Erasmus University Rotterdam
Erasmus School of Economics
MSc Economics & Business
Master specialisation Financial Economics

The Impact of Socially Active Executives

Abstract:

In this thesis, I examine the effect of usage of Twitter by CEOs of companies in the S&P 1500 on the value of the firm by means of an event study. I look at the effect of different sentiments (e.g. positive, neutral and negative) and content (e.g. work-related and non-work related) on the abnormal return of a company. We gathered and categorized a significant amount of tweets and regressed the results by OLS. I find significant results for positive and negative tweets at the cumulative abnormal return. Moreover, work-related tweets have significant impact on the cumulative abnormal return. The combination of negative tweets and work-related tweets have more impact on the cumulative abnormal return than the two variables taken independently. It is important for companies to know that social media is an important measure to trigger investors to invest in your company. Therefore, this thesis contributes to the growing literature of efficient social media communication.

Keywords: Social Media; Corporate disclosure; Stock Prices; Text Analysis; Event Study

JEL Classification: G12, G14, G38

Author: T.Lader
Student number: 356461
Thesis supervisor: S. Xia
Finish Date: 08-2017

Preface and Acknowledgements

The Erasmus University of Rotterdam and especially the Erasmus School of Economics provided me a lot of new useful insights for my future career and personal development. I am grateful and lucky that I was part of their willingness to educate and appreciate their help in many ways. Therefore, I would like to thank all the professors, teachers and other persons who were part of this.

Second, I would like to thank my supervisor, Shuo Xia, for helping me during the months when I worked on my Master thesis. Without his guidance and wise advices, I would not be able to complete the Master thesis in the way I finished it now. I also want to thank my girlfriend and friends for the support they gave during my thesis. They covered my back in difficult times, which often gave me extra motivation and encouraged me to finish the thesis. Last but not least, a special thanks to my parents. They always believed in me during my years at the Erasmus University of Rotterdam. This was not always an easy path and therefore I would like to dedicate my Master thesis to them. All in all, I can say that this thesis can be considered as a gift to them for their encouragements over the years.

Table of Contents

1. Introduction

Communication is as old as humanity exists and very important within a society. In the past, cavemen already spoke to each other in order to make themselves clear and warning others from possible dangers. Over the last centuries, the way people communicate changed drastically and transitioned into a digital format on the internet. The contantly increasing use of Internet as a way to communicate and becoming a source of information triggered an increasing online activity (Ranco, Aleksovski, Caldarelli, Grcar, & Mozetic, 2015). Twitter, Facebook and LinkedIn are nowadays the most popular platforms to get in contact with each other.

As a consequence of the fact that social media is the most common way to communicate with each other, it therefore gets the most interest of companies. Information diffusion by means of social media is relatively fast and efficient and reaches a wide audience. Twitter, for example, has become the most popular social media platform to disseminate financial reports (Nguyen, Varghese, & Barker, 2013). The executives of companies historically communicated via press releases, SEC fillings and public conference calls whereafter the news got facilitated through analyst- and media coverage (Blankespoor, Miller, & White, 2014). By means of the introduction of, for instance, Twitter, executives bypassed these coverages and communicated directly to investors and customers (Chen, Hwang, & Liu, 2015). This has lead to a quicker and more efficient way of responding to dispersed information. The introduction of Twitter as an efficient way to communicate with investors also has economic consequences with regard to the firm value.

This thesis investigates this impact of having a socially active CEO on the value of the firm. By means of investigating the stock price returns of the company, one aims to find correlations between the activity on Twitter, the content of the tweets and the sentiment represented within the tweets on the one hand and consequently the change in stock price returns on the other hand. By aggregating the stock price against the amount of outstanding stocks, one should be able to calculate the market capitalization of the firm and therefore the market value of the firm.

The thesis provides an analysis in order to give an answer to the following main research question:

''To what extent does the usage of Twitter by socially active CEOs have an impact on the value of a firm?''

The aim of this paper is to observe the economic consequences a firm might experience due to Twitter utilisation by the CEO. Due to the emergence of social media in the last years, executives have the opportunity to release information in the most efficient way. The biggest benefit of that, is the possibility that investors get a clear, quick and transparant insight into the company. As a consequence, investors react immediately and in the right way (e.g. buy, hold or sell) towards announced news which is facilitated by the firm. This means that when an executive releases information, investors exactly know when and how much to buy or sell in order to rebalance their portfolio correctly.

Another benefit contains the desire to reach out to investors and customers and increase their commitment to the firm (Chen, Hwang, & Liu, 2015). This information diffusion was different in the past when it took a detour by means of media coverage and press conferences (Chen, Hwang, & Liu, 2015). As a consequence, investors did not always handled in their most optimal way.

In order to find an answer to the research question, I observe the personal Twitter accounts of CEOs, who are listed with their firm on the S&P 1500 in the period of 2008 till 2015. Subsequently, I divide all the tweets sended by the CEOs into different sentiments and categories. Next to it, I gather financial information of stock prices, market returns, outstanding shares and the book values of the firms. Other relevant variables are the NAICS codes (to control for industry fixed effects) and the 4-factor model of Carhart (1997) to control for the calculation of the normal return. By means of T-tests (to test if the abnormal return is statistically significant to zero) and OLS regressions, I try to find an correlation between the sentiment and content of tweets and the corresponding stock price returns.

To preview our results, I observe that there are, on average, significant results for positive sentiments and negative sentiments with regard to the cumulative abnormal return for the different event windows.

I find significant positive results for a positive sentiment of 0.26% and significant negative results of -0.26% for a negative sentiment. Moreover, work-related tweets also seem to have significant impact on the cumulative abnormal returns. The sign of the impact changes from positive (0.12%) to negative (-0.22%) as the event windows gets broader. The combination of negative tweets with work-related tweets are statistically significant and has a larger impact on the cumulative abnormal return than negative tweets and work-related tweets summed up independently. Therefore, I can conclude that there is a significant impact of usage of Twitter by CEOs of specific variables on the market value of the firm, which is based on the stock price of a company.

This research can be considered as a contribution to the prior growing literature about the role of social media in financial markets. Other research did not had such effects of social media related to firm value and individual stock prices in scope, but rather to different economic aspects within a firm (e.g. bid-ask spread, trading activitiy and sales-over-assets) or stock market indices Das & Chen (2007). This thesis can also be important for firms, because it provides evidence of the effect of social media and its corresponding economic consequences on the value of the firm. Current literature within financial markets focuses on stock prices and entails the question if stock prices fully reflect the fundamental value of a firm (Sloan, 1996). Eventually, the analysis shows a confirmation to the continuous evolvement of communication and interaction within the financial market.

The rest of the paper has the following structure. The next section shows a thorough description of the prior literature. This relates to prior research of Twitter with regard to papers about the history of corporate disclosure, the content of tweets and other results of the effect of social media on stock prices. Section three discusses the data and methodology. It provides an overview of the data collection and the methodologies used in this research in order to answer the research question and hypotheses. Section four presents the results and main findings and answers the hypotheses. Section five provides a conclusion with an answer of the research question by means of the hypotheses. At last, the conclusion contains some limitations and drawbacks of this research on the one hand and recommendations to improve further research on the other hand.

2. Literature Review

This section describes the previous literature and research that has been conducted prior to this thesis. In this section, topics concerning corporate disclosure, twitter content, twitter sentiment and social media are frequently discussed by various papers. It also highlights the different constructive hypotheses, which will be briefly spoken and answered.

In the very beginning, researchers as Core (2001), Fields, Lys, & Vincent (2001), Lewellen, Park, & Ro (1996), Beyer, Cohen, Lys, & Walther (2010) and Healy & Palepu (2001) started to write about topics related to corporate disclosure. All these authors had, in their own way, a contribution to the start of the distinction between voluntary disclosure (management discussion and analysis) and mandatory disclosure (SEC filings). According to these authors, voluntary disclosure gives management the opportunity to voluntarily discuss the company's performance and future prospects. On the other hand, mandatory disclosure is an obligation and contains the earnings announcements and SEC filings, which return every single year.

Healy & Palepu (2001) started with a framework in order to analyze the reports of managers and look for the decisions to disclose in capital markets and consequently came up with well-formulated research questions for further research. They provided a broad overview where they asked big-picture questions about accounting information. Subsequently, Core (2001) reviewed the paper of Healy & Palepu (2001) and provided additional analyzes with regard to the empirical voluntary disclosure. Core (2001) focused on the voluntary disclosure, because it offered the greatest opportunity for large increases in the understanding of the role of accounting information in firm valuation and corporate finance. Healy & Palepu (2001) did not focused on the economic theory of voluntary disclosure and therefore Core (2001) attempted to complement their research by using a specific framework based on the economic theory of the firm.

In addition, Beyer, Cohen, Lys & Walther (2010) exposed certain conditions under which firms voluntarily disclose all their private information. These conditions include, for example; 1) the fact that these disclosures are costless; 2) investors know that firms have private information;

3) all investors interpret the firms' disclosure in the same way and firms know how investors will interpret that disclosure; 4) managers want to maximize their firms' share prices; 5) firms can credibly disclose their private information; and 6) firms can not commit ex-ante to a specific disclosure policy (Beyer, Cohen, Lys, & Walther, 2010).

After that, firms came up with the idea to disclose corporate information on their official website since the widespread use of Internet at the end of last century (Zhou, Lei, Wang, Fan, & Wang, 2015). The quality of the behaviour of reporting with respect to firms on their websites varied across countries, industries and firms. This had to due with factors, such as size, profitability, industry and regulation (Debreceny, Gray, & Rahman, 2002).

Thereafter, recent studies show that, at this time, approximately twelve percent of the surveyed firms use social media effectively (Harvard Business Review Analytic Services, 2010). This shows that the usage of social media still acts in its infancy and has a lot of potential to be more effective towards stakeholders. Authors as Kaplan & Haenlein (2010) tend to help firms in using social media more effectively by means of discussing challenges and opportunities and providing handy advices.

In order to look for the most interesting form of social media, Culnan, McHugh, & Zubillaga (2010) adopted four different platforms in their research (e.g. Twitter, Facebook, blogs and client-hosted forums) and found out that Twitter has been used most frequently (i.e. 53 percent). That being said, they discovered some significant industry differences among the different platforms. Hence, Twitter turned out to be the most popular form of social media on the internet and can therefore potentially be considered as the most efficient way to inform investors and disclose corporate information. Since firms send messages to their followers, who are interested in the announcements of the firm, they simultaneously share information with a bigger audience, due to retweets (i.e. re-posting of tweets) of the followers. The ease of receiving and sending information through Twitter allows firms to reach a wider range of stakeholders on a timely basis compared with other platforms (Zhou, Lei, Wang, Fan, & Wang, 2015).

Harvard Business Review Analytic Services (2010) came up with other kind of announcements, which are disseminated by social media, and especially Twitter.

According to them, firms specifically use social media to disclose annual reports as well as financial and event information, such as earnings releases and market-moving news. A different study added subsequently that the dissemination through Twitter reduces the information asymmetry between firms and investors and adopts negative capital market consequences, such as product recalls (Blankespoor, Miller, & White, 2014).

