Can Tweets predict intraday stock price movements?

The Honors Program
Senior Capstone Project
Student’s Name: Diwas Puri
Faculty Sponsor: Dave Louton
May 2017

Table of Contents

ABSTRACT

Twitter is a microblogging platform where over 320 million people post about things that matter to them in less than 140 characters. People post about their happiness, sadness, pride, disappointments, love, hate, expectations and other feelings. Nowadays, a popular trend has come up in which data from Twitter, can be gathered and analyzed in real time to see how people react to a particular situation including changes in stock markets. This research focuses on looking into the relationship, if any, between public mood from Twitter and the stock market returns by analyzing the Tweets about four major companies in consumer discretionary industry namely Amazon, Walt Disney, Home Depot and Comcast and their stock price over the duration of about a month. The research showed causal relationship between market sentiment corresponding to a particular stock and their stock returns.

INTRODUCTION

The rate of information dissemination and consumption has long played a crucial part in how a stock market performs. Good news drives a market upward while bad news leads to negative returns. Trading before the advent of computers used to be broker and dealer-centered whereby clients would go to brokers and dealers to buy stocks that they thought would do well. Today, in this day and age of information, the speed of communication has dramatically improved enabling all sections of the society to participate in the free market. This causes the data and chatter in the web and in the media. People post their status and pictures on Facebook, share posts with their loved ones and friends, Tweet about what is going on around them and search for information, which is literally at their fingertips. My topic which looks into the relationship between how people Tweet and how stock data perform is a hot topic in the financial market. Investment companies today such as Dataminr (Borzykowski) have started realizing the importance of information speed and have started using mediums such as Twitter, Facebook, and blogs to make judgments in their investments.

The examination of the relationship between Twitter mood and the stock market activity is vital for investors. They can make their best decision as to how to invest. The examination of the topics will involve the collection of Tweets for the duration of about one and half month, which will then be used to analyze mood and sentiment.

The motivation for this project came from the increasing popularity of big data and the rise of data science as an important skill in the workforce.

Background

Fama (1965) wrote about the relationship between past price charts and future price behaviors and it was one of the first papers in this field. Fama argued that when using a random-walk model there is no relationship between the past and the future price of the stock as the "successive price changes are independent, identically distributed random variables" (p.35). This meant that the price changes are not in any way related to its past behaviors.

Fama (1998) researched market efficiency, long-term monthly returns and behavioral finance demonstrating the idea that the market is efficient and that anomalies are random events.

Shiller (2003) challenged Fama's idea in the paper "From Efficient Markets Theory to Behavioral Finance" which stressed the various market anomalies in denying the efficient market theory. This paper reinforced the importance of remaining aware of the weaknesses of the efficient market theory and that prices do not always reflect genuine information (Shiller, 2003).

Antweiler and Frank (2004) gathered and analyzed 1.5 million messages posted on Yahoo! Finance and RagingBull and found stock messages helped predict stock market volatility although the effect is significant but small. This reaffirmed Shiller's random anomalies and confirmed the findings of Harris and Raviv (1993) about the idea that messages that are posted are not related to the increase in trading volume of companies (Antweiler & Frank, 2004).

Fama (1965, 1998), Shiller (2003), and Antweiler and Frank (2004) built the foundation of market behavior theory by showing that

1) past behavior and present behavior are related

2) prices do not always reflect genuine information and

3) market volatility and messages are related.

<u>Current state of the research</u>

Language and linguistic analysis are very important in behavioral finance, which seeks to study people's financial decision-making. Chen and Das (2007) developed a methodology to extract small investor sentiment from stock message boards. Their findings show the relationship between the stock values and the messages posted on the boards along with volatility and the volumes associated with the stock.

Not all stocks are of the same size and value. While companies like Google and Tesla have a huge footprint online, small companies with low market capitalization rarely make it to the online message boards. Baker studied this with Wurgler (2007) which found that such stocks were disproportionately sensitive to broad waves of investor sentiment.

Tetlock (2007) studied the daily content from a popular *Wall Street Journal* column and discovered that media pessimism predicts downward pressure on market prices. This further affirms that stock prices are related to market sentiments posted online.

