# The Effects of Donald Trump's Tweets on The Stock Exchange

Yardley, Ben

20 March 2020

# Dissertation

## The Effects of Donald Trump's Tweets on The Stock Exchange


Ben Yardley - 16008349


20th March 2020

BSc (Hons) Economics

**Word Count: 8201**

This dissertation is submitted in part fulfilment of the requirements of the BSc (HONS) Economics in the Department of Accounting, Finance and Economics at Manchester Metropolitan University.

## DECLARATION/ COPYRIGHT:

This dissertation is an original and authentic piece of work produced in fulfilment of my degree regulations. I have fully acknowledged and referenced all secondary sources. The dissertation has not been submitted in whole or part of assessment in another unit at this or any other university.

Signature: B Yardley

Date: 20/03/20

# ABSTRACT

Donald Trump is a huge personality in a unique social, political and financial situation. Through twitter he is able to influence his following instantly with no regulation, often with large unforeseen repercussions.

The aim of this report is to identify Trump's company specific tweets, quantify the mood and feeling of the tweets through sentiment analysis and then investigate the correlation between this and the effect on the stock prices of these companies. This will be done through a regression of the sentiment and the returns of the company over serval time periods. Further investigation will the take place in order to analyse how neutrally classified tweets are impacted, as well as the effect of the number likes and retweets a tweet has on the stock price of a company.

# ACKNOWLEDGEMENTS

Firstly, I would like to thank my dissertation supervisor Lefteris Giovanis for all his help throughout the year on this project. We have had many endless conversations on this topic which I believe we have both become very passionate about. Without his guidance and support, this report would not have been produced to the same standard, of which I am very proud.

I would also like to thank Gavin Brown, though not my dissertation supervisor, he also provided amazing and useful insight that helped to shape this paper.

Finally, I would like to thank my family, friends and girlfriend. Their help and support both direct and indirect made my work possible; and for that I am very grateful.

# TABLE OF CONTENTS

# 1. INTRODUCTION

On the 20th of January 2017 Donald Trump was inaugurated as the 45th President of the United States. He is a highly outspoken and vocal president; and although he divides opinion, his influence over both his admirers and critics is undeniable. The extent of this influence was highlighted when he was published in the Time100: Most Influential People in the World 2019 shortlist (TIME, 2019).

Twitter is a social media platform that began in 2006, the platform has over 330 million active monthly users, this number has grown consistently since its launch with a nearly 30% increase in the last 5 years (Statista, 2019).

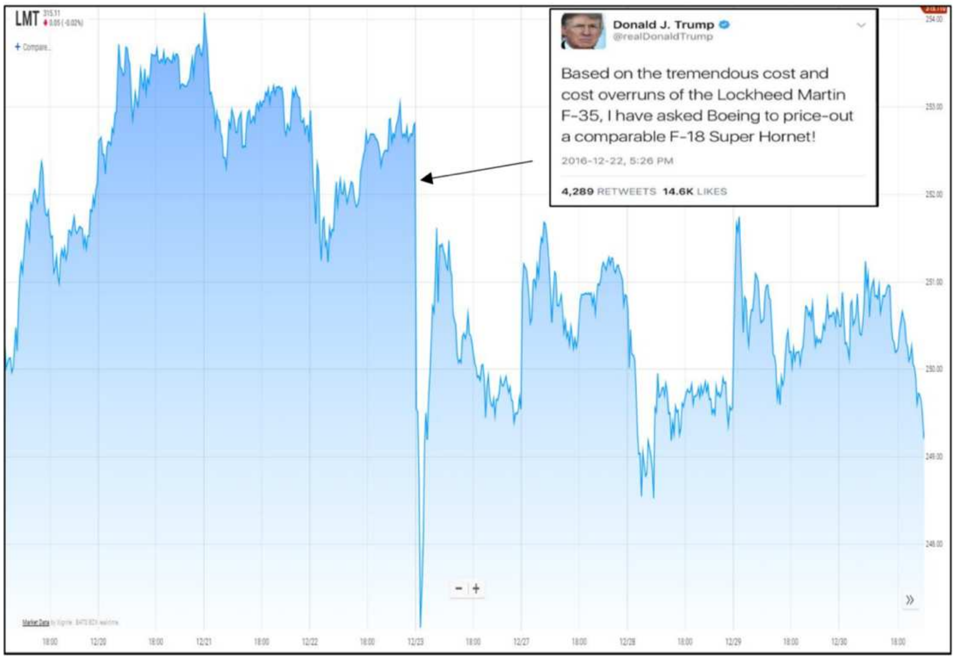
Trump is an avid Twitter user, using the site daily for everything from expressing his personal beliefs, to trying to block international trade deals and intimidate American corporation owners. By cross-referencing the Time100: Most Influential People in the World 2019 shortlist (TIME, 2019), with a list with a list of the most followed users on twitter (FriendorFollow, 2019); there are five individuals on both lists, see below.

| Name | Twitter Followers | Total Tweets |
|---|---|---|
| Taylor Swift | 86 million | 434 |
| Lady Gaga | 81 million | 9,110 |
| Donald Trump | 74 million | 49,910 |
| Ariana Grande | 72 million | 46,106 |
| LeBron James | 45 million | 6,911 |

Table 1: The Most Followed and Influential Individuals 2019

Of these individuals, Donald Trump and Ariana Grande are the only ones comparable in terms of followers and total tweets. Also, Trump is the by far the most financially and politically motivated, and most likely to be able influence traders in the opinion in the author; for these reasons Trump will be investigated in this report. There is however scope for investigation of the other individuals highlighted above.

Trump is able to impact beliefs directly through the platform, without any mandatory regulation and instantly connect to his 67 million followers (Twitter, 2019). His current twitter position is unique, he uses twitter daily and despite his influence, position and power, he does not separate the subjective from fact in his postings online. Here is an example of how one tweet can have an instant and significant impact in both the short and long term.




Source: (Rayarel, 2018)

Due to the above tweet the share price of Lockheed Martin fell 2% by the start of the next trading day, a total Market Cap depreciation of $1.2 billion. (Rayarel, 2018)

JPMorgan Chase have highlighted the significance of these movements through their creation of the VOLFEFE Index. This index is a measure of the volatility of US treasury bond yields, as affected by Trumps tweets (Bloomberg, 2019). Similarly, Citigroup stated that around 10% of the president’s tweets since his November 2016 election refer to subjects of importance to U.S. markets (Citigroup, 2019).

## 1.1 Research Question

What effects do Donald Trump's Tweets have on the stock exchange?

## 1.2 Research Aim

To analyse both the tweets pertaining to large corporations, and the effect this has on the share prices of these corporations.

## 1.3 Research Objectives

- To understand the possible influence of tweets on financial markets; consolidating an understanding of how this occurs in line with EMH assumptions.
- To appropriately quantify these effects in order to set a framework by which tweet related financial analysis can be carried out.
- To analyse the effects on stock prices, using the tweets and identifying the indicated significance of a relationship between a positive tweet and a positive increase in stock prices.
- To investigate the effects of likes and retweets on the tweets pertaining to companies and the relationship this has to the change of associated stock prices.

## 1.4 Rational

The author of this paper is a potential future trader. By understanding the effects these tweets have on financial markets and trading decisions. Using this understanding to form expectations in the present, the correct financial decisions in order to make profits and avoid losses can be made. Due to the increasing usage of twitter, it would not be unexpected if this was a more common occurrence in the future.

# 2. LITERATURE REVIEW

The Efficient market Hypothesis is a fundamental financial economic theory coined by Eugene Fama in the 1970s; it states that stocks always trade at their fair market price (Fama, 1970). This theory is built on the assumptions that Stocks are representative of all available information. Information includes not only what is currently known about a stock, but also any future expectations, such as earnings or dividend payments. Any new information is instantly communicated to all traders and all traders react in a rational and uniform way. The supply and demand of a stock is then instantly impacted, rationally and to the identical magnitude of that of the new information (Fama, 1998).

There are 3 forms or levels of the efficient market hypothesis that differ in what information is considered:

In the weak form, only past market trading information, such as stock prices, trading volume, and short interest are considered (Fama, 1970).

The semi-strong form extends the information to public information other than market data, such as news, accounting reports, company management, patents, products of the company, and analysts' recommendations. This is the most realistic of the forms as it encompasses the most realistic set of information that is available to traders through the utilisation of thorough research and sophisticated trading applications (Naseer & Tariq, 2015).