Zhou, Lei, Wang, Fan & Wang (2015) showed in their paper an example of dissemination of Facebook. Netflix CEO Reed Hasting posted a Facebook message on his personal account and stated that the monthly online viewing amount on Netflix exceeded one billion hours. Consequently, the stock price of the firm increased by 6.2 percent at that particular day. Reed, simultaneously, proved the influence of social media, and in this case Facebook, at stock prices. These authors also stated that the SEC (Securities and Exchange Commission), which guids the website disclosures of companies, claimed to cover the usage of social media channels.

There are also some papers showing concerns towards the positive effects of social media. The papers of Culnan, McHugh, & Zubillaga (2010) and Nair (2011) adopt cautious views while they discussed the factors that firms must consider in measuring the value of social media to justify the relevant costs. The factors include; 1) A mindful adoption of the media; 2) Building of a community; and the 3) Absortive capacity. However, these two papers still had few acceptance, due to different studies and aspects of these papers, which contained other potential advantages of social media in improving the customer engagement (Kaplan & Haenlein, 2010), enhancing promotion mixes (Mangold & Faulds, 2009), detecting customer complaints regarding product defects (Abrahams, Jiao, Wang, & Fan, 2012), and increasing equity and business values (Culnan, McHugh, & Zubillaga, 2010).

In order to revert the discussion about the distinction between voluntary disclosure and mandatory disclosure, one considers the disclosure by means of Twitter as a management discussion and analysis (Chen, Hwang, & Liu, 2015). However, tweets from CEOs differ from this management discussion and analysis due to the frequency and content. According to Chen, Hwang & Liu (2015), the disclosure of management discussion and analysis normally occur once a year or even once in a quarter. At the same time, tweets from CEOs occur on average once every five days.

They stated that Twitter leads to an increased frequency for the reason that executives not only send tweets for firm-specific reasons, but sometimes also have private implications.

The content of tweets is a related topic that is frequently discussed in the literature preceding the research of Chen, Hwang & Liu (2015). One other group of authors, i.e. Naaman, Boase, & Lai (2010), focused at the different kind of categories in which content can be divided. They seperated four different categories: 1) information sharing content; 2) opinions and complaints; 3) statements and random thoughts; and 4) tweets about the current status of the CEO. According to these authors, tweets about the current status of the CEO comprehend the largest frequency (e.g. 40%). All these categories have a significant influence on firm aspects as trading activity, bid-ask spreads, sales over assets, the retail shareholder base and stock prices (Chen, Hwang, & Liu, 2015).

Other Twitter effects found in previous studies can be explained by rational and irrational reasons. In the paper of Barber & Odean (2008), tweets are considered to have entertainment value and value-relevant content. The entertainment value in tweets tends to increase the trading activity and retail investor presence, due to the raised attention of investors. According to Bagehot (1971), this increased amount of retail investors leads to an increase in participation of relatively uninformed investors, which consequently results in a lower spread.

A dissimilar view with respect to the content of tweets contains the perspective towards non-related firm operations. Actually, the personal life of a CEO has an effect on the working status and vice versa. This means that if CEOs tweet about the good mood they are having in their private situation, this probably would reflect into a good decision in their life as an executive. These sort of tweets intensify the information flow between traders and managers in an unusual way. A tweet, for example, about the private situation or the mood of the CEO, simultaneously, incorporate information about operations within the firm. The tweet is obviously not directly related to the firm's operations (Chen, Hwang, & Liu, 2015).

Tweets also deliver information for strategic purposes (Chen, Hwang, & Liu, 2015). A tweet, for instance, about the good time the CEO has with friends will not occur on the verge of a negative earnings announcement. Tweets from executives regarding the personal life, lead to a change in, for instance, the trading activity.

An example of a study concerning the personal life of a CEO (Bollen, Mao, & Zeng, 2010) explored the content of the tweets with regard to the mood state and showed that these mood states relate to the value of the Dow Jones Industrial Average (DJIA). A couple of sentiments are, according to their results, predictive of the DJIA closing values. An other paper, written by Zhang, Fuehres, & Gloor (2011), showed a simillar pattern with the relationship between hope and fear on the one hand and the Dow Jones, S&P 500 and NASDAQ on the other hand. In this paper, the authors came up with a result which indicated that the emotional level of tweets was significant related to all three aggregated indicators.

These previous studies also have some drawbacks, due to the main focus they have on firms that meet certain criteria, such as categorization in the Fortune 500 or focusing at a specific industry. An example of such a limitation contains, for instance, the usage of randomised subsamples of all available tweets of the Twitter message stream (Sprenger, Tumasjan, Sandner, & Welpe, 2014). According to these authors, the majority of the tweets may not be related to stock market topics and therefore it is hard to conclude if the stock-specific information contained in tweets indeed associates with the indicators. An other limitation contains the fact that studies only explore the relationship between aggregate sentiment measures and aggregate stock market indices. They do not focus on the performance of individual stocks and therefore one could not relate the information in stock-related messages with this performance. This is, to some extent, due to the paper of Das & Chen (2007), who found a relationship between aggregated sentiment and the index returns to be much stronger than the correlation in the performance regarding individual stocks. The last drawback contains the result that not any single paper investigates the mechanism underlying the link between social media message sentiment and market prices (Sprenger, Tumasjan, Sandner, & Welpe, 2014).

This thesis, however, will provide a broader view towards the usage of social media with regard to corporate disclosure and the changes in the individual stock market prices. The findings have a more objective approach and are not influenced by specific criteria or exclude certain industries. This thesis collects a large group of CEOs and firms in different kind of industries and with various types of persons, which consequently gives a wider and more realistic sample. In order to find an answer to the research question, one firstly should formulate hypotheses to gather different parts of the final result:

Hypothesis 1: There is significant impact of work-related tweets at the stock price return in the following days.

In order to answer the first hypothesis, I made a distinction between tweets with a work-related implication and tweets with a non-work-related or personal implication as created by Chen, Hwang, & Liu (2015). Tweets which are having a work-related content are tweets which disseminate firm specific news, such as an announcement of a new product or a disappointing unexpected growth. Tweets which are having a non-work-related implication are tweets which disseminate the private situation of a CEO, such as a lunch break or a tweet about a sports game. Both these situations affect the stock price in a way which has been mentioned by Maniatis (2011). According to Maniatis (2011), the stock's closing price is a random walk and the best forecast for tomorrow's price, in a random walk, is today's price. The stock price is based on fundamentals, where the future cash flow is one example (Heaton & Lucas, 2000). It is plausible to say that future cash flows will reflect into the price today. A prediction or announcement about the firm which affects the future cash flows will therefore reflect in a change in today's stock price. This will end up in a positive or negative change, dependent on the type of announcement.

Hypothesis 2: There is significant impact of sentiment within tweets and the stock price return in the following days.

The second hypothesis deals with the mood of the content of the tweet, whether it is work-related or non-work-related. According to the theory of Chen, Hwang, & Liu (2015), tweets have a positive or negative mood. In addition, they showed that the mood a CEO has in his personal situation also will reflect in the firm specific decisions and announcements of the firm. Therefore, it would be arguable to say that positive moods always result in tweets with a thoroughgoing work-related implication. This means that positive tweets, irrespective if its work-related or non-work-related, always have an impact on the firm and subsequently change the stock price postively.

Hypothesis 3: The combination of a tweet with a positive sentiment and work-relevancy has a significantly larger impact on the stock return in the following days than both aspects summed up individually

The third hypothesis has to do with synergy. Synergy also occurs in mergers and acquisitions where the economic value of the newly merged company is higher than the sum of the economic value of the previous companies apart (Steiner, 1975). This occurrence possibly can also be the case when CEOs combine two different indicators of stock price return changes and gather a better result than both indicators summed up individually. This could be, for instance, the case for a positive sentiment combined with a work-relevant tweet. If this could be the case, CEOs can send more tweets with specific combinations that reach better and more often the range of investors, who subsequently are able to spread this information more often to users of Twitter. According to Ng & Wu (2006), Mizrach & Weerts (2009) and Hong, Kubik, & Stein (2005), the influence of word-of-mouth with respect to investors is very high. These investors have a desire to follow the advice of other investors and therefore react in the same way. one assumes that if more investors would be informed about a particular announcement, the bigger the impact would be on the price of the stocks. There are more investors who possibly react to the information and therefore correspond in the same specific way (e.g. buy, hold or sell).

3. Data

This part describes the different data that is used in this thesis and the way this data has been achieved. For instance, the CEO data, financial data and twitter data will briefly be discussed in this part. Firstly, it highlights regular information about Twitter in order to introduce the reader with this social platform. Further on, it describes all the other retrieved data (e.g. CEO date, financial data, twitter data and control variables data).

3.1 Twitter

Twitter was founded in 2006 and became part of a new online community, which is called social media. Twitter enables people to send messages to their own network of people with a maximum of 140 characters. These messages are called microblogs or ‘tweets’ and are sended to the ‘followers’. Followers are able to follow every public account they want, without having the consent of the person concerned. Registered users of Twitter are able to send messages and read the tweets of people they follow in their own timeline, but also have the opportunity to forward (e.g. ‘retweet’) tweets from people they follow. An other characteristic of Twitter includes the ‘hash-tag’ function. This function helps people to read certain topics, in which they are interested and wants to keep track on. It contains a word or message which is associated with a topic. Over the years, Twitter became one of the most popular social media websites in the United States with more than 300 million users and still keeps innovating in different ways.

3.2 CEO Data

In order to collect all the data, I first created a sample of socially active CEOs on Twitter. I downloaded a list, by means of Execucomp at WRDS, of all CEOs in the S&P 1500 in the period of 2008-2015. This time period is chosen because of the lack of more years within WRDS at the time of retrieval, which is caused by the relatively late emergence of Twitter in 2006 and the usage of Twitter by CEOs in 2008. The thesis uses the S&P 1500, due to the large variety and the huge amount of companies that are listed within this index (e.g. 90% of the market capitalization in the U.S.) and therefore having the highest and most divers amount of socially active CEOs.

This sample does not only contain CEOs of active companies, but also encompass companies which were once part of this index and tweeted during their time of operating. The full list of CEOs from Execucomp includes 2177 unique CEOs over approximately 2210 unique companies.

It is important to validate the accounts on their uniqueness and authenticity to prevent the research from biasses and hiatuses. Therefore, I verified the CEOs by their first name, middle name and last name and checked if there was a connection between the content of the tweets and the relation to the company of interest. I also checked for a correlation between characteristics of the Twitter profile (e.g. gender, function and company information) and the characteristics that used to belong to the CEOs. The Twitter accounts, which were in my opinion fake or contained an unsufficient amount of tweets (e.g. zero), were left out of the thesis and are therefore ignored in this research. In the end, this thesis contains 64 socially active CEOs divided over 60 unique companies. This means that some firms contain more than one socially active CEO. If the CEO got replaced by an other CEO of the same firm, a gap of one year is applied in the Twitter analysis in order to reformulate the effect to the stock return by the new CEO. The tweets from the CEOs are analysed during their period of leadership.

For each CEO, I gathered numerous data and made an overview, which is represented in Table 1. This data includes; 1) The Twitter alias of the CEO; 2) The first, middle and last name of the CEO; 3) The name of the company; 4) The date Twitter is activated by the CEO; 5) The total amount of tweets sended by the CEO; and 6) The total amount of followers as of May 2016. With respect to the text analysis of the tweets, I gathered data, such as; 1) The Twitter name of the CEO; 2) The first, middle and last name of the CEO; 3) The name of the company; 4) The date of the tweet sended; 5) The content of the tweet; and 6) The type of categorization of the tweet. The activations of the different accounts are widely spread over the total sample and are not related to a specific event.