Tetlock, Tsechansky, and Mackassy (2008) quantified language to measure a firm's fundamentals. They found that the linguistic media content captured hard to quantify aspects of firm's fundamentals, which quickly gets incorporated into stock prices. This helped make the study of market sentiment and stock prices easier as something as qualitative as language was converted to be used in forecasting models.

Bollen, Mao, & Zeng (2011) in the paper titled "Twitter mood predicts the stock market" published in the Journal of Computational Science sought to answer the question "Are public mood states derived from large-scale Twitter feeds correlated to the value of Dow Jones Industrial Average over time?" It used tools such as OpinionFinder and Google-Profile of Mood States to gather data about the mood states of the data gathered. Granger causality analysis and Self-Organizing Fuzzy Neural Network were used to investigate the mood states. This paper shed light on how to measure and quantify mood states from texts (Bollen, Mao, & Zeng, 2011).

Zhang, Fuehres and Gloor (2011) describe earlier works by doing a meta-analysis, which tried to predict stock market returns by analyzing Twitter posts. They gathered a randomized sample of Tweets over a period of 6 months and measured collective daily hope and fear indexes to check the correlation between these indices and stock market indicators. The study revealed that emotional Tweet percent displayed a significant correlation with the volatility index (Zhang, Fuehres, & Gloor, 2011).

Parsons, Garcia, Engelberg and Dougal (2012) found a relationship between financial reporting in *Wall Street Journal* and the stock market performance. The paper also studied the media's conditional effect whereby adding the effects of lagged returns increase the explanatory power by one-third (Dougal, Engelberg, García, & Parsons, 2012).

Arias, Arratia, and Xuriguera (2013) employed rigorous analytical methods to extract sentiment indicators from Tweets by building on research established by Chen and Das (2007) and Baker and Wurgler (2007). They differed by focusing on the stock market and movie box office revenue. Using the *summary tree* method to mine large datasets, they sought to improve the prediction power of the forecasting models (Arias, Arratia, & Xuriguera, 2013).

Yu, Duan, & Cao (2013) investigated the effects of social media and conventional media on the short-term firm stock market performances. By using the social media outlets and blogs of 824 publicly traded firms across 6 industries, they analyze the data gathered by using advanced sentiment analysis techniques that go beyond the number of mentions of each topic. This enables the authors to find the overall sentiment from each media source. Results suggest that social media has a strong correlation with the stock's short-term performance. They also find that the effect of different social media varies amongst each other. (Yu, Duan, & Cao, 2013).

Chen, De, Hu, & Hwang (2014) studied the investor opinions transmitted through social media to predict future stock returns. They analyzed the articles published in the social media about commentaries on finances. The paper found that the views expressed predict stock returns (Chen, De, Hu, & Hwang, 2014).

Sprenger, Andranik, Tumasjan, Sandner, and Welpe (2014) used the Twitter platform to study stock-related Tweets on a daily basis. They find an association between Twitter sentiment and stock returns. This further confirms the studies done earlier by researchers like Chen, De, Hu and Hwang.

Additionally, they presented a methodology for identifying a broad range of real-world news from Twitter by applying computational linguistics to a dataset of more than 400,000 Tweets related to the S&P 500 stocks. The authors find out that the returns prior to good news events are more pronounced than bad news events (Sprenger, Sandner, Tumasjan, & Welpe, 2014).

Ranco, Aleksovski, Caldarelli, Grčar, & Mozetič (2015) considered Twitter sentiment of 30 stocks from Dow Jones Industrial Average over the period of 15 months. From their analysis, they found low Pearson correlation and Granger causality between the corresponding time series over the entire time period. However, they find significant dependence between Twitter sentiment and the abnormal returns during the peak events (Ranco, Aleksovski, Caldarelli, Grčar, & Mozetič, 2015). This could mean that people react more when there is heavy trading.

In the paper titled "Stock market sentiment lexicon acquisition using microblogging data and statistical measures," the authors address the issue of Lexicon acquisition in sentiment analysis and presents a fast and novel approach for creating stock market lexicons. The authors make use of messages from StockTwits, a stock market microblog. The Twitter investor sentiment indicators generated are analyzed and show that the new microblogging indicators have a moderate correlation with popular investors intelligence (Oliveira, Cortez, & Areal, 2016).