The strong form extends the information further to include not only public information, but also private information, typically held by the those within a corporation in strategic roles or directly involved in business developments (Malkiel, 1989). Examples of this could be potential mergers, new product releases or any other news that is held within the business before public release. This means that even the most private and newly created information would be available publicly; and thus, impact trading decisions instantly (Yen & Lee, 2008). These are very strong and highly unrealistic assumptions as in practise, insider trading is the only instance by which private information has impact on financial markets. This act is illegal and surveyed the Securities and Exchange Commission (SEC).

The direct implication of this framework is that it is impossible for traders to consistently make supernormal profits, and thus “beat the market” (Fama, 1970). Believers of EMH would argue that there are no undervalued stocks, and that aiming to identify these through valuation, market timing or analysis of trends is arbitrary. Under these assumptions the only way to make higher returns is to take on additional risk (Malkiel, 1989).

Random walk theory states that in the short term, the prices of stock have no memory; and therefore, previous price movements cannot be used in order to predict future prices (Horne & Parker, 1967). In this model price changes occur with the same distribution but in an entirely unpredictable way.

Mathematically, one dimensional random walk theory demonstrates how an array of results can be possible, without any causal affect creating the result predictably. Assume the equation:

$$P_{t+1} = P_t + \mu + \varepsilon_{t+1}, \qquad \varepsilon \sim IID(0, \sigma^2)$$

This model explains how in the short run, the future price ($P_{t+1}$); is the sum of the current price ($P_t$), the expectation of price change, and an error term ($\varepsilon$), or size of the random walk. $IID(0, \sigma^2)$ denotes that the term $\varepsilon$ is independently and identically distributed, with a mean of 0 and a variance of $\sigma^2$. This interdependence mathematically demonstrates how changes in price are uncorrelated and independent of any predictable effect (Alrabadi, 2010).

Much like the efficient market hypothesis, this assumes price changes of a stock are completely independent of any previous changes or trends. Even under efficient market hypothesis assumptions, there are variances in the prices of stocks in the short term that cannot be explained by rational analysis. Random walk theory adds colour to the possibility of variance in the price of stocks in the short term within the acceptance of the efficient market hypothesis (Fama, 1995).

The idea of a trader not being able to beat the market is also built on; short-term buy-hold trading strategies are entirely ineffective within this structure. The short-term future value of a stock is the sum of its value and expectations plus a random walk. In an efficient market, the only way to know beat the market is through private information (Cheng & Deets, 1971).

Despite the Strong logical approach of the efficient market hypothesis, it has many critics who claim it does not stand true in the real world. Behavioural economists disagree with the EMH framework, they believe that stocks are not only impacted by information relating to the true value of a company, but also by perception and opinion. That traders are likely to both over and under react to new information; and that the herd mentality of traders can over affect the prices of stock to remain above or below that of the true market value (Malkiel, 2003).

A deeper look into this reveals how the effect of new information hugely depends on the quality of delivery of information and the impact this has on an individual basis. This means that a bigger impact will result from less accurate information, that is well reported, from a reputable source. Compared to correct information that is poorly reported from a less reputable source. Persuasion bias also states that an individual may be influenced by the number of times they hear the same information and that this can be confused for the value or truth of news (DeMarzo, et al., 2003).

Herd mentality theory states that as the price of a stock experiences a significant change, traders will tend to ignore their own beliefs and any of the other information available to them. And instead of acting as an individual would under EMH assumptions; follow the decisions of other traders without rational thought (Shantha, 2018). This can be caused by information from a completely uninformed backing. The Warren Buffet effect explains this, where due to his consistent success, when he announces his buying or selling a position the price of this stock is over inflated hugely in the associated direction to his announcement. There is also asymmetry between the impacts of herd mentality in an up market compared to a down market (Dang & Lin, 2016). Where there is consistent financial prosperity, traders are far more likely to be influenced by the decisions made by others and be less critical in their decision making. This is thought to be due to an improvement to trader rationale in recessionary periods, due to an increase in the importance of decisions (Dang & Lin, 2016).

Studies have shown how Twitter can be used in order to capture the attitudes and traders and consumers. This is done through sentiment and mood analysis tools, which look at a live feed of all tweets over a period. The findings of this analysis have then been used to give a numeric value for the “twitter mood” on a given day. (Bollen & Mao, 2011) This is then correlated against the movements of a broad stock market index.

The findings of this report concluded that the utilisation of these tools can be effective in numerically summarising qualitative or subjective matters. It also found that there is a strong positive relationship between the positivity of the “twitter mood” and an increase in the broad stock market index (Bollen, et al., 2011).

It is worth note that these findings could be falsely representative. Reverse causality would mean that instead of the twitter mood being able to predict the movements of the stock market, it is the movements of the stock market that will impact the postings of individuals on twitter. This study also doesn’t consider peoples incentive or likelihood to post on social media when comparing the negative to positive. If, for example, individuals are more likely to post negative thoughts and feelings. Then this would skew the findings of the twitter mood value and effect the conclusions of the report (Chen & Lazer, 2013).

Another study took postings on a yahoo finance stock message board; and used the sentiment of the posts in order to predict the price movements of associated stocks. The intention of this study is logically sound, however it concluded that due to the high volumes of spam and subjectivity in these message boards. There was no significant correlation between the sentiment of the postings and stock price movements (Rechenthin, et al., 2013).

While carrying out this study it was highlighted how individuals post false news in order to artificially inflate the price of a stock they wish to sell, this trading method is coined as a “pump and dump”. Herd mentality also contributes to this as some traders believe the news and buy the stock increasing the price, however other traders will see the increase in the price of this stock and further inflate it above the rational level of EMH assumptions.

A famous case of “pump and dump” involved 15-year old high school student Jonathan Lebed, who would purchase stocks, and then send spam messages on Yahoo message boards on the same day to inflate the value of these stocks. His six-month trading profits amounted to $800,000 (Rechenthin, et al., 2013).

The growing relevance of this study is shown, as according to Aite Group, a financial services consulting company, as of the end of 2010, 35% of professional trading firms were exploring the use of sentiment analysis in their models, up from 2% in 2008 (Bowley, 2010).

## 2.1 Contribution

From a review of the literature and research carried out so far, it is not clear as to any investigation that has been carried out on one individual’s tweets and the effect this has on the stock exchange. This research aims to do just this and provide insight of the effects that twitter can have when powerful and influential individual posts to a huge audience.

# 3. DATA COLLECTION & METHODOLOGY

## 3.1 Data Collection

In order to test the effects Donald Trump's tweets have on the company's he tweets about; firstly, the tweet data must be collected. Trump Twitter Archive is a website which provides a database of all Donald Trump's Tweets and is updated live. There is additional information available on this site, including the number of likes and retweets. In this analysis, all Donald Trump's tweets since his inauguration on the 20th of January 2017 will be used, this totals 12,835 tweets as original data sample (trumptwitterarchive, 2019).

These tweets will be used alongside a list of S&P 500 companies downloaded from Datahub (Datahub, 2019). This list is used along with *Formula 1* (see appendix). This formula returns a 1 if the entire content of the company name is found anywhere in the tweet, and a 0 if it is not. Next, *Formula 2* (see appendix) will be used; this returns the name of the company that returns a 1 from the previous formula with the matrix of tweets by companies. This allows the company specific tweets to be identified, and also the specific company to be returned which is being tweeted about.

However, this original search only captures tweets which have the entirety of a company's official name within them. This causes difficulties as Trump is unlikely to tweet this full name every time he references a company, for example, explicitly mentioning "Incorporated" or "Corporation". Due to this, a second search term is used, the simplified or shorthand version of the company name, using the following *Formula 3* (see appendix).

See the examples below:

| Full Name | Shorthand Name |
|---|---|
| Amazon.com Inc. | Amazon |
| JPMorgan Chase & Co. | JPMorgan |
| Visa Inc. | Visa |

Table 2. Shorthand Name Example

This method however raises its own problems as generic company names that are searched for could incorrectly flagged as a company specific tweet. Examples of this are a tweet which contains “Amazon” being flagged as a company specific tweet about Amazon when it may be about the Amazon rainforest. There is no remedy to this problem in excel so the tweets that are found will be checked one by one to ensure they are company specific.