3.3 Financial Data

The daily stock prices of the firms are collected from CRSP at WRDS and are related to the corresponding company in the period of 01/01/2008 till 31/12/2015.

There are a bunch of companies, that ended up in default or started tweeting later on and therefore were not available in every year of the time sample.These companies will be analysed during their own specific time of operating. Table 4 presents an overview of all the tickers of the companies.

First, in order to calculate the returns, daily stock prices are needed and should be adjusted with a factor in order to cope with dividend returns and stock splits. Therefore, this return includes dividend, due to the fact that dividend deduction is pre-settled. The thesis uses normal returns instead of log-returns in order to be coherent with previously written papers, such as Campbell, Lo, & Mackinlay (1997) and MacKinlay (1997), who wrote about event studies. These stock returns are value-weighted and are used for the analysis concerning the sentiment and the value relevancy. The corresponding formula for the return encompass the difference between the closing values of the stock price of day T and day T-1 divided by the closing value of T-1. Moreover, dividend is added and alters the amount of return. The actual return is calculated as follows:

$$R = \frac{Close\ T - Close\ T-1}{Close\ T-1} \qquad (2)$$

3.4 Twitter Data

The Twitter data is manually collected and consist of daily tweets from CEOs. The period of interest covers the years 2008 untill 2015. The sentiment of a tweet possibly predicts the direction of the stock return of the corresponding firm. Therefore, it is very important that the tweet and the company involved are matched. A way to match these aspects is to take a thorough look at the content of the tweets and see if they match with the main activities of the firm. Moreover, non-work-related tweets such as CEOs ‘watching sports’ or CEOs ‘having a dinner’ are also collected in order to have a broad sample of tweets in different kind of situations.

Hyperlinks and user mentions (e.g. @lwlang) are removed from the content of tweets to generalize the model and to make the tweets independent of specific users. Moreover, URLs do not contain relevant information and the relevant content mostly conceals within the link. Retweets, pictures and videos are filtered from the sample and are not considerd in the research. Slang language is translated into normal english and letter repetitions are deleted.

After pre-processing the tweets, a couple of other methods are used in order to restructure the tweets. These methods includes, for example, that the tweets are tokenised in order to to classify each word individually. Moreover, the tweets are lemmatised in order to simplify the words.

3.5 Other Data

In order to conduct this research properly, I also retrieved other data which is related to the thesis. For every single company, I retrieved the PERMNO, Ticker symbol, NAICS code and CUSIP header in order to identify and correlate all the date and the corresponding companies with each other. Next to all the financial and twitter data, I gathered information about the total amounts of shares outstanding for every company in order to calculate the market capitalization, which is used as a control variable for size effects. I used the logarithm amounts in order to cope with large outliers. For the control variables of the normal return I collected the high-minus-low factor (HML), the small-minus-big (SMB) factor and the momentum (MOM) factor for every single day in the research in order to control for size, value and momentum effects and measure the normal return. Next to it, I also retrieved the value-weighted AMEX-NYSE-NASDAQ market index returns which includes dividend amounts. All these factors are also described by Carhart (1997) as an extension of the Fama and French 3-factor model Fama & French (1993). These are all used to calculate the normal return for the calculation of the abnormal return. Finally, I also captured a company specific variable like the price-to-book ratio, which controls for over- and under valuation of the stocks. This ratio is calculated by dividing the market value by the book value of a stock. The market value contains the closing price of a specific day and the book value is calculated as follows;

$$Book\ Value = \frac{Total\ Assets - Total\ Liabilities}{Shares\ Outstanding} \quad (3)$$

4. Methodology

This part describes the methods being used in this thesis to get from the collection of tweets to the end results. It presents the division of the different types of categories for the classification of sentiment and shows the used methods to test for significance for every single classification and combinations of classification (e.g. positive tweet combined with value-relevancy).

4.1 Tweet Categorization

Before answering the hypotheses, where the different aspects of tweets with respect to stock price returns will be analysed, one should first divide the different tweets into categories. Based on data mining and pre-processing of the tweets in Rapidminer and the dictionary of Harvard-IV (Jorgenson & Vu, 2005), this thesis follows the way Tetlock (2007) and Da, Engelberg, & Gao (2011) started their analysis of tweets regarding sentiment classifications. The Harvard-IV dictionary is often used by online news articles and in the press.

The thesis seperates tweets into three different sentiments; positive, neutral and negative. This division is shown in Table 2 by means of some examples of tweets. Due to two different algorithms of tools in Rapid Miner (i.e. Alien text analysis) and Rosette (i.e. Rosette text analysis) and a manual check of the word list of Harvard-IV, this thesis succeeds to sort these different sentiments into the previous mentioned categories, including having robustness, due to the usage of different checks. Rapidminer also provides a confidence level of polarity and subjectivity in order to show the likelihood that a specific word certainly belongs into the right category, which is used as background support.

The other way this thesis categorizes, includes the divison between value relevant tweets and non-value relevant tweets. Company-related news announcements and other work-related tweets belong to the category ‘value-relevant’ and non-work-related tweets belong to the category ‘non-value relevant’. Table 3 presents some examples of tweets of CEOs related to value-relevant or non-value relevant categories.

4.2 Pre-processing of Data

In order to process all the data and to conduct an event study for the regressions, one should first pre-process all the data. As already mentioned before, all the tweets were tokenized and lemmatised. This implicates that words, phrases and sentences are grouped together and are seen as single independent items. Moreover, the different verbs of words are simplified to classify these sentences easier of the algorithms. I converted the dates into numerical amounts in order to correlate the date. For all the different kind of classifications, different contents and fixed effects dummy variables are made. For example, if a specific tweet has a positive load, the dummy variable "dPOS" shows a "1". Next to it, tweets made during the weekend are aggregated and shifted towards the first trading day in order to still come up with an effect with regard to the stock price of the firm. Events with more than 170 trading days before and 130 trading days after the event are also excluded from the sample in order to keep the dataset more efficient and clean the set for noise. Further, I clean the dataset for events with prices lower than one dollar and very small returns. I winsorize at the 1% and 99% percentile in order to control for outliers. After pre-processing of the data, I end up with approximately 5000 events for the analysis of the event study.

4.3 The Event Study

The method of research in this thesis is called 'the event study'. The event study is a method that is mostly used in financial papers and is based on a statistical method to calculate the effect of an abnormal event by subtracting the normal return from the realized return. The objective of the study is to find the stock price effects which are related to and affected by the abnormal event, which is in our thesis a single tweet.

A key task when conducting a event study consists of identification of the events and measuring the time period over which the stock prices got examined. Every single tweet is considered as an event. The thesis analyses approximately 5000 events. The time window in which the events will be analysed and calculated for abnormal returns is called ''the event window'' (Ranco, Aleksovski, Caldarelli, Grcar, & Mozetic, 2015). There is not any perfect parameter to measure this event window (Konchitchki & O'Leary, 2011). However, in this thesis the event window consists of the main event (tweet) and a few days before and after the event (T-1 and

T+1). These days are included because of the possible information that could be received by the market before the actual tweet has been made.
On the other hand, you do not want to have a event window which is too large so that it possibly could cause unnecessary noise. Moreover, in order to test for robustness, I also test additionally at the T-3 and T+3 and T-5 and T+5 level.

The abnormal return is defined as the difference between the actual return and the return of the chosen benchmark (Dimson & Marsh, 1986). The abnormal return is presented in the following formula:

$$AR = R - E(R) \tag{1}$$

At first, the abnormal return contains the ex post return (R). This is the actual return in presence of the abnormal conditions. The second part of the formula contains the expected return (normal return) and incorporates a return which is calculated in absence of the event. The benchmark within the normal return is crucial and determines the quality of the event study (Dimson & Marsh, 1986). The benchmark in this thesis is included in the multi-factor model of Carhart (1997). This model includes the same factors (HML, SMB and RM) as Fama and French (1993), but additionally corrects for momentum effects in the normal return. The benchmark represents the market return and is in our case the AMEX-NYSE-NASDAQ index. The thesis consists of a wide variety of companies and therefore this benchmark is suitable. The estimation window for the normal return consists of 170 days before until 21 days before the event. The model for the calculation of the normal return looks as follows:

$$E(R)_{i,t} = \alpha + \beta_{i,RM}RM + \beta_{i,SMB}SMB + \beta_{iHML}HML + \beta_{iMOM}MOM + \epsilon_{i,t} \tag{2}$$

In the thesis, I also conduct the cumulative abnormal returns (CAR). Not only the performance of the event date, but also the period surrounding the tweet is interesting and important. The cumulative abnormal return is the sum of the abnormal returns in consecutive days around the event date. This should be done, in order to control for variation of the returns and draw for overall interferences for the event of interest in the daily abnormal returns (MacKinlay, 1997).

It could be possible that we expect for some firms and CEOS to have positive stock returns after an positive announcement but unintentionally ending up at negative returns. According to MacKinlay (1997), tests with one event observation are not useful and therefore it will be necessary to aggregate the abnormal returns. The corresponding event windows consist of T-1 and T+1, T-3 and T+3 and T-5 and T+5. The formula for the CAR is as follows:

$$CAR\ (t1, t2) = \sum_{t1}^{t2} AR_t \tag{3}$$

The third method in the thesis is to test for significance. One is computing a t-statistic to check if the CAAR and AAR are significantly different from zero. One uses the normal parametric t-test as developed by Patell (1976) in order to see if the absolute value of the test is larger than 1.96. If that is the case, than the average abnormal return is significantly different to zero at 5%. The formula for the calculation of the t-test and the AAR and CAAR are as follows:

$$T - test = \frac{\frac{\sum AR}{N}}{(AR_{SD}/\sqrt{N})} \tag{4}$$

$$AAR_t\ =\ \frac{1}{N}\sum_{i=1}^{N} AR_{it} \tag{5}$$

$$CAAR\ (T_1T_2) =\ \frac{1}{N}\sum_{i=1}^{N} CAR\ (T_1T_2) \tag{6}$$

The fourth method in the thesis is the OLS regression. I will use OLS regressions to test for significance for the independent variables with the CAR for different event windows as the dependent variable. The regression will be controlled by independent variables as the sentiment of the tweet, the logarithm of the market capitalization and the price to book ratio of the firms. The market capitalization and the P/B value control for market shocks and size differences. CEO fixed effects industry fixed effects and year fixed effects are also present in the regression.

The OLS regression will be conducted as follows:

$$\mathrm{C}AR_t = \alpha + \beta_{1i} SENTIMENT + \beta_{2i} WORKRELEVANCY + \beta_{3i} \log(MARKETCAPITALIZATION) + \beta_{4i}\ P/B\ RATIO \quad (7)$$

5. Results

In this section, I will present the results with respect to the three hypotheses and additionally give an answer to the research question. It presents the different methods that are used and determines whether the results are significant or not. The thesis additionally tests with different event windows for CAR in order to be coherent with previously written papers. These results will also be present in this part.