This literature review has established that this area of interest is a new and growing field of study. Many scholars have attempted to study trends in Twitter to analyze real world events.

In fact, Derwent Capital, a UK-based hedge fund, in partnership with Bollen et al. began trading a $40 million hedge fund using the Twitter predictor. This is a growing field, which is being tapped for various uses and has a lot of potential to grow.

METHODOLOGY

The methodology I will be using comes directly from the research done by Sprenger, Andranik, Tumasjan, Sandner, and Welpe (2014). I used a cloud server to gather data. The data was then downloaded to a personal computer for analysis purposes. This was done so as to ensure the continual download of Tweet stream without exposure to extensive disruption in the download process.

My study differs with respect to the data being studied, which are the four main stocks of consumer discretionary industry. These stocks are namely Walt Disney (DIS), Comcast (CMCSA), Amazon (AMZN) and Home Depot (HD). While most of the research mentioned above such as the ones by Chen, De, Hu and Hwang and Sprenger, Andranik, Tumasjan, Sandner, and Welpe (2014) predict positive correlation between market sentiment and stock returns, I expect similar results with market sentiment being positively impacted by the stock performance and vice versa.

This study has shown that the potential of behavioral finance and computational informational analysis for investing is immense. If one realizes the benefit of such a relationship, one might be able to reap the benefits of data science and sentiment analysis of the social media and other relevant domains.

Company	**Index Weight**	**Market Capitalization**
Amazon.com Inc.	13.64%	404.08B
Home Depot Inc.	7.28%	178.91B
Comcast Corp A	7.14%	175.98B
Walt Disney Co	6.65%	177.7B

Table 1 Shows the top four companies in consumer discretionary industry (Yahoo! Finance)

In order to convert the textual Tweets into quantifiable numbers that can be measured and compared with price movements, it is important to understand how the two are going to be studied. First, the individual Tweets will be parsed to output their individual sentiment. Then the resulting Twitter sentiment with a time tag, will be merged with their corresponding stock price. We then wish to study if the stock price returns of the companies above either lead or lag the market sentiment and that how significant is the causal relationship if there is any.

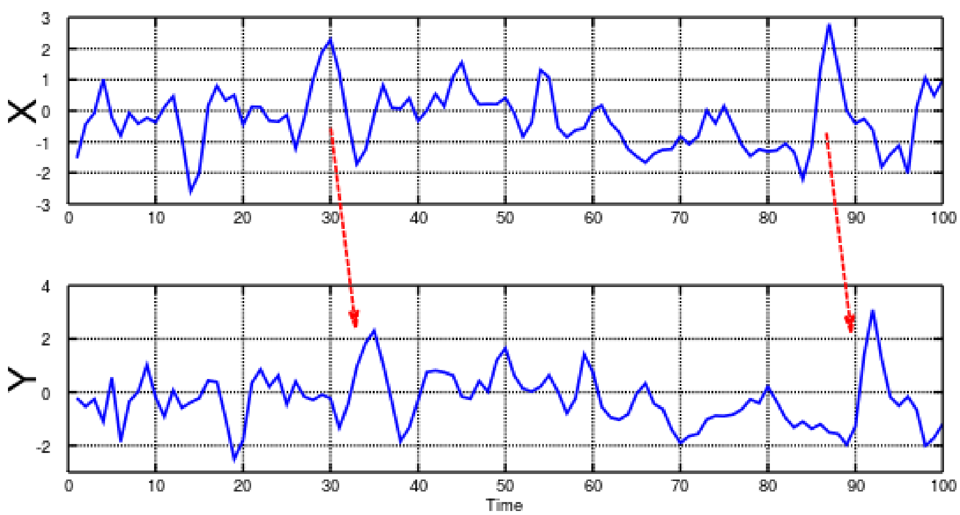


Figure 1 Illustrating granger causality between X and Y. ("Granger Causality Illustration")

Vader : Sentiment Analysis on Social Media Texts

VADER stands for Valence Aware Dictionary for sEntiment Reasoning (2014). It is a tool that is used extensively for analysing texts in the social media for sentiment gathering purposes.