There are also cases where Trump will mention more than one company in a tweet. Where this happens, *Formula 3* (see appendix) will return two number ones in the matrix. *Formula 2* (see appendix) will return the name of the first company alphabetically in the tweet. This is a limitation of the formula; however, a measure is in place to highlight tweets with more than one company in them, and where this occurs I will manually duplicate the tweet, assigning one to each of the companies mentioned, becoming two separate observations.

With the specific tweets and companies that have been tweeted about identified, they become each individual observation of this report. For each of the observations, the tweets will be taken and a score for the sentiment will be calculated. Furthermore, a percentage change in the share price as a result of the tweet will be found.

Sentiment analysis involves the computational processing of a body of text of any length in order to give a numeric score that captures the level of positivity or negativity. Sentiment refers to the mood or feeling of text; objectivity refers to how opinion based it is. The scores are given through the use of natural language processing, text analysis, computational linguistics, and biometrics to systematically identify, extract and quantify affective states and subjective information.

A sentiment analysis tool from MonkeyLearn will be used (MonkeyLearn, 2020). This tool takes anybody of text and returns a sentiment classification; positive, neutral or negative. And also, a confidence score given a percentage of how certain the tool is in its sentiment classification. See below a table with generic statements and the results that are given by the MonkeyLearn sentiment analysis tool:

| Statement | Classification | Confidence % |
|---|---|---|
| I like you | Positive | 76.6% |
| I don't have an opinion on you | Neutral | 75.1% |
| I hate you | Negative | 96.4% |

Table 3. Sentiment Example

This shows how the classification and confidence scores are given. Also see how the confidence score is higher for “I hate you” compared to “I like you” as the statement is stronger.

In order to appropriately measure the impact of the tweets, the change in the share prices of the companies that are tweeted about must be quantified. In order to do this, we will collect historical stock price information from Yahoo finance, and the following formula will be used:

$$\Delta P_{c,t} = \frac{P_{t+n} - P_{t-n}}{P_{t-n}}$$

$\Delta P_{c,t}$ is the change in stock price, for company $c$, at the day of the tweet $t$. It is calculated as the percentage change between the adjusted closing price $P$ for company from $n$ number of days after the tweet $t$. Minus the adjusted closing price $P$ for company from $n$ number of days after the tweet $t$. Divided by the adjusted closing price $P$ for company from $n$ number of days after the tweet $t$.

This however does not isolate the effect of the tweet as it is the total price movement of the stock and will also include any broad market movements which may also be impacting the stock price. To remedy this problem, the tweet specific return on the share price of a company will be classified as:

$$R_{c,t} = \frac{(P_{t+n} - P_{t-n})/P_{t-n}}{\beta_c} - (P_{t+n} - P_{t-n})/P_{t-n}$$

This can be simplified to:

$$R_{c,t} = \frac{\Delta P_{c,t}}{\beta_c} - \Delta I_t$$

Which is the change in share price $\Delta P_{c,t}$, divided by the Beta of that company $\beta_c$ and minus the change in the S&P 500 market index adjusted close price $\Delta I_{c,t}$. This adjusts the returns of the stock price to amplify or contract them by the beta of that company, and then subtracts the movement of the market. This will be the returns values used in the analysis of this report.

The Beta of a company ($\beta_c$) is a standard financial measure with an average value of 1 that is used to approximate the riskiness of a company in comparison to the market. It is calculated as the covariance between the returns of the company and the market divided by the variance of the market:

$$\beta_c = \frac{Cov(R_c\ , R_m)}{Var(R_m)}$$

This is used in the tweet specific returns formula as a catalyst to amplify or contract returns. This would mean that a unrisky company with a Beta value of 0.4 for example, would have its daily returns divided by 0.4 which would increase them, as this company is generally involatile, so returns would normally be lower and less erratic.

Note below an example of tweets, the associated sentiment classification and confidence. As well as the returns, this is the standard information across all of the data in this report.

| Tweet | Classification | $Sen_{c,t}$ | $R_{c,t}$ |
|---|---|---|---|
| General Motors which was once the Giant of Detroit is now one of the smallest auto manufacturers there. They moved major plants to China BEFORE I CAME INTO OFFICE. This was done despite the saving help given them by the USA. Now they should start moving back to America again? | Negative | -0.617 | -2.91% |
| I was just informed by Marillyn Hewson CEO of Lockheed Martin of her decision to keep the Sikorsky Helicopter Plant in Coatesville Pennsylvania open and humming! We are very proud of Pennsylvania and the people who work there. Thank you to Lockheed Martin one of the USA's truly great companies! | Positive | 0.987 | 0.43% |
| Harley-Davidson should stay 100% in America with the people that got you your success. I've done so much for you and then this. Other companies are coming back where they belong! We won't forget and neither will your customers or your now very HAPPY competitors! | Negative | -0.614 | -2.65 |
| I promised that my policies would allow companies like Apple to bring massive amounts of money back to the United States. Great to see Apple follow through as a result of TAX CUTS. Huge win for American workers and the USA! | Positive | 0.454 | -0.48% |
| Big announcement by Ford today. Major investment to be made in three Michigan plants. Car companies coming back to U.S. JOBS! JOBS! JOBS! | Positive | 0.497 | 0.92% |

Table 4. Tweet Sentiment and Returns Example

With the required factors identified a simple regression analysis will be carried out for both the sentiment of the tweets and the effect they have on $R_{c,t}$, the return on the share price.

## 3.2 Methodology

The observations in this regression will be all tweets that are given the sentiment classification of positive or negative; neutral tweets will be removed. This is because neutral tweets about a company should have no impact on the price of its stock. Also, if a tweet is classified as neutral and given a confidence score for this classification. There is no way of identifying which side of neutral the tweet lays on, meaning it cannot be used as a continuous measure. The sentiment score ($Sen_{c,t}$), will range from -1 to 1, with the tweets classified as positive being the confidence percentage score of that classification. Similarly, for negatively classified tweets, their sentiment score will be being the negative of the confidence score of that classification.

A one-day and a five-day time period will be analysed, in order to try show the effects of the tweets differ over time. To carry out this analysis, a regression will be done where $t = 1$ and where $t = 5$. Theory would suggest that the greater $t$ is, the less the effect on the return of a tweet will be. This is due to criticisms of EMH that state traders may overreact to news.

A 1-day time period is used to capture the instant impact of the tweet, this is because a shorter time window is not available on yahoo finance. The instant effect should be the most significant as traders react to new information.

A 5-day time period is used as example of a longer view of how stock prices are impacted. This in theory should be a long enough window of time to allow the market to stabilise be at its new true value after a potential overreaction from the market. 5 days is not of significant importance and only used as an example. If a 5-day time period has a significant relationship then this will be extended to 10 days, then 15 days and 30 days.

In order to analyse the sentiment of a tweet and its effects on share prices, the regression will be:

$$R_{c,t} = \beta_0 + \beta_1 Sen_{c,t} + \varepsilon_{c,t}$$

Using this regression framework, we expect a strong positive correlation between the sentiment of the stock and the stock return over a 1-day time period. This would be to say that, the more positive the tweet is, the greater the return on the stock would be. The t-statistic of this regression will be used to test the significance of the sentiment effect.

However, we expect less or no correlation between the sentiment of the stock and the stock return when looking at a 5-day time period. This would be to say that, the price of a stock is unaffected over a 5-day time period due to the effects of tweets. The t-statistic of this regression will be used to test the significance of the sentiment effect.

The Second analysis that will be carried out will be looking at positively and negatively classified tweets and comparing their impacts on share prices against one another. This will be to investigate if there is a similarity in the way positive and negative tweet affect share prices.

To do this two data sets; positive and negative will be classified, and the confidence percentage scores used to as the sentiment variable ($Sen_{c,t}$), a value between 0 and 1. This will be slightly different to the above so that all sentiment scores are positive, being the confidence in their classification or "how negative or positive the tweet is". Also, the returns for negative tweets will be multiplied by -1, this will make the return variable a "returns in line with the sentiment classification" percentage. By doing this the positive and negative tweets can be compared on an even playing field.

In order to analyse the sentiment of a tweet and its effects on share prices, the regression will be:

$$I_{c,t,s} = \beta_0 + \beta_1 Sen_{c,t,s} + \varepsilon_{c,t,s}$$

The coefficients and standard errors will then be compared in order to see whether positive or negative tweets are more impactful on the change in the share price.