The descriptive statistics (Table 5 in Appendix) show that the amount of positive tweets consist of 33% of the whole sample. Most of the tweets are considered to be neutral (57%), where negative tweets consist of the fewest percentage (10%). CEOs prevent theirselves for possible negative effects on stock prices, whereby they mostly tweet positive or neutral. With regard to value-relevancy of the tweets, which means whether they are work-related or not, one can see that most tweets contain work-related content (57%). This is plausible, due to the fact that CEOs want to trigger investors to be active by means of social media contact. Outliers are already winsorized at the 1% and 99% level and are therefore not substantially present in the sample.

In order to test hypotheses one and two by testing for significance, one should first calculate the (cumulative) abnormal returns for every type of sentiment and tweet content. After having calculated all the abnormal returns, which is only applicable for days within the different event windows, one should test for significantly difference to zero for the (cumulative) average abnormal returns by means of the T-test.

Table 6 in the Appendix shows the abnormal return for tweets with different sentiments at the event date. It appears that tweets which contain a negative sentiment, on average, end up with the highest abnormal return (0.16%). This is in contradiction with the paper of Chen, Hwang, & Liu (2015) who argued that tweets with a negative sentiment end up with a decrease in stock price returns. On the other hand, Bollen, Mao, & Zeng (2010) stated that negative sentiments are not predictive for the DJIA (Dow Jones Industrial Average). Moreover, the effect of the abnormal return can also negatively revert in the following days after the event. This will be further investigated in the T-test calculation and OLS regressions of the cumulative abnormal returns, where I additionally test for the mediator effect.

A third reason could be the relative small amount of negative tweets, which leads to a worser prediction in comparison with the large proportion of positive and neutral tweets. Positive tweets result, on average, in an increase of the abnormal return of 0.11%. This positive result is in accordance with previously written papers such as Bartov, Faurel, & Mohanram (2017) and Jame, Johnston, Markov, & Wolfe (2016). Neutral tweets end up, at the event date, with the lowest average abnormal return. This is plausible, due to the fact that investors can not relate if the announcement turns out to be positive or negative, which therefore could be considered as to risky to invest in.

In Table 7 of the Appendix the distinction between work-relevant tweets and non-work relevant tweets and its effect on the abnormal return is shown. One can derive from the table that the abnormal return for work-related tweets (0.12%) is higher than tweets with a non-work-related content (0.08%). This seems plausible as investors are more likely to react to company-related news than towards personal circumstances of the CEO. This is contradictionary to the paper of Chen, Hwang, & Liu (2015), who argued that the personal mood of a CEO also has an impact on stock price returns.

In Table 8 of the Appendix the cumulative abnormal returns are shown with respect to the different sentiments of tweets. A further distinction is made with respect to different event windows. One can conclude that the more days involved in the event window of CAR for negative tweets, on average, the more negative the cumulative abnormal return ends up (from 0.21% untill -0.20%). As already mentioned, it seems to be that negative tweets revert in the following days after the event and that investors consider the content of the tweet seriously (whether it is work-related or not). Positive tweets have, on average, the highest cumulative abnormal return at the t-1 and t+1 level (0.15%) . This return decreases if the event window is larger ([t-3,t+3] and [t-5,t+5]). Neutral tweets seems to have the highest cumulative abnormal return at the t-5 and t+5 level. This seems to be surprising as neutral tweets are unpredictable. One plausible explanation of the small differences between the different sentiments and event windows and the fact that the effect of tweets revert, is the primarily incite naïve reaction of investors (Chen, Hwang, & Liu, 2015). An other explanation is the frequent amount of sarcasm in tweets, which could cause biased results for specific days (Reyes, Rosso, & Veale, 2013).

In Tabel 9 of the Appendix the cumulative abnormal returns are shown with respect to the work relevancy of the tweets. Work-related tweets have, on average, a higher abnormal return than non-work-related tweets. This is applicable for all the event windows, where one can observe that non-work-related tweets barely got an abnormal return. This is in contrast with previous papers, where they stated that tweets about the personal life of a CEO (which are non-work-related) also have impact on the stock price returns (Chen, Hwang, & Liu, 2015).

In Table 10 of the Appendix, the average mean of abnormal returns and the T-statistics are shown for both work-related and non-work related tweets for every date in the range of [-10,10]. With respect to hypothesis one, this table helps us to determine the significant impact of work-related tweets on stock price returns and to determine if they are significantly different to zero. However, this table only shows if there is significant impact on event window days surrounding the event without taking other control variables and fixed effects into account. According to the table, it seems that this is present for work-related tweets at [t=0] where we have, on average, a significant impact (1% level) and an abnormal return of 0.13%. Moreover, it seems to be that there was insider information as there were signals of company-related information before the tweet was disseminated, as there are six significant abnormal returns before the tweet date. Further on, it seems to be that the effect of the tweet dissapears after it was send, as there is no convinced indication of significant impact at days after the event. With respect to non-work related tweets, one could say that there is no presence of significance with regard to the stock price returns. Despite of having significant days at [t+1] and [t+8], it is plausible to say that this is not sufficient to consider non-work related tweets as significantly different to zero.

In Table 11 of the Appendix, the cumulative average abnormal returns per content of the tweets are shown with respect to different event windows. This table confirms that work-related tweets have significant impact at the stock price returns in the environ of the event date. Significant t-values for CAR[-1,1], CAR[-3,3] and CAR[-5,5] show that work-related tweets, on average, have a positive impact on the stock price returns. Work-related content seems to trigger investors to react positively towards the announcements of the company. The highest cumulative average abnormal return for work-related tweets is 0.46% [t-5, t+5].

When one looks at non-work related tweets, again, there is not any significant impact on the stock price returns observable. Therefore, it is plausible to say that non-work related tweets, on average, do not have any significant impact on stock price returns.

Table 12 shows the average abnormal return for the different sentiments within the event window [-10, 10] and determines if these averages are significant by means of the T-test. Just as in Table 10, all other control variables and fixed effects are omitted in the calculations. One can see that all sentiments seems to be significant at the eventdate of the tweet [t=0]. For negative tweets this is, on average, a positive abnormal return of 0.16%, which seems to be surprising as we expect that negative tweets cause, on average, an negative abnormal return. However, after the event date, most of the abnormal returns for tweets with a negative sentiment revert into negative returns due to the incite reaction of investors at the event date, which is consistent with the paper of Chen, Hwang, & Liu (2015). Moreover, these returns are, except of [t+3], not significant. Positive tweets are, at the event date, significant at a 10% level and have an average abnormal return of 0.11%, which is lower than negative tweets. Neutral tweets are at the event date significant at a 5% level and have an average abnormal return of 0.09%. Next to it, one also can conclude that the neutral announcements were already known by insiders due to the significant positive reactions prior to the events. In conclusion, it is plausible to say that for every sentiment the significant impact of the tweet only exists at the event date itself and that the effect disappears in the successive days after the event date.

In Table 13, one observes the cumulative average abnormal return of tweets with different sentiments at different event windows. One can see in this table that neutral tweets have, on average, a significant positive impact on the stock prices returns and are therefore correlated. Neutral tweets mostly incorporate company announcements and other work-related news and are therefore a good source of information on the long run, which triggers investors to invest. Other remarkable results are the turn of negative tweets at the CAR [-3, 3] level into negative average abnormal returns. It seems to be that investors primarily react wrong and wait for the impact of a negative tweet so that they can react more correctly after the event date.

In Table 14, the OLS regression for the CAR [-1,1] is presented where one, in contraction with the T-tests, corrects for year-, industry- and ceo fixed effects.

These fixed effects capture the time-series trends in the years of observation, the peaks in different industries and specific CEOs, who, for instance, tweet very often or freuently in a specific mood. This increases the reliability of the research substantially (increase of R2 towards 11%-12%). Therefore, we mostly conclude for the hypotheses on the basis of these regressions instead of the T-tests. One can see in the table, that we made six different regressions per CAR, which contain different variables for each regression. As one can see in the different regressions for CAR[-1,1], is that positive sentiments seem to be highly statistically significant for, on average, positive cumulative abnormal returns (e.g. 0.26%). These results are consistent with previously written papers. Tweets with a negative sentiment seems to be highly statistically significant with, on average, negative cumulative abnormal returns (-0.26%). When one adds the mediator effect “negative*work” to the regression, one can see that the negative sentiment variable reverts into a positive variable. This implicates that non-work related negative tweets have a positive influence on the cumulative abnormal return. If the tweet is work-related, negative sentiments still end up with a negative result (-1.1%). Further on, one observes that work relevant tweets are positively significant and are, on average, positively correlated to the cumulative abnormal returns (0.12%). Work-relevant tweets seem to trigger investors to invest in the company. At last, we also encounter, on average, significant results for the Ln Marketcap (-0.4%) and P/B ratio (0.03%). This seems plausible as value stocks are more likely to earn returns independently of a tweet and are not sensitive to sentiment. The significant increase of the CAR by means of a smaller market capitalization, has to do with the fact that small firms are more risky than large firms (Vassalou & Xing, 2004). It is therefore plausible to say that investors are more likely and willing to stay concerned with the firm as they want to know every single detail about the company and the related risk of their investment.

Table 15 presents the CAR [-3,3] with the same regressions as Table 14 has. The previously made fixed effects are also applicable in this table. One can see that most of the results are significant and have the same signing of the coefficients as the table of CAR [-1,1]. It seems to be that the longer the event window, the stronger the effects of the variables on the CAR will be. It seems to be that work-related tweets are the exception, as they have a negative influence on the CAR [-3,3].

This is in contradiction with the previous table and the T-test tables, where we showed that the relation between work-related tweets and the cumulative abnormal return is positive. This is also applicable for Table 16, where I encountered the same results as in Table 15.

In conclusion, one can say with respect to hypothesis one, that we accept the hypothesis and conclude that work-related tweets have, on average, a significant impact on the stock price at the event date and in the cumulative days of the event window. This is both the case for the t-tests as for the regressions, which both show that there is significant impact. However, the OLS regressions show that work-related tweets for CAR [-1,1] have a positive impact and for [-3,3] and [-5,5] a negative impact. One plausible explanation of this inconsistency, could be one the shortcomings of the research, which we discuss in the conclusion. These results are therefore not expected. With regard to hypothesis two, one can say that this hypothesis is accepted for both positive tweets and negative tweets. The regressions show that there are, on average, positive cumulative abnormal returns for positive tweets and, on average, negative cumulative abnormal returns for negative tweets. These results are consistent with the paper of Chen, Hwang, & Liu (2015) who found the same results. With respect to the third hypothesis, one can say that this hypothesis is accepted for the combination between negative sentiments and work-related tweets. This cumulative abnormal return ends up, on average, with a more significant negative return than the two variables summed up independently. This is applicable for CAR[-1,1], CAR[-3,3] and CAR[-5,5]. The results for positive and work-related tweets are not significant for CAR[-1,1] and CAR[-5,5]. Therefore, I conclude that it only makes sense for the cumulative abnormal stock price return when you combine work-related tweets with negative tweets.

6. Conclusion

In this section, I will give an answer to the research question and will briefly summarize and conclude the main results of this thesis. I will describe how this thesis contributes to the existing literature and discuss the limitations and short comings of the research. Further on, I will present the recommendations for improvement in further research.