The tool was used to find the sentiment polarity of the Tweets as to what degree of positive, negative, neutral a particular Tweet was. One of the primary motivations of using this tool was because it was very fast and already had in its arsenal, a gold standard corpus which was particularly suited to the social media domain. Please refer to Appendix A for some sample sentiment extraction from VADER.

Intraday stock price was gathered from Barchart onDemand ("Barchart OnDemand | Powerful Financial Market Data APIs That Are Easy to Use.") a Norwegian company that provides financial market data at a reasonable price.

timestamp	**tradingDay**	**open**	**high**	**low**	**close**	**volume**
2016-11-11 09:30:00-05:00	11/11/2016	96.43678	96.75434	95.39481	96.25816	296341
2016-11-11 09:31:00-05:00	11/11/2016	96.30779	96.33755	95.04749	95.53374	347672
2016-11-11 09:32:00-05:00	11/11/2016	95.53374	96.24824	95.53374	96.15893	306000

Table 2 showing sample stock data for Walt Disney (DIS).

Tweets

The Tweets were gathered from November 6, 2016 through December 22, 2016. Some sample Tweets about Walt Disney is as follows.

user_name	**followers_count**	**text**	**created_at**	**reTweet_count**
jordan	334	RT @javy23baez: Disney full of @cubs fans. #incredible #JB9	Sat Nov 05 20:57:07 +0000 2016	45
Joerg Iversen (US)	45	RT @Karla_garrison: What's Your Disney Princess IQ? https://t.co/oyqax2vWsr	Sat Nov 05 20:57:10 +0000 2016	10
sefiat 💗🔥💦	47	RT @ThatsSarcasm: Disney Channel made my childhood 😭 https://t.co/OWxZEF6dqx	Sat Nov 05 20:57:10 +0000 2016	2153

Table 3 showing sample Tweets about Walt Disney (DIS).

The Tweets were first cleaned to remove html links. After that *VADER* sentiment analysis was computed on them to output each Tweet's positive, negative and neutral sentiment polarity. Subsequently, a combined sentiment data encompassing the Twitter sentiment polarities and stock prices at their particular point in time were produced for final analysis. Please refer to Appendix A for further information on the extraction process.

The data was produced in 10 Minute interval data by averaging out the sentiments and the stock price for that particular bin. The sample combined price data and sentiment for Walt Disney is as follows.

date	pos	neg	neu	comp	tvol	open	high	low	close	volume
11/11/2016 14:30	0.075578	0.105905	0.818507	-0.05835	410	96.26516	96.51816	95.96244	96.26412	204611
11/11/2016 14:40	0.075882	0.079587	0.844531	-0.00397	407	96.97166	97.16119	96.86846	97.06644	195677
11/11/2016 14:50	0.12907	0.036876	0.834056	0.194341	355	97.14803	97.21876	96.99522	97.09318	118964
11/11/2016 15:00	0.092438	0.081955	0.825597	0.025472	536	96.95866	97.02922	96.8754	96.91967	67200.7
11/11/2016 15:10	0.083697	0.047052	0.869231	0.089875	446	96.59121	96.66585	96.52692	96.59073	60929.8

Table 4 showing the combined sentiment and stock data by 10 minute interval.

<u>The variables under consideration</u>

We were interested in a relationship between price returns, compound Twitter sentiment and Twitter volume.

The price returns were calculated by the following formula;

$$R_t = \frac{Close - Open}{Open}$$

The compound Twitter sentiment, as mentioned earlier, were extracted using VADER and it is normalized sum of the valence scores of each word in the VADER lexicon. It is also called 'normalized weighted composite score' and it provides us with a single unidimensional measure of sentiment as opposed to giving us various degrees of different sentiment.

The Twitter volume is simply the number of Tweets in each of the 10 minute interval.

<u>Test for Stationarity and Optimal Lag Length</u>

The Augmented - Dickey Fuller (ADF) test performed for stationarity check on the variables for the four stocks under consideration is as follows.

The test for stationarity reveals that all of the variables under consideration are stationary at level. For the four stocks, an optimal lag length was found for Vector Auto Regression by using likelihood ratio (LR) test as opposed to other tests such as Final Prediction Error, Akaike information criterion, Schwarz information criterion and Hannan-Quinn information criterion which were giving suspiciously small lag lengths.