Thirdly, analysis will be carried out to see how neutrally classified tweets impact the share prices of the companies that are tweeted about. To test this, the neutrally classified tweets will be used, and their confidence scores used as the sentiment variable ($Sen_{c,t}$), a value between 0 and 1. This will be regressed against the impact ($I_{c,t}$) on the share price. This will be calculated as the absolute value of the return ($R_{c,t}$):

$$I_{c,t} = |R_{c,t}|$$

This is used as the size of the movement in share price is important in this case, rather than the direction of price movement.

In order to analyse the sentiment of a tweet and its impact on stock prices, the regression will be:

$$I_{c,t} = \beta_0 + \beta_1 Sen_{c,t} + \varepsilon_{c,t}$$

Another analysis that will be carried out will be to see the effect of the number of likes on a company specific, positively or negatively classified tweet $Like_{c,t}$ has on the stock price impact to that company $I_{c,t}$. Theory would suggest that the more likes a tweet has, the more people agree with it and the more likely it is to impact share price movements. To do this, tweets with a positive or negative classification will be used. These observations will then be used to form the following regression:

$$I_{c,t} = \beta_0 + \beta_1 Like_{c,t} + \varepsilon_{c,t}$$

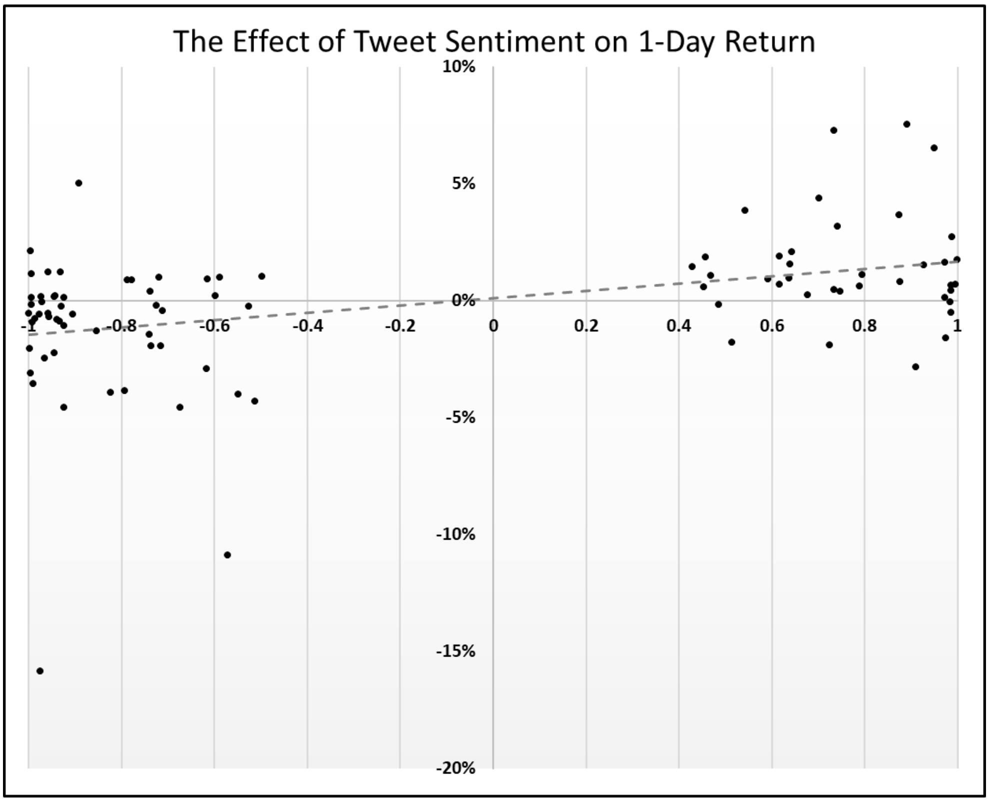
This will also be compared to the same analysis done on the number of retweets a tweet has. To see the effect of the number of retweets a positive or negative tweet $Retweet_{c,t}$ has on the share price impact $I_{c,t}$. Theory would also suggest that the more retweets a tweet has, the more people agree with it and the more likely it is to impact share price movements. To do this, the same tweets as above will be used. These observations will then be used to form the following regression:

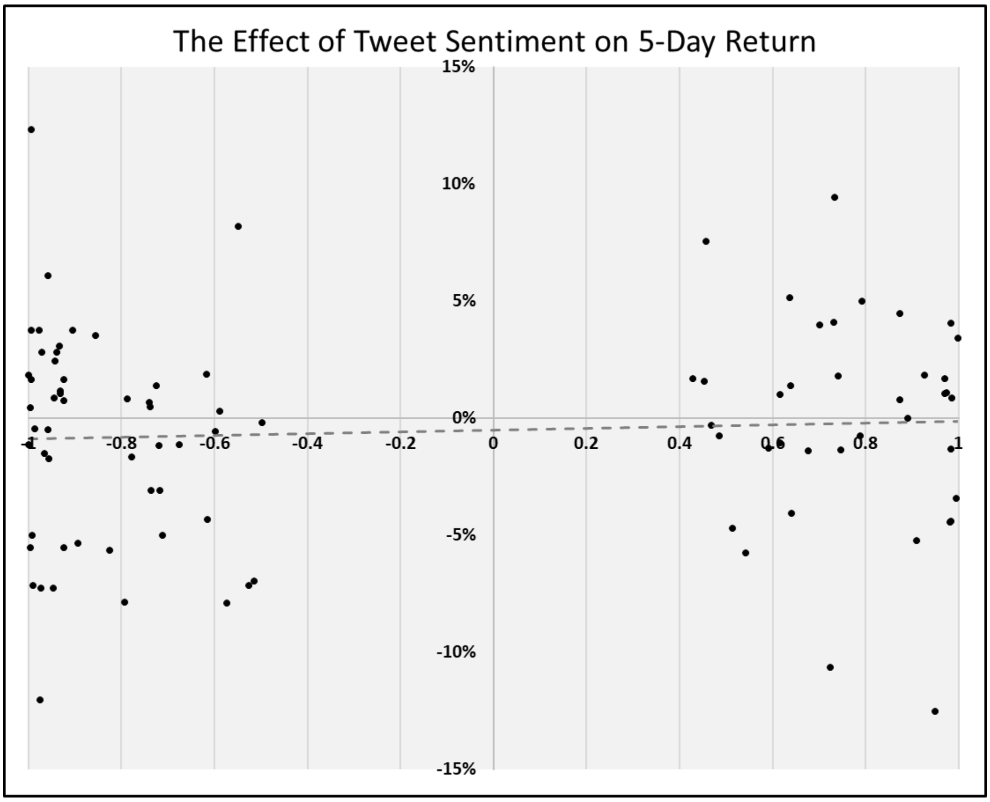
$$I_{c,t} = \beta_0 + \beta_1 Retweet_{c,t} + \varepsilon_{c,t}$$

These two regressions will also be compared, in order to see if the number of likes or retweets are more impactful.

# 4. RESULTS AND DISCUSSION

## 4.1 The Effect of Twitter Sentiment on 1-Day and 5-Day Returns



| | 1-Day | 5-Day |
|---|---|---|
| **Coefficient (Std Err.)** | 0.0156 (0.00367) | 0.00381 (0.00589) |
| **P>\|t\|** | 0.000 (3sf) | 0.519 |
| **No of observations.** | 90 | 90 |
| **R-squared** | 0.170 | 0.0047 |

Results 1: The Effect of Twitter Sentiment on 1-Day and 5-Day Returns
(see Appendix 6.2)