Based on financial information and the usage of Twitter by CEOs, this thesis finds results about the impact of social activity of CEOs on the value of the firm. We summed up all the results and determined if the hypotheses are significant in the previous section. Therefore, we are able to give an answer to the research question of the thesis in this section. We conducted T-tests and multiple regressions were we corrected for year-, industry- and CEO fixed effects and used different event windows for cumulative abnormal returns. In conclusion, in order to answer the research question, one can say that there is a significant matter of impact of Twitter on the stock price returns. Based on hypotheses 1 and 2, we can conclude that work-related, positive and negative tweets have significant impact on the stock price returns and therefore on the value of the firm (i.e. market value of the firm). This is also the case when we take different event windows for the cumulative abnormal return into account and check for robustness for the different sentiments and content. We found out that the longer the period of the event window, the more significant, on average, the impact of positive and negative tweets to the cumulative abnormal return is. In contradiction of the paper of Chen, Hwang, & Liu (2015), we observed that non-work related tweets are not significantly important for stock price returns. Basically, it is plausible to say that investors do not take personal tweets serious and do not relate them to the performance of the company. Based on hypothesis 3, we concluded that the combination of work-related and negative tweets leads to a more significant outcome to the stock price returns than if you take both variables independently. This outcome is not significant for the combination of positive and work-related, but present for some event windows.

This research contributes to the prior literature about the role of social media in financial markets. This could be important for firms as it delivers some evidence of the effect of Twitter towards the stock price returns of the firm. It is therefore possible for firms to control a small part of the stock price return by means of Twitter.

Eventually, the analysis confirms that there is a continous evolvement of online communication and interaction within the financial market.

Nevertheless, our analysis also has a number of limitations. First of all, we did not correct for the whole crisis period, where I expect that most CEOs tweet about the negative economic circumstances during that period. This could bias some of the results for negative tweets. Secondly, I used a sample, which only consists of US publically listed firms. A suggestion would be to extend the research to more countries and to see if the impact of tweets in other countries would provide additional evidence for significance. Third, it would better to categorize the tweets perfectly with better statistic methods, as we now did not take, for instance, sarcasm into account. However, this would cause even more time than it already took now. At last, it would be better if the tweets did not overlap each other, so that you can determine the impact of every tweet individually. Due to a lack of time, this was unfortunately not possible.

Reference List

Abrahams, A., Jiao, J., Wang, G., & Fan, W. (2012). Vehicle defect discovery from social media. *Decision Support Systems, 54 (1)* , 87-97.

Bagehot, W. (1971). The only game in town. *Financial Analysts Journal 27(4)* , 28-35.

Barber, B., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behaviour of individual and institutional investors. *Review of Financial Studies 27(4)* , 785-818.

Bartov, E., Faurel, L., & Mohanram, P. (2017). Can Twitter help predict firm level earnings and stock returns? *Rotman School of Management Working Paper No 2631421* .

Beyer, A., Cohen, D., Lys, T., & Walther, B. (2010). The financial reporting environment: Review of the recent literature. *Journal of Political Economy, 111 (5)* , 296-343.

Blankespoor, E., Miller, G., & White, H. (2014). The role of dissemination in market liquidity: Evidence from firms' use of Twitter. *The Accounting Review 89(1)* , 79-112.

Bollen, J., Mao, H., & Zeng, X. (2010). Twitter mood predicts the stock market. *Journal of Computational Science, 2(1)* , 1-8.

Campbell, J., Lo, A., & Mackinlay, A. (1997). *The econometrics of financial markets.* Princeton University Press.

Carhart, M. (1997). On Persistence in Mutual Fund Performance. *Journal of Finance, 52(1)* , 57-82.

Chen, H., Hwang, B., & Liu, B. (2015). The Economic Consequences of having ''Social'' Executives. *Working Paper* .

Core, J. (2001). A review of the empirical disclosure literature: Discussion. *Journal of Accounting and Economics 31(1-3)* , 441-456.

Culnan, M., McHugh, P., & Zubillaga, J. (2010). How large U.S. companies can use Twitter and other social media to gain business value. *MIS Quarterly Executive, 9 (4)* , 243-259.

Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. *The Journal of Finance, 66(5)* , 1461-1499.

Das, S., & Chen, M. (2007). Yahoo! for Amazon: Sentiment extraction from small talk on the web. *Management Science, 53(9)* , 1375-1388.

Debreceny, R., Gray, G., & Rahman, A. (2002). The Determinants of Internet financial reporting. *Journal of Accounting and Public Policy 21 (4-5)* , 371-394.

Dijkman, R., Ipeirotis, P., Aertsen, F., & van Helden, R. (2015). Using Twitter to Predict Sales: A Case Study. *Beta Working Paper series 471* , 1-14.

Dimson, E., & Marsh, P. (1986). Event study methodologies and the size effect: The case of UK press recommendations. *Journal of Financial Economics, 17(1)* , 113-142.

Fama, E., & French, K. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33(1)* , 3-56.

Fields, T., Lys, T., & Vincent, L. (2001). Empirical research on accounting choice. *Journal of Accounting and Economics 31(1-3)* , 255-307.

Harvard Business Review Analytic Services. (2010). *The New Conversation: Taking Social Media from Talk to Action.*

Healy, P., & Palepu, K. (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics 31 (1-3)* , 405-440.

Heaton, J., & Lucas, D. (2000). Stock Prices and Fundamentals. *NBER Macroeconomics Annual 1999, Vol. 14* , 213-264.

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)* , 2801-2824.

Jame, R., Johnston, R., Markov, S., & Wolfe, M. (2016). The value of crowdsourced earnings forecasts. *Journal of Accounting Research, 54 (4)* , 1077-1110.

Jorgenson, D., & Vu, K. (2005). Information technology and the World Economy. *The Scandinavian Journal of Economics, 107(4)* , 631-650.

Kaplan, A., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons, 53 (1)* , 59-68.

Konchitchki, Y., & O'Leary, D. (2011). Event studies methodologies in information systems research. *International Journal of Accounting Information Services, 12(2)* , 99-115.

Lewellen, W., Park, T., & Ro, B. (1996). Self-serving behavior in managers discretionary information disclosure decisions. *Journal of Accounting and Economics, 21(2)* , 227-251.

MacKinlay, A. (1997). Event studies in economies and finance. *Journal of Economic Literature, 35(1)* , 13-39.

Mangold, C., & Faulds, D. (2009). Social Media: The new hybrid element of the promotion mix. *Business Horizons, 52 (4)* , 357-365.

Maniatis, P. (2011). Individual stock market risk and price valuation: the case of Titan S.A. *Managerial Finance, 37 (4)* , 347-361.

Mizrach, B., & Weerts, S. (2009). Experts online: an analysis of trading activity in a public Internet chat room. *Journal of Economic Behavior & Organization, 70(1-2)* , 266-281.

Naaman, M., Boase, J., & Lai, C.-H. (2010). Is it really about me? Message content in social awareness streams. *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work* , 189-192.

Nair, M. (2011). Understanding and measuring the value of social media. *Journal of Corporate Accounting & Finance, 22 (3)* , 45-51.

Ng, L., & Wu, F. (2006). Peer effects in the trading decisions of individual investors. *Financial Management, 39(2)* , 807-831.

Nguyen, V., Varghese, B., & Barker, A. (2013). The royal birth of 2013: Analysing and visualising public sentiment in the UK using Twitter. *IEEE International Conference on Big Data* , 46-54.

Patell, J. (1976). Corporate forecasts of earnings per share and stock price behavior: Empirical test. *Journal of Accounting Research, 14 (2)* , 246-276.

Ranco, G., Aleksovski, D., Caldarelli, G., Grcar, M., & Mozetic, I. (2015). The Effects of Twitter Sentiment on Stock Price Returns. *PLoS ONE, 10(9)* , e0138441.

Reyes, A., Rosso, P., & Veale, T. (2013). A multidimensional approach for detecting irony in Twitter. *Language Resources and Evaluation, 47(1)* , 239-268.

Sloan, R. (1996). Do Stock Prices Fully Reflect Information in Accruals and Cash Flows about Future Earnings. *The Accounting Review 71 (3)* , 289-315.

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

Steiner, P. (1975). *Mergers: Motives Effects. Policies.* Ann Arbor: University of Michigan Press.

Tetlock, P. (2007). Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance, 62(3)* , 1139-1168.

Vassalou, M., & Xing, Y. (2004). Default Risk in Equity Returns. *The Journal of Finance, 59(2)* , 831-868.

Zhang, X., Fuehres, H., & Gloor, P. (2011). Predicting stock market indicators through Twitter - ''I hope it is not as bad as I fear''. *Procedia-Social and Behavioural Sciences, 26* , 55-62.

Zhou, M., Lei, L., Wang, J., Fan, W., & Wang, A. (2015). Social Media Adoption and Corporate Disclosure. *Journal of Information Systems, 29 (2)* , 23-50.

Appendix

Table 1: List of socially active CEOs

This table presents socially active CEOs in the period of 2008-2015 out of the S&P 1500. This table represents: 1) The Twitter name of the CEO; 2) The first, middle and last name of the CEO; 3) The name of the company; 4) The TAA of the CEO (i.e. Twitter Account Activation date)

Twitter name	CEO Name	Company Name	TAA
openJonathan	Jonathan Swartz	ORACLE AMERICA CORP	01-12-2007
gweston	Graham Weston	RACKSPACE HOSTING INC	01-03-2008
lnapier	A. Lanham Napier	RACKSPACE HOSTING INC	01-03-2008
marcchardon	Marc E. Chardon	BLACKBAUD INC	01-04-2008
shirleyedge	Shirley Singleton	EDGEWATER TECHNOLOGY INC	01-04-2008
reedhastings	Reed Hastings	NETFLIX INC	01-06-2008
loudinardo	Louis Dinardo	EXAR COMP	01-07-2008
gcolony	George F. Colony	FORRESTER RESEARCH INC	01-08-2008
jonasprising	Jonas Prising	MANPOWERGROUP	01-08-2008
manpowergroupjj	Jeffrey A. Joerres	MANPOWERGROUP	01-04-2009
briandunn	Brian J. Dunn	BEST BUY CO INC	01-09-2008
marissamayer	Marissa A. Mayer	YAHOO INC	01-11-2008
jmcaughlin173	John P. Mclaughlin	PDL BIOPHARMA INC	01-11-2008
roblocascio	Robert P. Locascio	LIVEPERSON INC	01-12-2008
jdjr2009	James DeGraffenreidt	WGL HOLDINGS INC	01-01-2009
dkirchhoff	David P. Kirchhoff	WEIGHT WATCHERS INTL INC	01-02-2009
finkd	Mark Zuckerberg	FACEBOOK INC	01-02-2009
megWhitman	Margaret C. Whitman	HEWLETT PACKARD ENTERPRISE	01-02-2009
rapino99	Michael Rapino	LIVE NATION ENTERTAINMENT	01-02-2009
carlbass	Carl Bass	AUTODESK INC	01-02-2009
stevebmicrosoft	Steven A. Ballmer	MICROSOFT CORP	01-02-2009
satyanadella	Satya Nadella	MICROSOFT CORP	01-02-2009
kaufer	Stephen Kaufer	TRIPADVISOR	01-02-2009
johnriccitiello	John S. Riccitiello	ELECTRONIC ARTS INC	01-02-2009
mr2matt	Matthew E. Rubel	COLLECTIVE BRANDS INC	01-02-2009
grimshawstuart	Stuart Ian Grimshaw	EZCORP INC – CL A	01-02-2009

