M. M. and Faraj, S. (2005). Why should I share? Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice. *Management Information Systems Quarterly* 29 (1), 35-57.

Wasko, M. M., Teigland, R. and Faraj, S. (2009). The Provision of Online Public Goods: Examining Social Structure in an Electronic Network of Practice. *Decision Support Systems* 47 (3), 254-265.

Weng, J., Lim, E.-P., Jiang, J. and He, Q. (2010). TwitterRank: Finding Topic-Sensitive Influential Twitterers. In: *Proceedings of the 3rd ACM International Conference on Web Search and Data Mining.* New York City: United States, 261-270.

Wooldridge, J.M. (2009). *Introductory Econometrics. A Modern Approach*. Mason, OH: South-Western.

Wooldridge, J.M. (2011). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Yahoo!Finance (2013). Carl Icahn's Multibillion-Dollar Tweet Boosts Apple Stock. Available online at http://finance.yahoo.com/blogs/the-exchange/carl-icahn-multibillion-dollar-tweet-boosts-apple-stock-205938760.html, checked on 3/21/2015.

Zhang, X., Fuehres, H. and Gloor, P. A. (2011). Predicting Stock Market Indicators through Twitter "I hope it is not as bad as I fear". *Procedia - Social and Behavioral Sciences* 26, 55-62.

Zheludev, I., Smith, R. and Aste, T. (2014). When can Social Media lead Financial Markets? *Scientific Reports* 4.