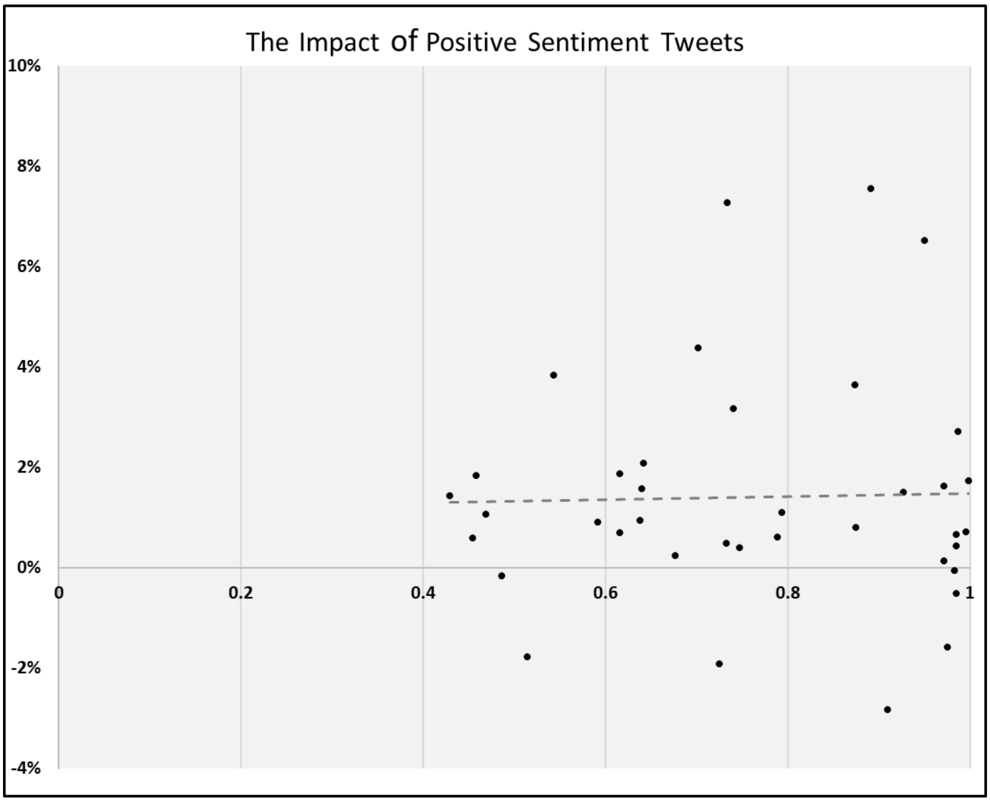
The two scatterplots above show the correlation between the sentiment of Donald Trump's tweets (X-axis) and the return of the company tweeted about (Y-axis), over a one-day and a five-day time-period respectively. The table shows a comparison between key statistics of the regression carried out of these two variables. By observation, a clear trend can be seen over a one-day time-period. It shows that generally negative sentiment tweets produce negative returns and positive sentiment tweets produce positive returns.

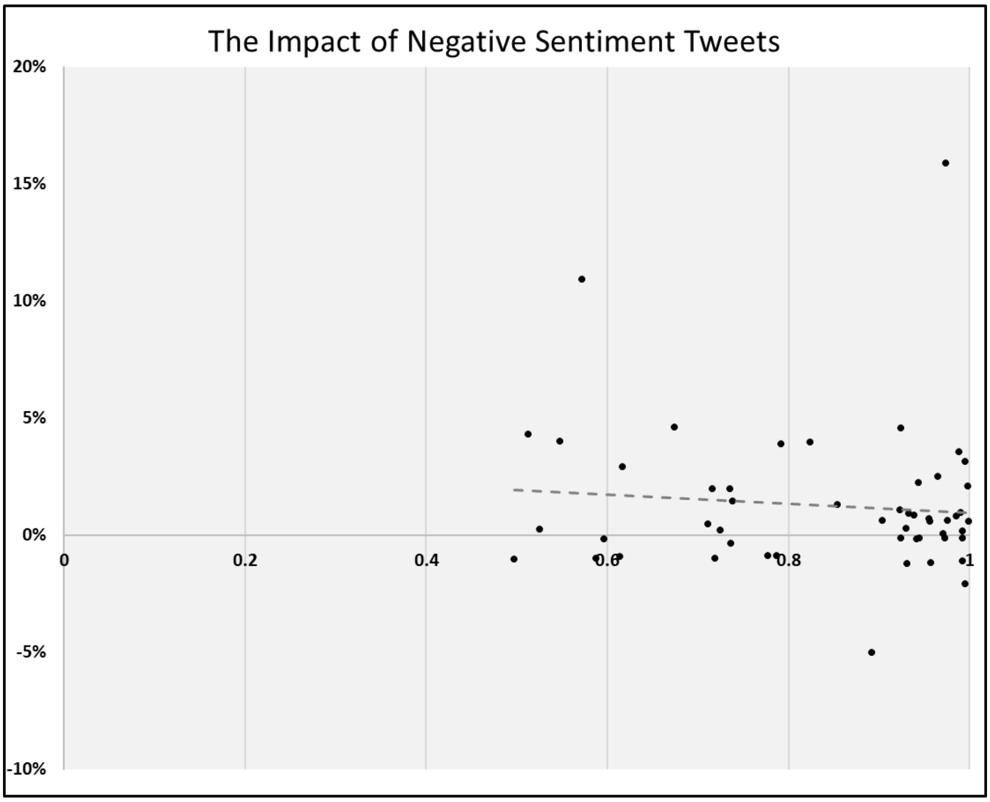
This is affirmed by the P value of the one-day return as it shows a valid relationship between the sentiment of Donald Trump's tweets and the return of the company tweeted about. The coefficient of this regression would show that on average, across this data set, a tweet with a perfectly positive or negative sentiment $Sen_{c,t} = 1$ or $Sen_{c,t} = -1$ would result in a 1.56% increase or decrease in the price of the stock of the company tweeted about. This however is only an average estimation and not entirely representative of the portfolio, the scatter plot shows a lot of observations where the returns are around 0%, it also shows extreme examples where returns are as great as -16% and +8%.

This meets the expectations set out in this report, as over a number of observations of company specific tweets there is a significant impact in the movement of the share price of those companies. This could be explained by herd mentality as individuals irrationally react to the decisions of others. It can also be explained by rational traders reacting to new information in the marketplace that is tweeted by Trump (Wilhelm, 2011).

By observation of the five-day time-period there is no clear relationship between the sentiment of Donald Trump's tweets and the return of the company tweeted about. Through the regression we can see that the P-value is far too high to be accepted at even a 10% confidence level. This is inline with expectations as the market price recovers to its true value after an initial over reaction to new information or option given by Trump over twitter. Price movements in this framework can be explained by random walk theory as well as the other numerous factors that affect the price of a company's stock. For example, other news released about the company, dividend payments, interest rates and herd mentality (Spitzer, 2013).

## 4.2 The Impact of Positive and Negative Sentiment Tweets on Stock Prices



| | Positive | Negative |
|---|---|---|
| **Coefficient (Std Err.)** | 0.00279 (0.0200) | -0.0195 (0.0276) |
| **P>\|t\|** | 0.890 | 0.483 |
| **No of observations.** | 38 | 52 |
| **R-squared** | 0.0005 | 0.0099 |

Results 2: The Impact of Positive and Negative Sentiment Tweets on Stock Prices (see Appendix 6.2)

The two scatterplots above show the correlation between the sentiment of Donald Trump's tweets (X-axis) and the Impact on the of the company tweeted about (Y-axis), over a one-day time-period for positive and negative tweets respectively. The table shows a comparison between key statistics of the regression carried out of both positive and negative tweets. The observations in this regression are the same 90 observations of the original regression, but this time split by the positively and negatively classified tweets.

By observation there is no clear relationship between the strength of sentiment and a price stock price movement in the direction of the sentiment. Both scatter plots show all by no correlation.

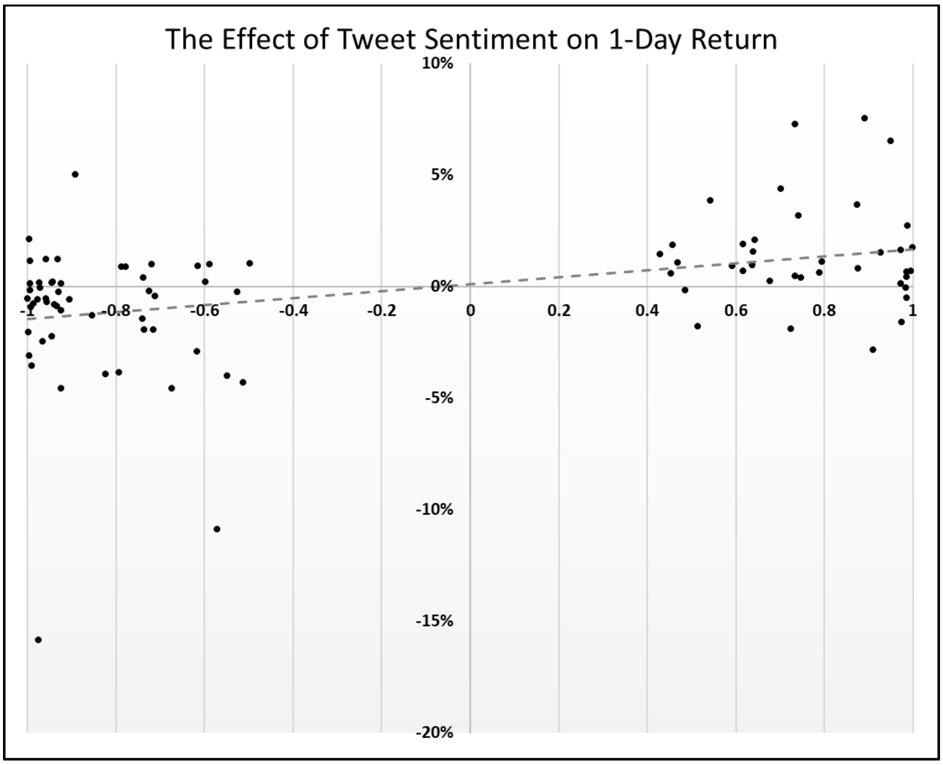
This is highlighted by the P values of both the positive and negative regression. As can be seen, both P values cannot be accepted at the 10% confidence level, indicating the lack of explanatory power. It is also shown by the negative coefficient of the negative tweet analysis. This would imply that the more negative the sentiment of Trump's company specific tweets, the less the impact on negative returns. This would certainly not be expected as there is no rational reason why the share price would decrease less as the strength of the negative sentiment increases.