Table 1. Continued.

dpassmanii	S. David Pasmann III	CARMIKE CINEMAS INC	01-02-2009
vivek	Vivek Y. Ranadive	TIBCO SOFTWARE INC	01-03-2009
gianforte	Gregory R. Gianforte	RIGHTNOW TECHNOLOGIES IN	01-03-2009
johnheyman	John H. Heyman	RADIANT SYSTEMS INC	01-03-2009
rsdearth	Randall S. Dearth	CALGON CARBON CORP	01-03-2009
scarley01	Stephen E. Carley	RED ROBIN GOURMET BURGERS	01-03-2009
lwlang	Laura Wicke Lang	TIME INC	01-04-2009
m_abraham	Magid M. Abraham	COMSCORE INC	01-04-2009
iridiumboss	Matthew J. Desch	IRIDIUM COMMUNICATIONS INC	01-04-2009
gregmarcus	Gregory S. Marcus	MARCUS CORP	01-04-2009
rdstiley	Randall D. Stilley	PARAGON OFFSHORE PLC	01-04-2009
melkarmazin	Mel Karmazin	SIRIUS XM HOLDINGS INC	01-04-2009
pehongchen	Pehong Chen	BROADVISION INC	01-04-2009
dkhos	Dara Khosrowshahi	EXPEDIA INC	01-05-2009
dave_wentz	David A. Wentz	USANA HEALTH SCIENCES INC	01-05-2009
rmeeusen	INCRichad Meeusen	BADGER METER INC	01-05-2009
colemaned	J. Edward Coleman	UNISYS CORP	01-05-2009
bslobodow	Brian Slobodow	U S SILICA HOLDINGS INC	01-05-2009
bobhagerty	Bob C. Hagerty	POLYCOM INC	01-05-2009
timbiltz	Timothy G. Biltz	LUMOS NETWORKS CORP	01-05-2009
irvingazoff	Irving L. Azoff	TICKETMASTER ENTERTNMNT INC	01-06-2009
stevessingh	Sudhir Steven Singh	CONCUR TECHNOLOGIES INC	01-06-2009
tradrules	Salomon Sredni	TRADESTATION GROUP INC	01-07-2009
mtberg	Mark T. Bertolini	AETNA INC	01-09-2009
craigherkert	Craig R. Herkert	SUPERVALU INC	01-09-2009
robopoh	Robert C. Pohlad	PEPSIAMERICAS INC	01-09-2009
ericschmidt	Eric E. Schmidt	ALPHABET INC	01-12-2009
timarmstrongaol	Timothy M. Armstrong	AOL INC	01-01-2010
learningmoment	Garry O. Ridge	WD-40 CO	01-01-2010
donahoe_john	John J. Donahoe	EBAY INC	01-02-2010
tkdendallhunt	T. Kendall Hunt	VASCO DATA SEC INTL INC	01-02-2010
jbaliff	Jonathan E. Baliff	BRISTOW GROUP INC	01-03-2010

Table 1. Continued.

sergematta	Serge Matta	COMSCORE INC	01-05-2010
great_again	Henry R. Nothhaft	TESSERA TECHNOLOGIES INC	01-06-2010
karmanosunplggd	Peter Karmanos Jr	COMPUWARE CORP	01-10-2010
maryjunck	Mary E. Junck	LEE ENTERPRISES INC	01-12-2010
michael_saylor	Michael J. Saylor	MICROSTRATEGY INC	01-01-2011
jefferyrgardner	Jeffery R. Gardner	WINDSTREAM HOLDINGS INC	01-01-2011
bkrunner	Brian M. Krzanich	INTEL CORP	01-02-2011
healthmgmtceo	Gary D. Newsome	HEALTH MANAGEMENT ASSOC	01-02-2011
wymanroberts	Wyman T. Roberts	BRINKER INTL INC	01-03-2011
msiegal1	Michael D. Siegal	OLYMPIC STEEL INC	01-03-2011
dpstockert	David P. Stockert	POST PROPERTIES INC	01-03-2011
diane_irvine	Diane M. Irvine	BLUE NILE INC	01-04-2011
rambo1724	Randy L. Ortiz	LOJACK CORP	01-07-2011
amolinaroli	Alex A. Molinaroli	JOHNSON CONTROLS INC	01-08-2011
davidfieldetm	David J. Field	ENTERCOM COMMUN. CORP.	01-08-2011

Table 2: Tweet Sentiment Examples

This table shows tweets with a positive, neutral or negative sentiment. This relates to the amount of ‘negative’ words a message has and in which context these words are placed. Negative words are shown in red. Positive words are shown in green. Moreover, this table shows: 1) The content of the tweets; 2) The date of the tweets; 3) The name of the CEO; 4) The name of the company; and 5) The sentiment which belongs to the mood of the tweet.

Tweets	Tweet Date	CEO Name	Company Name	Sentiment
Congratulations to Jason and Ryan. Milwaukee’s first culture carrier. Thanks Milwaukee team for hosting me.	14-03-2016	Tim Yates	Monster World Wide	Positive
"Taxi!" Caught a cab in Beijing this morning with Didi Chuxing's Jean Liu.	16-05-2016	Tim Cook	Apple INC	Neutral
Without a doubt we have lost the greatest innovator of our time.	06-10-2011	John Donahoe	EBAY INC	Negative

Table 3: Tweets Value-relevancy Examples

This table presents tweets with a value relevant content or a non-value relevant content. News announcements and work-related tweets are associated with value relevant content and non-work-related tweets are associated with non-value relevant. Moreover, this table shows: 1) The content of the tweets; 2) The date of the tweets; 3) The name of the CEO; 4) The name of the company; and 5) The category of the tweets, whether it's a news announcement, work-related or non-work-related.

Tweets	Tweet Date	CEO Name	Company Name	Category
"Taxi!" Caught a cab In Beijing this morning with Didi Chuxing's Jean Liu.	16-05-2016	Tim Cook	Apple INC	Non-work-related
Good talk w/women leads @chicago office - tons of challenges + opportunities to grow our female leaders in #Expedia	10-03-2016	Dara Khosrowshahi	EXPEDIA INC	Work-related
Calgon Carbon: An Environmental Growth Stock With High Risk And Reward http://seekingalpha.com/a/wlls $CBT $DHR $HWKN $MWV $XYL $CCC	03-07-2016	Randall Dearth	CALGON CRBON GRP	News Announc.

Table 4: Stock Market Tickers

This table presents the companies included in the thesis with their corresponding ticker symbol. Most companies are represented in the complete time sample (e.g. 2008-2015). A few companies are active in a part of the sample, due to ending up in default or de-indexing.

Ticker Symbol	Company Name
ORLC	ORACLE AMERICA CORP
RAX	RACKSPACE HOSTING INC
BLKB	BLACKBAUD INC
EDGW	EDGEWATER TECHNOLOGY INC
NLFX	NETFLIX INC
FORR	FORRESTER RESEARCH INC

Table 4. Continued.

MAN	MANPOWERGROUP
BBY	BEST BUY CO INC
YHOO	YAHOO INC
PDLI	PDL BIOPHARMA INC
LPSN	LIVEPERSON INC
WGL	WGL HOLDINGS INC
MSFT	MICROSOFT CORP
FB	FACEBOOK INC
HPE	HEWLETT PACKARD ENTERPRISE
LYV	LIVE NATION ENTERTAINMENT
ADSK	AUTODESK INC
TRIP	TRIPADVISOR
EA	ELECTRONIC ARTS INC
CKEC	CARMIKE CINEMAS INC
TIME	TIME INC
IRDM	IRIDIUM COMMUNICATIONS INC
MCS	MARCUS CORP
PGNPQ	PARAGON OFFSHORE PLC
SIRI	SIRIUS XM HOLDINGS INC
BVSN	BROADVISION INC
EXPE	EXPEDIA INC
USNA	USANA HEALTH SCIENCES INC
BMI	BADGER METER INC
UIS	UNISYS CORP
PLCM	POLYCOM INC
LMOS	LUMOS NETWORKS CORP
AET	AETNA INC
SVU	SUPERVALU INC
GOOG	ALPHABET INC
WDFC	WD-40 CO
EBAY	EBAY INC
VDSI	VASCO DATA SEC INTL INC

Table 4. Continued.

BRS	BRISTOW GROUP INC
SCOR	COMSCORE INC
TSRA	TESSERA TECHNOLOGIES INC
LEE	LEE ENTERPRISES INC
MSTR	MICROSTRATEGY INC
WIN	WINDSTREAM HOLDINGS INC
INTC	INTEL CORP
EAT	BRINKER INTL INC
ZEUS	OLYMPIC STEEL INC
PPS	POST PROPERTIES INC
NILE	BLUE NILE INC
JCI	JOHNSON CONTROLS INC
ETM	ENTERCOM COMMUNICATIONS CORP
LOJN	LOJACK CORP
TRAD	TRADESTATION GROUP INC
AOL	AOL INC
RADS	RADIANT SYSTEMS INC
CPWR	COMPUWARE CORP
PAS	PEPSIAMERICAS INC
CNQR	CONCUR TECHNOLOGIES INC
PSS	COLLECTIVE BRANDS INC
TIBX	TIBCO SOFTWARE INC

Table 5: Summary statistics for key variables used in the analysis

Variable	N	Mean	Std. Dev	Min	Max
Positive	1.426.784	0.33	0.47	0	1
Neutral	1.426.784	0.57	0.49	0	1
Negative	1.426.784	0.10	0.30	0	1
Relevant	1.426.784	0.57	0.49	0	1
Non-Relevant	1.426.784	0.43	0.49	0	1
Log Mark. Cap	1.426.784	15.08	2.10	9.46	19.93
P/B ratio	1.426.784	4.26	19.06	-121.00	1672.00
Abnormal Ret	472.363	0.0001	0.23	-0.49	1.07
Predicted 3FF	472.363	0.0007	0.15	-0.15	0.15
CAR [-1,1]	5027	0.0014	0.04	-0.41	0.43
CAR [-3,3]	5027	0.0023	0.06	-0.47	0.53
CAR [-5,5]	5027	0.0027	0.07	-0.55	0.61

The table above shows the summary statistics of the key variabels used in the sample. Positive is the variable that indicates if a tweet is considered as positive. Neutral is the variable that indicates if a tweet is considered as neutral. Negative is the variable that indicates if a tweet is considered as negative. Relevant is the variable that indicates if a tweet is considered as value-relevant. Non-relevant is the variable that indicates if a tweet is considered as non-value relevant. Log Mark. Cap is the logarithm of the market capitalization of a firm. P/B ratio is the price to book value and indicates the market value of a stock divided by the book value of the stock. Abnormal Ret is the abnormal return which is calculated as the realised return minus the normal return. Predicted 3FF is the normal return including small-minus-big, high-minus-low, momentum and the market index. CAR [-1,1] is the cumulative abnormal return in the event window [-1, +1]. CAR [-3, 3] is the cumulative abnormal return in the event window [-3, +3]. CAR [-5, 5] is the cumulative abnormal return in the event window [-5, +5].