Vector Auto Regression is used to capture the relationship between a variable and its past lagged values. A k^{th} order VAR with is as shown below;

$$y_t = c + A_1 y_{t-1} + A_2 y_{t-2} + \cdots + \ + A_k y_{t-k} + \ e_t$$

Where k = number of lags, t = time

	comp	tvolume	return	Optimal Lag Length
AMZN	✓	✓	✓	21 1p1f
CMCSA	✓	✓	✓	21 p
DIS	✓	✓	✓	16 f
HD	✓	✓	✓	28 p

Table 5 showing stationarity test on each of the variables for the four concerned stocks.

Test for Granger Causality

Granger Causality test was applied to the VAR model to see if the two independent variables namely Twitter volume and sentiment caused stock returns. 5% level of significance was chosen. The results are as follows for each of the stocks.

DIS - Walt Disney

VAR Granger Causality/Block Exogeneity Wald Tests	
Date: 03/21/17 Time: 22:44	
Sample: 11/06/2016 00:10 12/23/2016 23:50	
Included observations: 656	

Dependent variable: RETURN			
Excluded	Chi-sq	df	Prob.
TVOLUME	7.049383	16	0.9723
COMP	19.21332	16	0.2578
All	27.9097	32	0.6738

For DIS stock, Twitter volume and sentiment doesn't appear to granger cause returns.

<u>AMZN - Amazon.com, Inc.</u>

VAR Granger Causality/Block Exogeneity Wald Tests			
Date: 03/21/17 Time: 22:28			
Sample: 11/06/2016 00:10 12/23/2016 23:50			
Included observations: 960			
Dependent variable: RETURN			
Excluded	Chi-sq	df	Prob.
TVOLUME	16.44075	6	0.0116

COMP	3.230739	6	0.7794
All	19.89937	12	0.069

For AMZN stock, Twitter volume seems to granger cause return but overall, the null hypothesis that Twitter volume and compound sentiment does not granger cause stock returns is not rejected.

CMCSA - Comcast

VAR Granger Causality/Block Exogeneity Wald Tests			
Date: 03/21/17 Time: 22:42			
Sample: 11/06/2016 00:10 12/23/2016 23:50			
Included observations: 340			
Dependent variable: RETURN			
Excluded	Chi-sq	df	Prob.
TVOLUME	35.23817	21	0.0266
COMP	35.34465	21	0.0259
All	71.82076	42	0.0028

HD - Home Depot

VAR Granger Causality/Block Exogeneity Wald Tests			
Date: 03/21/17 Time: 22:46			
Sample: 11/06/2016 00:10 12/23/2016 23:50			
Included observations: 198			
Dependent variable: RETURN			
Excluded	Chi-sq	df	Prob.
TVOLUME	42.77556	28	0.0366
COMP	50.77923	28	0.0053
All	80.19432	56	0.0187

For CMCSA and HD stocks, Twitter volume and compound sentiment appears to granger cause returns.

DISCUSSION AND CONCLUSION

While it is not conclusive, there appears to be a causal relationship between Twitter sentiment and Tweets volume for a particular stock and stock price returns as indicated by the results shown above. This is with 5% statistical significance level and shouldn't be confused with economic significance. We attribute our discrepancy in results to crude data gathering process and the fact that only 10% of the real Tweets were gathered via Twitter's streaming api (see Appendix B for number of Tweets gathered by company).

It is recommended that the future researchers choose a wider pool of companies to study this relationship. Improvement on sentiment analysis tools might also lead to better results. Also, for better results, it might be wise to choose companies that have no ambiguity in their names so as to avoid noise in the data.