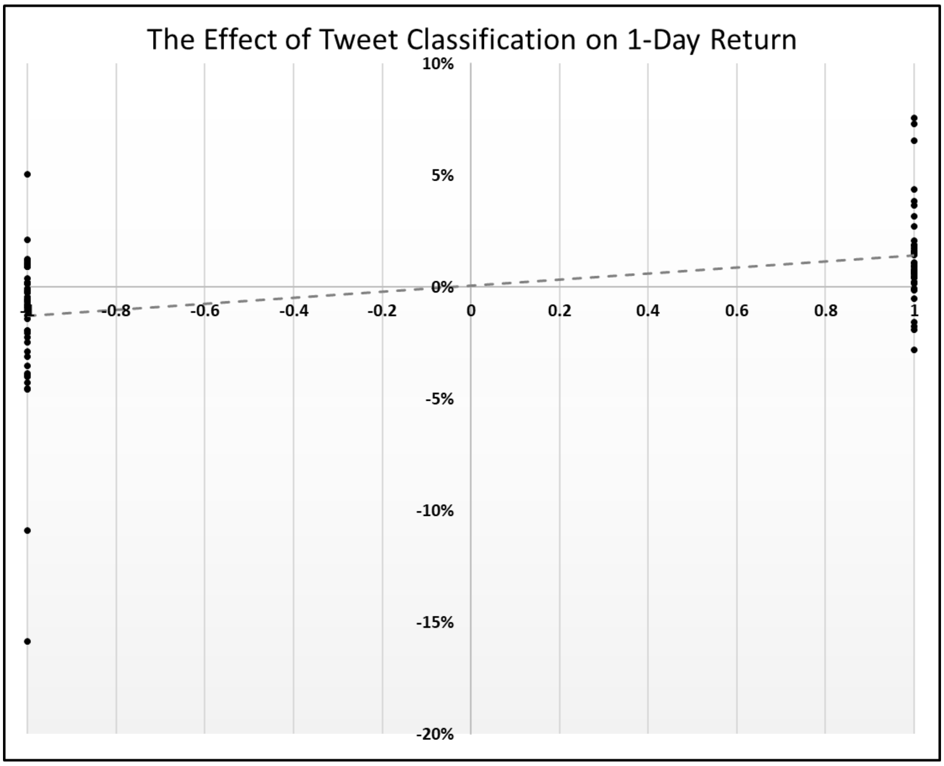
This Analysis was originally carried out to test whether the effects of positive and negative tweets impact the share price of the company are similar or different. However, after carrying out this analysis, it is apparent that when the sentiment classification is isolated; the strength of sentiment is not correlated to the impact it has on the share price of the company. This means that for example, a stronger negative sentiment doesn't directly imply a greater decrease in company share price.

Given the results of the previous regression, where it was shown that there is a relationship between the sentiment of a tweet and the change in share price this causes. The results of this test would imply that the sentiment classification can be used broadly to predict the direction of stock price movement. However, the specific scores cannot be used cannot be used to accurately predict the size of the movement. This is likely due to the number of factors that will impact the price of a company's stock as mentioned previously.

To remedy this a number of control variables could be introduced the model. This would help to further isolate the tweet specific impact on the stock price. In addition to the S&P 500 index which is already accounted for in the return's formula, this could include an industry specific index for the company being tweeted about, the number of news headlines or articles about a company in a given time period, or any other proxy to capture the overall perception of traders at the time they are tweeted about (De Long, 1988).

## 4.3 A Comparison of Tweet Sentiment and Tweet Classification



| | Sentiment Strength | Sentiment Classification |
|---|---|---|
| **Coefficient (Std Err.)** | 0.0156 (0.00367) | 0.0134 (0.00298) |
| **P>\|t\|** | 0.000 (3sf) | 0.000 (3sf) |
| **No of observations.** | 90 | 90 |
| **R-squared** | 0.170 | 0.187 |

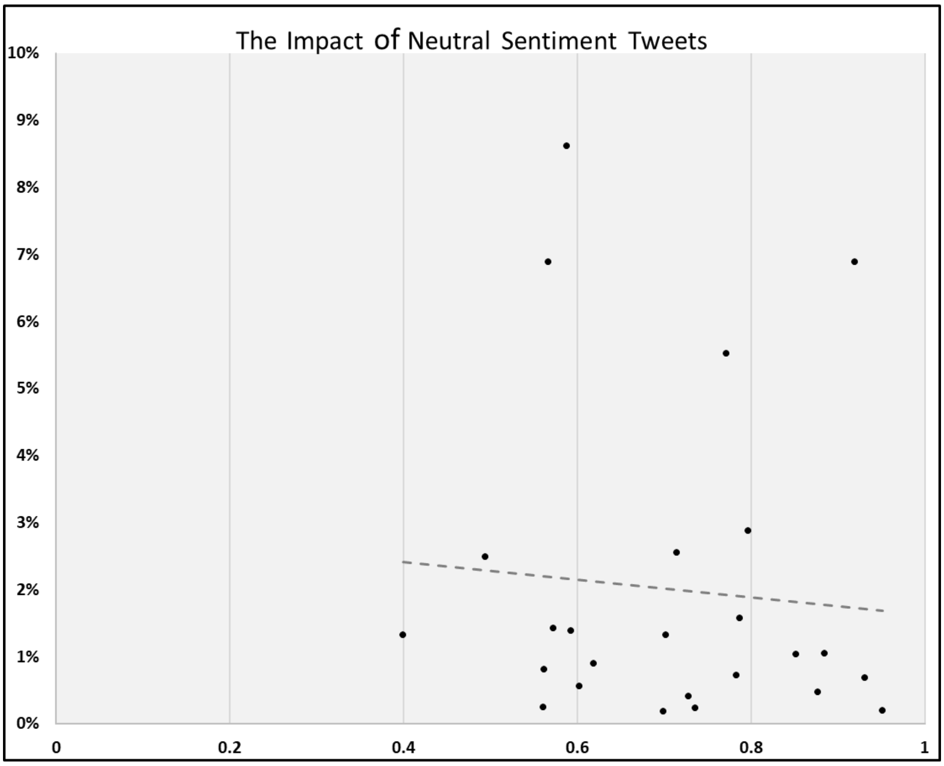
Results 3: A Comparison of Tweet Sentiment and Tweet Classification
(see Appendix 6.2)

The above regression is now carried out to compare models that use the sentiment strength to predict the size of the price movement; to a model that uses a dummy variable for positive and negative to indicate price direction. The scatter plots show the same returns, simply with all sentiment scores given as $Sen_{c,t} = 1$ or $Sen_{c,t} = -1$ in the second model.

By comparing the key statistics of the two models, it can be seen that they are very similar, both in terms of the coefficients and R-Squared values produced by the regression analysis. The R-squared values imply that the Dummy variable model is a slightly more suitable model, however this difference is so marginal that it is arbitrary to use this value to assume one model is better than the other. The coefficients are also similar showing that a perfectly positive or negative sentiment would result in an increase or decrease of 1.56% or 1.34%. Again, these coefficients do not offer much separation between the models.