Table 6: Abnormal Returns (t=0) for Sentiment

	Mean	Std. Error	Max	Min	N
ARpos	0.0011	0.00059	0.2474	-0.2777	1638
ARneut	0.0009	0.00039	0.2407	-0.2038	2888
ARneg	0.0016	0.00073	0.8499	-0.0684	501

The table above shows the abnormal returns for every type of sentiment of the tweet. ARpos is the abnormal return for tweets having a positive load. ARneut is the abnormal return for tweets having a neutraal load. ARneg is the abnormal return of tweets having a negative load. Mean is the average abnormal return of the corresponding sentiment. Std. Error is the standard error of the abnormal return. Max is the maximum abnormal return within the sample of the corresponding type of sentiment. Min is the minimum abnormal return within the sample of the corresponding type of sentiment. N is the amount of observations with the corresponding type of sentiment

Table 7: Abnormal Returns (t=0) for work-relevancy

	Mean	Std. Error	Max	Min	N
ARwork	0.0012	0.00037	0.2474	-0.1615	2883
ARnonwork	0.0008	0.00050	0.2474	-0.2777	2144

The table above shows the abnormal returns for every type of work-relevancy of the tweet. ARwork is the abnormal return for tweets having a content containing work-relevancy. ARnonwork is the abnormal return for tweets having a content containing non-work relevancy. Mean is the average abnormal return of the corresponding work-relevancy. Std. Error is the standard error of the abnormal return. Max is the maximum abnormal return within the sample of the corresponding type of work-relevancy. Min is the minimum abnormal return within the sample of the corresponding type of work-relevancy. N is the amount of observations with the corresponding type of work-relevancy.

Table 8: Cumulative Abnormal Returns for Sentiment

	Mean	Std. Error	Max	Min	N
$CAR[-1,1]_{POS}$	0.0015	0.00097	0.4257	-0.3192	1638
$CAR[-1,1]_{NEUT}$	0.0012	0.00149	0.3462	-0.4122	2888
$CAR[-1,1]_{NEG}$	0.0021	0.00065	0.2410	-0.1091	501
$CAR[-3,3]_{POS}$	0.0009	0.00149	0.5341	-0.3339	1638
$CAR[-3,3]_{NEUT}$	0.0038	0.00102	0.5150	-0.4677	2888
$CAR[-3,3]_{NEG}$	-0.0009	0.00219	0.1971	-0.1755	501
$CAR[-5,5]_{POS}$	-0.0007	0.00189	0.6142	-0.3532	1638
$CAR[-5,5]^{NEUT}$	0.0055	0.00127	0.6125	-0.5540	2888
$CAR[-5,5]_{NEG}$	-0.0020	0.00277	0.2752	-0.2427	501

The table above shows the cumulative abnormal returns for every type of sentiment of the tweet. The division is made between positive, negative and neutral and the different event windows for the cumulative abnormal return. CAR[-1,1] pos is the cumulative abnormal return for tweets with a positive sentiment between t-1 and t+1. CAR[-1,1] neut is the cumulative abnormal return for tweets with a neutral sentiment between t-1 and t+1. CAR[-1,1] neg is the cumulative abnormal return for tweets with a negative sentiment between t-1 and t+1. CAR[-3,3] pos is the cumulative abnormal return for tweets with a positve sentiment between t-3 and t+3. CAR[-3,3] neut is the cumulative abnormal return for tweets with a neutral sentiment between t-3 and t+3. CAR[-3,3] neg is the cumulative abnormal return for tweets with a negative sentiment between t-3 and t+3. CAR[-5,5] pos is the cumulative abnormal return for tweets with a positive sentiment between t-5 and t+5. CAR[-5,5] neut is the cumulative abnormal return for tweets with a neutral sentiment between t-5 and t+5. CAR[-5,5] neg is the cumulative abnormal return for tweets with a negative sentiment between t-5 and t+5. Mean is the average of the cumulative abnormal return of the corresponding sentiment. Std. Error is the standard error of the cumulative abnormal return. Max is the maximum cumulative abnormal return within the sample of the corresponding type of sentiment. Min is the minimum cumulative abnormal return within the sample of the corresponding type of sentiment. N is the amount of observations with the corresponding type of sentiment

Table 9: Cumulative Abnormal Return for work-relevancy

	Mean	Std. Error	Max	Min	N
$CAR[-1,1]_{WORK}$	0.0024	0.00065	0.3462	-0.4122	2883
$CAR[-1,1]_{NONWORK}$	0.0000	0.00081	0.4257	-0.3192	2144
$CAR[-3,3]_{WORK}$	0.0038	0.00103	0.5150	-0.4677	2883
$CAR[-3,3]_{NONWORK}$	0.0004	0.00124	0.5341	-0.3338	2144
$CAR[-5,5]_{WORK}$	0.0046	0.00129	0.6125	-0.5540	2883
$CAR[-5,5]_{NONWORK}$	0.0002	0.00156	0.6142	-0.3517	2144

The table above shows the cumulative abnormal returns for every type of work-relevancy of the tweet. CAR[-1,1] work is the cumulative abnormal return for tweets having a content containing work-relevancy between t-1 and t+1. CAR[-1,1] nonwork is the cumulative abnormal return fot tweets having a content containing non-work relevancy between t-1 and t+1. CAR[-3,3] work is the cumulative abnormal return for tweets having a content containing work-relevancy between t-3 and t+3. CAR[-3,3] nonwork is the cumulative abnormal return fot tweets having a content containing non-work relevancy between t-3 and t+3. CAR[-5,5] work is the cumulative abnormal return for tweets having a content containing work-relevancy between t-5 and t+5. CAR[-5,5] nonwork is the cumulative abnormal return fot tweets having a content containing non-work relevancy between t-5 and t+5. Mean is the average cumulative abnormal return of the corresponding work-relevancy. Std. Error is the standard error of the cumulative abnormal return. Max is the maximum cumulative abnormal return within the sample of the corresponding type of work-relevancy. Min is the minimum cumulative abnormal return within the sample of the corresponding type of work-relevancy. N is the amount of observations with the corresponding type of work-relevancy.

Table 10: T-test Average Abnormal Return per content for event window days

Date	AR_{WORK}	$T\text{-}test_{WORK}$	$AR_{NONWORK}$	$T\text{-}test_{NONWORK}$
-10	0.00074*	(1.69)	-(0.00015)	-(0.34)
-9	0.00175***	(4.08)	(0.00010)	(0.23)
-8	0.00088**	(2.19)	-(0.00021)	-(0.45)
-7	0.00049	(1.29)	(0.00068)	(1.33)
-6	0.00067*	(1.75)	(0.00050)	(1.02)
-5	0.00096**	(2.48)	-(0.00033)	-(0.77)
-4	0.00011	(0.28)	(0.00081)	(1.61)
-3	-0.00014	-(0.38)	(0.00029)	(0.64)
-2	0.00041	(0.94)	-(0.00010)	-(0.22)
-1	0.00103**	(2.55)	(0.00016)	(0.37)
0	0.00125***	(3.41)	(0.00083)	(1.63)
1	0.00011	(0.28)	-(0.00096)**	-(2.27)
2	0.00082**	(2.14)	-(0.00025)	-(0.51)
3	0.00030	(0.84)	(0.00050)	(1.12)
4	-0.00000	-(0.01)	-(0.00029)	-(0.67)
5	-0.00023	-(0.66)	-(0.00049)	-(1.13)
6	0.00019	(0.48)	-(0.00023)	-(0.53)
7	0.00073*	(1.89)	-(0.00030)	-(0.67)
8	0.00047	(1.24)	(0.00109)**	(2.31)
9	-0.00010	-(0.26)	-(0.00018)	-(0.27)
10	-0.00040	-(1.13)	(0.00138)***	(2.92)

This table reports the average abnormal return per event window date for work-related and non-work-related tweets with the corresponding T-test statistics. The event window consists of t-10, t+10 with 0 as the event date. AR_{WORK} is the average abnormal return for tweets with a work-related content. $AR_{NONWORK}$ is the average abnormal return for tweets with a non-work-related content. $T\text{-}test_{WORK}$ is the T value of work-related tweets at the corresponding event window date. $T\text{-}test_{NONWORK}$ is the T value of non-work-related tweets at the corresponding event window date. Absolute T-statistics are in parentheses. Significant at 1% (***), 5% (**), 10% (*) levels.

Table 11: T-test Cumulative Average Abnormal Return per content for different event windows

Period	CAR_{WORK}	$T\text{-test}_{CARWORK}$	$CAR_{NONWORK}$	$T\text{-test}_{CARNONWORK}$
-1, +1	0.00240***	(3.67)	0.00002	(0.03)
-3, +3	0.00378***	(3.67)	0.00046	(0.37)
-5, +5	0.00462***	(3.59)	0.00002	(0.11)

This table reports the average cumulative abnormal return for the event windows [t-1, t+1], [t-3,t+3] and [t-5,t+5] with respect to work-related and non-work-related tweets with the corresponding T-test statistics. CAR_{WORK} is the cumulative average abnormal return for tweets with a work-related content. $CAR_{NONWORK}$ is the cumulative average abnormal return for tweets with a non-work related content. $T\text{-test}_{CARWORK}$ is the T value of work-related tweets at the corresponding event window date. $T\text{-test}_{CARNONWORK}$ is the T value of non-work related tweets at the corresponding event window date. Absolute T-statistics are in parentheses. Significant at 1% (***), 5% (**), 10% (*) levels.

Table 12: T-test Average Abnormal Return per sentiment for event window days

Date	AR_{NEG}	$T\text{-test}_{NEG}$	AR_{POS}	$T\text{-test}_{POS}$	AR_{NEUT}	$T\text{-test}_{NEUT}$
-10	-0.00078	-(0.81)	0.00064	(1.02)	0.00040	(1.03)
-9	0.00199*	(1.84)	0.00068	(1.38)	0.00109***	(2.66)
-8	-0.00080	-(0.79)	-0.00029	-(0.56)	0.00103**	(2.54)
-7	0.00019	(0.19)	0.00030	(0.56)	0.00079*	(1.95)
-6	0.00114	(1.20)	0.00059	(1.00)	0.00051	(1.37)
-5	-0.00042	-(0.49)	0.00031	(0.57)	0.00061*	(1.66)
-4	0.00104	(1.20)	-0.00066	-(1.17)	0.00091**	(2.19)
-3	-0.00034	-(0.36)	-0.00075	-(1.41)	0.00056	(1.55)
-2	-0.00018	-(0.19)	-0.00042	-(0.73)	0.00060	(1.48)
-1	0.00151	(1.62)	0.00058	(1.10)	0.00056	(1.47)
0	0.00163**	(2.23)	0.00113*	(1.92)	0.00094**	(2.44)
1	-0.00101	-(1.20)	-0.00019	-(0.33)	-0.00031	-(0.82)
2	-0.00060	-(0.75)	0.00022	(0.40)	0.00617	(1.56)
3	-0.00191**	-(2.34)	0.00033	(0.62)	0.00081**	(2.33)
4	-0.00118	-(1.45)	-0.00039	-(0.74)	0.00020	(0.56)
5	-0.00058	-(0.79)	-0.00086*	-(1.66)	-0.00000	-(0.15)
6	0.00107	(0.97)	-0.00063	-(1.25)	0.00019	(0.51)
7	0.00138	(1.56)	0.00043	(0.84)	0.00002	(0.05)
8	0.00132	(1.50)	0.00012	(0.22)	0.00098***	(2.62)
9	-0.00113	-(1.31)	-0.00046	-(0.53)	0.00022	(0.59)
10	-0.00122	-(1.51)	0.00076	(1.51)	0.00040	(1.06)

This table reports the average abnormal return per event window date for positive, negative and neutral tweets with the corresponding T-test statistics. The event window consists of t-10, t+10 with 0 as the event date. AR_{NEG} is the average abnormal return for tweets with a negative sentiment. AR_{POS} is the average abnormal return for tweets with a positive sentiment. AR_{NEUT} is the average abnormal return for tweets with a neutral sentiment. $T\text{-test}_{NEG}$ is the T value of negative tweets at the corresponding event window date. $T\text{-test}_{POS}$ is the T value of positive tweets at the corresponding event window date. $T\text{-test}_{NEUT}$ is the T value of neutral tweets at the corresponding event window date. Absolute T-statistics are in parentheses. Significant at 1% (***), 5% (**), 10% (*) levels.