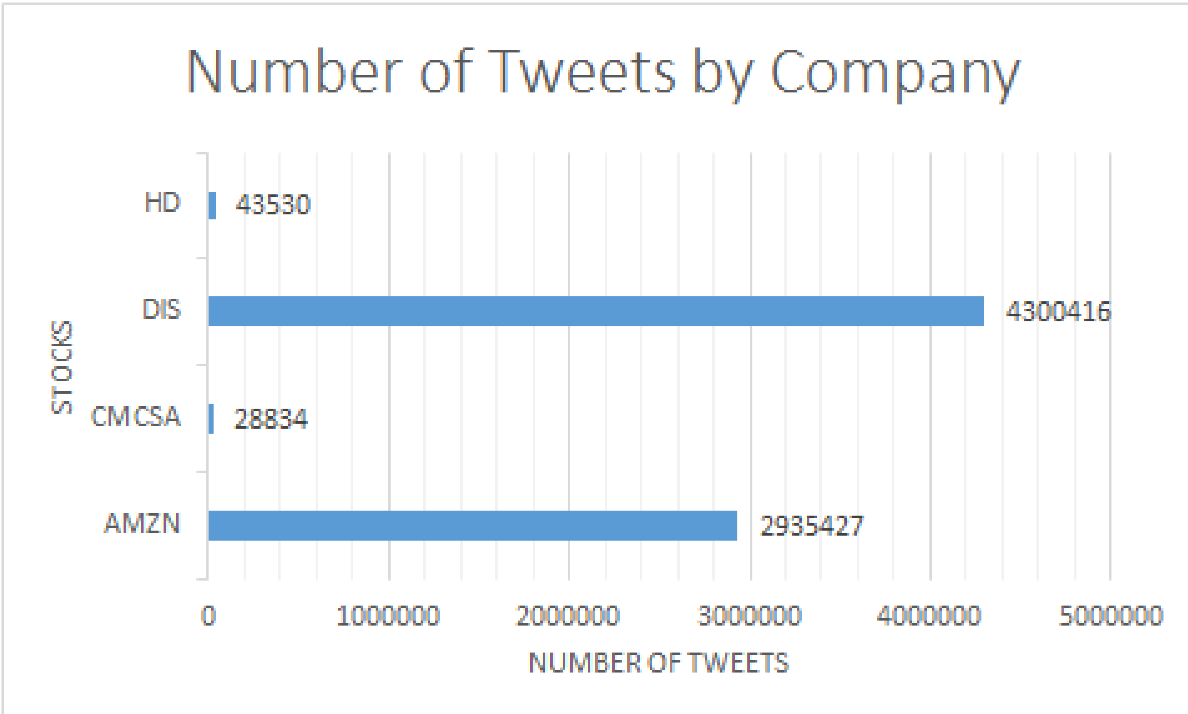
APPENDICES

Appendix A – Sentiment extraction using VADER

The sentiment extraction using VADER provides us with the 'pos', 'neu' and 'neg' scores and are the proportions of text that fall in each category. Thus, in the sum of their scores is 1. The compound score is the most important score in our case as it is provides us with a single measure of how much of sentiment for a Tweet and it ranges from -1 to +1 (2014).

```
VADER is smart, handsome, and funny.----------------------------- {'neg': 0.0,
'neu': 0.254, 'pos': 0.746, 'compound': 0.8316}

VADER is not smart, handsome, nor funny.------------------------- {'neg': 0.646,
'neu': 0.354, 'pos': 0.0, 'compound': -0.7424}

VADER is smart, handsome, and funny!----------------------------- {'neg': 0.0,
'neu': 0.248, 'pos': 0.752, 'compound': 0.8439}

VADER is very smart, handsome, and funny.------------------------ {'neg': 0.0,
'neu': 0.299, 'pos': 0.701, 'compound': 0.8545}

VADER is VERY SMART, handsome, and FUNNY.------------------------ {'neg': 0.0,
'neu': 0.246, 'pos': 0.754, 'compound': 0.9227}

VADER is VERY SMART, handsome, and FUNNY!!!---------------------- {'neg': 0.0,
'neu': 0.233, 'pos': 0.767, 'compound': 0.9342}

VADER is VERY SMART, uber handsome, and FRIGGIN FUNNY!!!--------- {'neg': 0.0,
'neu': 0.294, 'pos': 0.706, 'compound': 0.9469}

The book was good.----------------------------------------------- {'neg': 0.0,
'neu': 0.508, 'pos': 0.492, 'compound': 0.4404}

The book was kind of good.--------------------------------------- {'neg': 0.0,
'neu': 0.657, 'pos': 0.343, 'compound': 0.3832}

The plot was good, but the characters are uncompelling and the dialog is not
great. {'neg': 0.327, 'neu': 0.579, 'pos': 0.094, 'compound': -0.7042}

At least it isn't a horrible book.------------------------------- {'neg': 0.0,
'neu': 0.637, 'pos': 0.363, 'compound': 0.431}

Make sure you :) or :D today!------------------------------------ {'neg': 0.0,
'neu': 0.294, 'pos': 0.706, 'compound': 0.8633}
```