Due to the similarity of the two summary statistics the best model should be decided by theory. As was previously explained in the analysis of the positive and negative tweets, it would appear that the specific sentiment strengths cannot be used in order to predict specific price movements. Due to this, despite the similarity of the summary statics, it would appear that using the sentiment classification of “positive” or “negative” would be a more suitable way to predict the way stocks are affected by Donald Trump’s tweets.

## 4.4 The Impact of Neutral Sentiment Tweets on Stock Prices



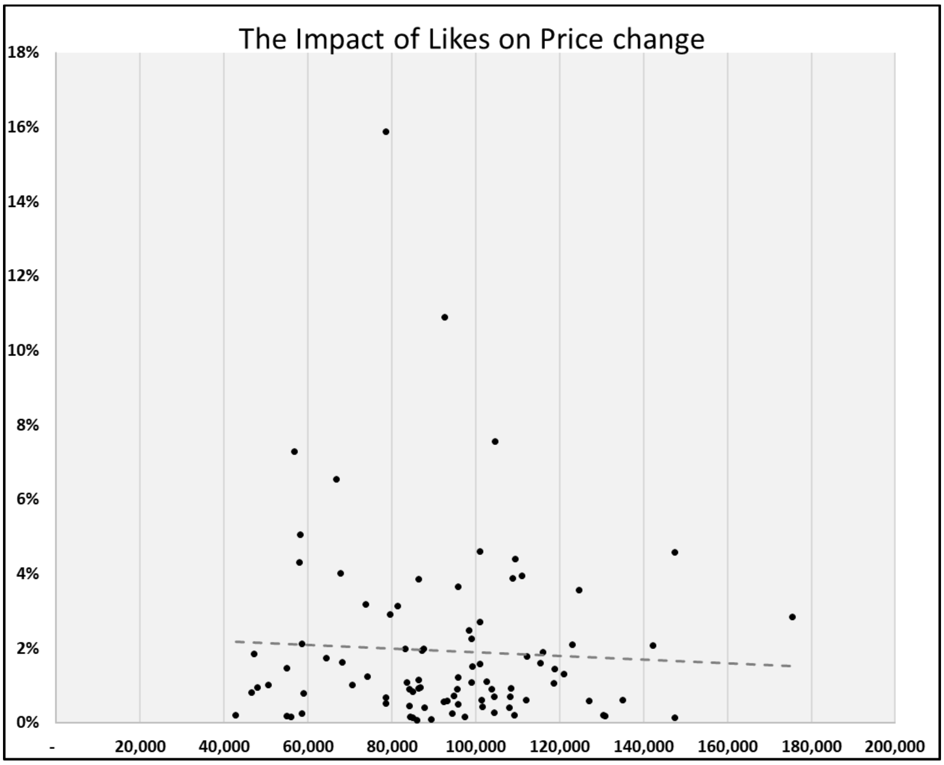

| | |
|---|---|
| **Coefficient (Std Err.)** | -0.0131<br>(0.0332) |
| **P>\|t\|** | 0.695 |
| **No of observations.** | 25 |
| **R-squared** | 0.0068 |

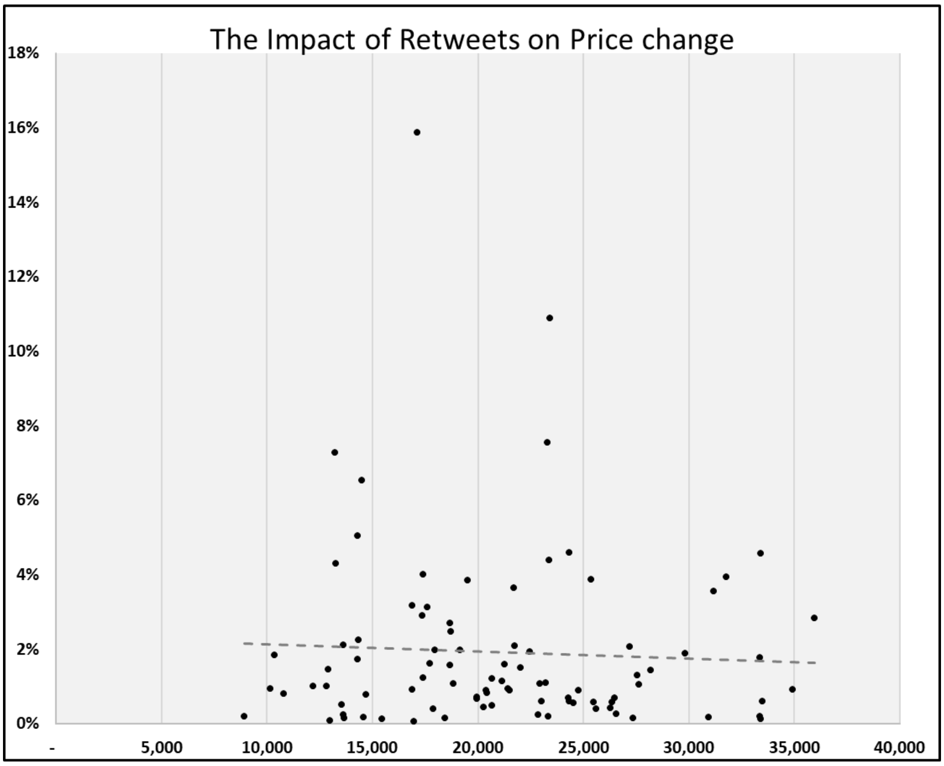
Results 4: The Impact of Neutral Sentiment Tweets on Stock Prices
(see Appendix 6.2)

The above is an analysis of Trump's Neutrally classified tweets about companies, investigating the impact this has on stock prices of the companies that are tweeted about. Both the scatterplot and the summary statistics in the table below show no correlation between the strength of the neutral sentiment and the return on stock price. Supporting this is the huge P-value of 0.934 and the very low R-squared value of 0.0003, which would indicate almost no explanatory power of how neutral a tweet is when looking at the impact it has on a stock price change.

The coefficient of this regression is however negative, in line with expectations and implying that the more neutral a tweet is the less the impact on the price of a company's stock. When looking at these results, it is hard to ignore the number of other factors that will affect the stock price and the potential reasons trump may have for tweeting neutrally about a company. It could be the case the when there is news concerning a company which would impact their stock price, Trump may tweet about it but without clearly being positive or negative. This would mean the underlying news or cause would be the reason for the price movement, rather than Trumps tweet. Due to the P-value and R-squared value being so low in the model, indicating large inefficiencies; this is a much more likely scenario.

## 4.5 The Impact of Likes and Retweets on Stock Prices



| | Likes | Retweets |
|---|---|---|
| **Coefficient (Std Err.)** | -0.000000489 (0.000000101) | -0.000000697 (0.000000508) |
| **P>\|t\|** | 0.628 | 0.637 |
| **No of observations.** | 90 | 90 |
| **R-squared** | 0.0027 | 0.0025 |

Results 5: The Impact of Likes and Retweets on Stock Prices
(see Appendix 6.2)

The above is an analysis of the number of likes and retweets a tweet has, when it is a positive or negative tweet from Trump pertaining to a company, and the impact this has on stock price. Both the scatterplots above and the summary statistics table show no evidence of a relationship between the number of likes and retweets to the impact on stock price. This observation is upheld by the respective P-values of the two regressions of 0.628 and 0.637. These are high P-values which means the cannot be accepted at even the 10% confidence interval. Also, the respective R-squared values are 0.0027 and 0.0025 which indicate highly inefficient models with no explanatory power.

By observation of both the scatterplots and the summary statistics, and also through theoretical reason there is a strong correlation between the number of like and retweets. This means if both like and retweets are introduced into the same model there would be a high level of multicollinearity between the two variables. This would mean if the variables are to be include in a similar future model then only one should be included due to potential inefficiencies. Despite this, it would not be recommended that either like are retweets be included in a future model as there is no apparent relationship in this model.

The coefficients in this analysis are both negative, however both are so close to 0 that this would be assumed to be a no relation coefficient. This is not in line with expectations as we would expect that the more likes or retweets a tweet has; the more people agree with its message and the greater the impact would be on the stock price.

There may also be no relationship as the number of likes are not an accurate proxy for the attitudes and agreement of traders. This would mean that the twitter audience of “likers” and “retweeters” may not necessarily represent traders that will actually impact the price of a company's stock. There is also the problem of the number of likes and retweets not being the number as of the day after the tweet. This would be a more accurate way of estimating price movements as future likes and retweets after the price impact is measured would not affect this price impact.