Table 13: T-test Cumulative Average Abnormal Return per sentiment for different event windows

Period	CAR_{POS}	$T\text{-test}_{CARPOS}$	CAR_{NEG}	$T\text{-test}_{CARNEG}$	CAR_{NEUT}	$T\text{-test}_{CARNEUT}$
-1, +1	0.00152	(1.57)	0.00213	(1.43)	0.00118*	(1.82)
-3, +3	0.00090	(0.60)	-0.00090	-(0.41)	0.00377***	(3.69)
-5, +5	-0.00069	-(0.37)	-0.00202	-(0.73)	0.00548***	(4.32)

This table reports the average cumulative abnormal return for the event windows [t-1, t+1], [t-3, t+3], [t-5,t+5] with respect to positive, negative and neutral tweets with the corresponding T-test statistics. CAR_{POS} is the cumulative average abnormal return for tweets with a positive sentiment. CAR_{NEG} is the cumulative average abnormal return for tweets with a negative sentiment. CAR_{NEUT} is the cumulative average abnormal return for tweets with a neutral sentiment. $T\text{-test}_{CARPOS}$ is the T value of positive tweets at the corresponding event window date. $T\text{-test}_{CARNEG}$ is the T value of negative tweets at the corresponding event window date. $T\text{-test}_{CARNEUT}$ is the T value of neutral tweets at the corresponding event window date. Absolute T-statistics are in parentheses. Significant at 1% (***), 5% (**), 10% (*) levels.

Table 14: The impact of sentiment on cumulative abnormal returns for [t-1, t+1]

Dependent variable: *Cumulative abnormal return [t-1, t+1]*

Variable	OLS(1)	OLS (2)	OLS(3)	OLS(4)	OLS(5)	OLS(6)
Intercept	0.03537	0.03825*	0.03679*	0.03944	0.03644*	0.04498**
	(1.61)	(1.74)	(1.67)	(1.79)*	(1.66)	(2.05)
Positive sentiment	0.00288**		0.00265***		0.00323***	
	(3.42)		(3.13)		(3.42)	
Negative sentiment		-0.00288***		-0.00265***		0.00356***
		-(3.42)		-(3.13)		(3.41)
Work relevant			0.00120**	0.00120**	0.00172***	0.00262***
			(2.33)	(2.33)	(2.70)	(4.91)
Neutral sentiment	-0.00063	-0.00351***	-0.00082	-0.00347***	0.00094	-0.00353***
	-(0.79)	-(6.72)	-(1.03)	-(6.65)	-(1.18)	-(6.78)
Positive*Work					-0.00138	
					-(1.38)	
Negative*Work						-0.01515***
						-(10.08)
Ln Market Cap	-0.00393***	-0.00393***	-0.00401***	-0.00401***	-0.00400***	-0.00468***
	-(4.14)	-(4.14)	-(4.23)	-(4.23)	-(4.21)	-(4.93)
P/B Ratio	0.00029***	0.00029***	0.00029***	0.00029***	0.00029***	0.00029***
	(5.11)	(5.11)	(5.11)	(5.11)	(5.03)	(5.12)
Year Fixed Effects	YES	YES	YES	YES	YES	YES
Industry Fixed Effects	YES	YES	YES	YES	YES	YES
CEO Fixed Effects	YES	YES	YES	YES	YES	YES
F-statistic	40.93	40.93	40.47	40.47	39.97	41.44
Sample Size	23382	23382	23382	23382	23382	23382

This table presents the impact of sentiment on the cumulative abnormal return in the event window [t-1, t+1) by means of mulitple OLS regressions. The dependent variable is the cumulative abnormal return of [t-1, t+1]. Each regression has different independent variables. Positive sentiment is the variable that represents a tweet with a positive sentiment. Negative sentiment is the variable that represents a tweet with a negative sentiment. Work relevant is the variable that represents a tweet with work relevant content. Neutral sentiment is the variable that represents a tweet with a neutral sentiment. Positive*Work is the variable that represents a tweet which has a positive sentiment and has work relevant content. Negative*Work is the variable that represents a tweet which has a negative sentiment and has work relevant content. Ln Market Cap is the variable that represents the logarithm of the market capitalization. P/B ratio is the variable that represents the price-to-book ratio. In all the regressions, Year, Industry and CEO fixed results are used. The total amount of observations is 23382. Absolute T-statistics are in parentheses. Significant at 1% (***), 5% (**), 10% (*) levels.

Table 15: The impact of sentiment on cumulative abnormal returns for [t-3, t+3]

Dependent variable : *Cumulative abnormal return [t-3, t+3]*

Variable	OLS(1)	OLS(2)	OLS(3)	OLS(4)	OLS(5)	OLS(6)
Intercept	0.20090***	0.20428***	0.19826***	0.20207***	0.19920***	0.21189***
	(5.98)	(6.07)	(5.90)	(6.01)	(5.93)	(6.31)
Positive sentiment	0.00339***		0.00381***		0.00243*	
	(2.64)		(2.94)		(1.69)	
Negative sentiment		-0.00339***		-0.00381***		0.00721***
		-(2.64)		-(2.94)		(4.52)
Neutral sentiment	-0.00142	-0.00481***	-0.00107	-0.00487***	-0.00078	-0.00498***
	-(1.17)	-(6.03)	-(0.87)	-(6.10)	-(0.63)	-(6.25)
Work relevant			-0.00224***	-0.00224***	-0.00349***	0.00027
			-(2.84)	-(2.84)	-(3.57)	(0.34)
Positive*Work					0.00330**	
					(2.17)	
Negative*Work						-0.02687***
						-(11.71)
Ln Market Cap	-0.01854***	-0.01854***	-0.01839***	-0.01839***	-0.01843	-0.01958
	-(12.79)	-(12.79)	-(12.68)	-(12.68)	-(12.71)	-(13.51)
P/B Ratio	0.00072***	0.00072***	0.00072***	0.00072***	0.00073***	0.00072***
	(8.33)	(8.33)	(8.34)	(8.34)	(8.44)	(8.36)
Year Fixed Effects	YES	YES	YES	YES	YES	YES
Industry Fixed Effects	YES	YES	YES	YES	YES	YES
CEO Fixed Effects	YES	YES	YES	YES	YES	YES
F-statistic	53.03	53.03	52.45	52.45	51.84	53.86
Sample Size	23382	23382	23382	23382	23382	23382

This table presents the impact of sentiment on the cumulative abnormal return in the event window [t-3, t+3) by means of mulitple OLS regressions. The dependent variable is the cumulative abnormal return of [t-3, t+3]. Each regression has different independent variables. Positive sentiment is the variable that represents a tweet with a positive sentiment. Negative sentiment is the variable that represents a tweet with a negative sentiment. Work relevant is the variable that represents a tweet with work relevant content. Neutral sentiment is the variable that represents a tweet with a neutral sentiment. Positive*Work is the variable that represents a tweet which has a positive sentiment and has work relevant content. Negative*Work is the variable that represents a tweet which has a negative sentiment and has work relevant content. Ln Market Cap is the variable that represents the logarithm of the market capitalization. P/B ratio is the variable that represents the price-to-book ratio. In all the regressions, Year, Industry and CEO fixed results are used. The total amount of observations is 23382. Absolute T-statistics are in parentheses. Significant at 1% (***), 5% (**), 10% (*) levels.

Table 16: The impact of sentiment on cumulative abnormal returns for [t-5, t+5]

Dependent variable : *Cumulative abnormal return [t-5, t+5]*

Variable	OLS(1)	OLS(2)	OLS(3)	OLS(4)	OLS(5)	OLS(6)
Intercept	0.07915*	0.09195**	0.07387*	0.08751**	0.07455*	0.09930**
	(1.82)	(2.11)	(1.70)	(2.01)	(1.71)	(2.28)
Positive sentiment	0.01280***		0.01364***		0.01252***	
	(7.68)		(8.13)		(6.70)	
Negative sentiment		-0.01280***		-0.01364***		-0.00041
		-(7.68)		-(8.13)		-(0.20)
Neutral sentiment	0.00461***	-0.00818***	0.00533***	-0.00831***	0.00557***	-0.00843***
	(2.94)	-(7.91)	(3.37)	-(8.02)	(3.50)	-(8.16)
Work relevant			-0.00449***	-0.00449***	-0.00550***	-0.00147
			-(4.39)	-(4.39)	-(4.35)	-(1.39)
Positive*Work					0.00267	
					(1.36)	
Negative*Work						-0.03225***
						-(10.83)
Ln Market Cap	-0.01067***	-0.01067***	-0.01038***	-0.01038***	-0.01040***	-0.01180***
	-(5.68)	-(5.68)	-(5.52)	-(5.52)	-(5.53)	-(6.28)
P/B Ratio	0.00192***	0.00192***	0.00192***	0.00192***	0.00193***	0.00192***
	(17.13)	(17.13)	(17.14)	(17.14)	(17.19)	(17.18)
Year Fixed Effects	YES	YES	YES	YES	YES	YES
Industry Fixed Effects	YES	YES	YES	YES	YES	YES
CEO Fixed Effects	YES	YES	YES	YES	YES	YES
F-statistic	51.82	51.82	51.43	51.43	50.79	52.54
Sample Size	23382	23382	23382	23382	23382	23382

This table presents the impact of sentiment on the cumulative abnormal return in the event window [t-5, t+5) by means of mulitple OLS regressions. The dependent variable is the cumulative abnormal return of [t-5, t+5]. Each regression has different independent variables. Positive sentiment is the variable that represents a tweet with a positive sentiment. Negative sentiment is the variable that represents a tweet with a negative sentiment. Work relevant is the variable that represents a tweet with work relevant content. Neutral sentiment is the variable that represents a tweet with a neutral sentiment. Positive*Work is the variable that represents a tweet which has a positive sentiment and has work relevant content. Negative*Work is the variable that represents a tweet which has a negative sentiment and has work relevant content. Ln Market Cap is the variable that represents the logarithm of the market capitalization. P/B ratio is the variable that represents the price-to-book ratio. In all the regressions, Year, Industry and CEO fixed results are used. The total amount of observations is 23382. Absolute T-statistics are in parentheses. Significant at 1% (***), 5% (**), 10% (*) levels.