Appendix B – Number of Tweets gathered by Company



REFERENCES

Antweiler, W., & Frank, M. Z. (2004). Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards. *The Journal of Finance*, *59*(3), 1259–1294.

Arias, M., Arratia, A., & Xuriguera, R. (2013). Forecasting with Twitter data. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *5*(1), 8.

Baker, M., & Wurgler, J. (2007). *Investor Sentiment in the Stock Market* (Working Paper No. 13189). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w13189

Borzykowski, Bryan. "How Investors Are Using Social Media To Make Money". *CNBC*. N.p., 2016. Web. 26 Sept. 2016.

Chen, H., De, P., Hu, Y. (Jeffrey), & Hwang, B.-H. (2014). Wisdom of Crowds: The Value of Stock Opinions Transmitted Through Social Media. *Review of Financial Studies*, *27*(5), 1367–1403. https://doi.org/10.1093/rfs/hhu001

Darbar, S. M., & Deb, P. (1995). Does trading volume have a unit root?. *Applied Economics Letters*, *2*(5), 144-147.

Das, S. R., & Chen, M. Y. (2007). Yahoo! For Amazon: Sentiment Extraction from Small Talk on the Web. *Management Science, 53*(9), 1375–1388.

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

Fama, E. F. (1965). The Behavior of Stock-Market Prices. *The Journal of Business*, *38*(1), 34–105.

Gilbert, C. H. E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. In *Eighth International Conference on Weblogs and Social Media*

(ICWSM-14). Available at (20/04/16) http://comp. social. gatech. edu/papers/icwsm14. vader. hutto. pdf.

Granger Causality Illustration. (n.d.). Retrieved March 28, 2017, from https://upload.wikimedia.org/wikipedia/commons/thumb/7/7d/GrangerCausalityIllustration.svg/640px-GrangerCausalityIllustration.svg.png?1490672210359

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

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

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

Sprenger, T. O., Tumasjan, A., Sandner, P. G., & Welpe, I. M. (2014). Tweets and Trades: the Information Content of Stock Microblogs. *European Financial Management*, *20*(5), 926–957. https://doi.org/10.1111/j.1468-036X.2013.12007.x

Sprenger, T. O., Sandner, P. G., Tumasjan, A., & Welpe, I. M. (2014). News or Noise? Using Twitter to Identify and Understand Company-specific News Flow. *Journal of Business Finance & Accounting*, *41*(7/8), 791–830. http://doi.org/10.1111/jbfa.12086

Tetlock, P. C. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of Finance*, *62*(3), 1139–1168. https://doi.org/10.1111/j.1540-6261.2007.01232.x

Tetlock, P. C., Saar-Tsechansky, M., & Macskassy, S. (2008). More Than Words: Quantifying Language to Measure Firms' Fundamentals. *The Journal of Finance*, *63*(3), 1437–1467. https://doi.org/10.1111/j.1540-6261.2008.01362.x

Yahoo Finance - Business Finance, Stock Market, Quotes, News. (n.d.). Retrieved March 28, 2017, from https://finance.yahoo.com/

Yu, Y., Duan, W., & Cao, Q. (2013). The impact of social and conventional media on firm equity value: A sentiment analysis approach. *Decision Support Systems*, *55*(4), 919–926. http://doi.org/10.1016/j.dss.2012.12.028

Zhang, X., Fuehres, H., & Gloor, P. A. (2011). Predicting Stock Market Indicators Through Twitter "I hope it is not as bad as I fear." *Procedia - Social and Behavioral Sciences*, *26*, 55–62. http://doi.org/10.1016/j.sbspro.2011.10.562