# 5. CONCLUSION

Firstly, the research in this report is an estimation of the effects of Trump's tweets on the companies tweeted about. It is therefore important to emphasise that the results of this research may offer insight in this area but are not definitive. Despite this, there certainly appears a relationship between his comments and the price movement, however it would seem that these movements cannot be calculated accurately by the sentiment, likes or retweets of the associated tweets. This means that Sentiment classification can be used as signalling mechanism of the direction of price movements, rather than individual sentiment scores being used being to reflect the exact movement of the stock price. Initially this would seem to be in line with EMH assumptions, as traders react to new available information, buying and selling stock appropriately and therefore affecting the price of the stock (Fama, 1998).

However, there appears to be a difference in the ways in which traders react within a 1-day and a 5-day time period. This would suggest an initial overreaction by traders and some elements of a herd mentality before a readjustment of views days after the tweet (Hoffmann, 2015). EMH assumptions would this should not happen, and that traders will always react instantly to the correct magnitude of the given new information. This then gives scope for a buy-sell strategy within the stock exchange as the actions of traders and movements of become more predictable and consistent within this framework.

Estimating stock price movements is an extremely complicated and imperfect science due to the number of factors and facets that dictate prices, which come from bother rational reasoning. This highlights the importance and significance of the findings of this report, as its research has found an innate behaviour of traders which is predictable in the short term. Where behaviour is predictable it can be exploited by traders to make a profit (Chiang, 1988).

## 5.1 Calls for Future Research

The main purpose of these findings should be to inform and influence a buy-sell strategy. This would be to approximate and capitalise on the aggregate actions of traders following a company specific tweet with a strongly positive or negative sentiment. In order to do so, the findings of this research would have to be developed and made more sophisticated so that more accurate pricing predictions can be calculated. With this more sophisticated model, tweets could be analysed in at the time of tweeting and be used to calculate buy-sell conditions.

A development of the sentiment tools used could be necessary in a more sophisticated model. This is because if a tweet has a general positive sentiment, for example, phrases like “the great American people”; however, a company is slandered within this tweet. This could result in a positive sentiment classification, despite the company specific sentiment being negative and a likely negative impact of the stock price of the company. This could be called a misclassification as the sentiment analysis tool does not look at the sentiment explicitly pertaining to the companies. A more sophisticated sentiment analysis tool may be able to separate the general sentiment of a text from the company specific sentiment. This would lead to more accurate findings in the analysis of this report.

Another development of the text analysis methods used in this model would be subjectivity analysis. EMH assumptions would suggest that all traders are rational and therefore able to separate the subjective from the objective. Traders should understand that a more subjective tweet is based on opinion rather than facts and substantial information. This would mean that the more subjective a tweet is the less the impact should be on the stock price movement. With this considered subjectivity analysis could also be carried out and used in order to make estimations of future stock prices more accurate.

Fleets have also been recently introduced to twitter, this is a tweet which disappears after 24 hours, this could also be introduced into the model as a dummy variable in order to assess the effect this condition has on its effect on traders. This is an example of how such a model would have to adapt to changes in the platform in order to remain effective.

Finally, many more variables would need to introduce into the model in order to try account for the many other factors that can influence the price of a stock. In a perfect model, variables would account for news headlines and other new information pertaining to the company would be considered. Also, a proxy for the confidence of traders towards both the company and industry would be included. There are other additional factors which would influence stock price such as the timing of dividend payment and macro variables like interest rates and inflation.

The scope of further investigation in this area is huge, several other individuals could be analysed to develop the findings of this report. In a world of increasing influencers this could be useful to understand if these effects are consistent amongst a host of influential individuals. This would also broaden the range of tweets which could be analysed and lead to a more frequently active buy-sell tool. In future research I would recommend a similar investigation of the other individuals in *Table 1*. Who are both on the Time100: Most Influential People in the World shortlist and the one of the 100 most followed individuals on twitter. Additionally, any other well followed individual who posts information and opinion pertaining to companies and financial markets.

This analysis could also then be carried out on news stories, financial reports, interviews and any other new information available to traders with text that can be analysed and that may have influence on the decisions of traders.

Overall this report and research accomplished its aims and objectives in identifying a relationship between the sentiment of Donald Trump's tweets and the impact this has on the stock price of the company that is tweeted about. This is evidence to the contrary of Fama's Efficient market hypothesis. However, I feel these findings are just the start of a far more sophisticated analysis by which a trading strategy can be built, trading in real time as tweets are other information are published to the public.

# 6. APPENDIX

## 6.1 Formulas

**Formula 1:**
=SUMPRODUCT(--ISNUMBER(SEARCH(E1,$D3)))

Where;

E1= Full Company Name

$D3= the tweet

**Formula 2:**
=@INDIRECT(ADDRESS(1,MATCH(1,A1:S1,0)))

Where;

A1:S1= The range of outputs from *Formula 1*

**Formula 3:**
=IF(SUMPRODUCT(--ISNUMBER(SEARCH(E$2,$D3)))=0,SUMPRODUCT(--ISNUMBER(SEARCH(E$1,$D3))),SUMPRODUCT(--ISNUMBER(SEARCH(E$2,$D3))))

Where;

E1= Full Company Name

E2= Shorthand Name

$D3= the tweet

## 6.2 Tables

*Disclaimer* Unfortunately the Stata output was unable to be inserted because of the closure of university facilities due to the Coronavirus

All tables were already created as a result of these Stata outputs, highlighting the most useful and necessary information.

# 7. References


Alrabadi, D. W., 2010. Systematic Liquidity Risk and Stock Price Reaction to Large One-Day Price Changes: Evidence from London Stock Exchange.. *University of Bradford,* pp. 1-34.

Bloomberg, 2019. *JPMorgan Creates 'Volfefe' Index to Track Trump Tweet Impact.* [Online] Available at: https://www.bloomberg.com/news/articles/2019-09-09/jpmorgan-creates-volfefe-index-to-track-trump-tweet-impact

Bollen, J. & Mao, H., 2011. Twitter mood as a stock market predictor. *Computer,* pp. 91-94.

Bollen, J., Mao, H. & Zeng, X., 2011. Twitter mood predicts the stock market. *Journal of computational science,* pp. 1-8.

Bowley, G., 2010. *Computers That Trade on the News,* s.l.: New York Times.

Cheng, P. & Deets, M., 1971. Portfolio returns and the random walk theory. *The Journal of Finance,* pp. 11-30.

Chen, R. & Lazer, M., 2013. Sentiment analysis of twitter feeds for the prediction of stock market movement. *Stanford edu Retrieved.*

Chiang, R. &. V. P. C., 1988. Insider holdings and perceptions of information asymmetry: A note. *The journal of finance,* 43(3), pp. 1041-1048.

Citigroup, 2019. *JPMorgan Creates 'Volfefe' Index to Track Trump Tweet Impact.* [Online] Available at: https://www.bloomberg.com/news/articles/2019-09-09/jpmorgan-creates-volfefe-index-to-track-trump-tweet-impact

Dang, H. & Lin, M., 2016. Herd mentality in the stock market: On the role of idiosyncratic participants with heterogeneous information. *International Review of Financial Analysis,* pp. 247-260.

Datahub, 2019. S&P 500 Companies with Financial Information. [Online]
Available at: https://datahub.io/core/s-and-p-500-companies#data

De Long, J. B. S. A. S. L. H. &. W. R. J., 1988. The survival of noise traders in financial markets. *The Journal of Business,* Volume 64.

DeMarzo, P. M., Vayanos, D. & Zwiebel, J., 2003. Persuasion Bias, Social Influence, and Unidimensional Opinions. *The Quarterly Journal of Economics,* pp. 909-968.

Fama, E. F., 1970. Efficient capital markets: A review of theory and empirical work. *The journal of Finance,* 25(2), pp. 383-417.

Fama, E. F., 1995. Random walks in stock market prices. *Financial analysts journal,* pp. 75-80.

Fama, E. F., 1998. Market efficiency, long-term returns, and behavioral finance. *Journal of financial economics,* 49(3), pp. 283-306.

FriendorFollow, 2019. *Twitter: Most Followers.* [Online]
Available at: https://friendorfollow.com/twitter/most-followers/
[Accessed 21 November 2019].

Hoffmann, A. P. T. a. P. J., 2015. How investor perceptions drive actual trading and risk-taking behavior. *Journal of Behavioral Finance,* 16(1), pp. 94-103.

Horne, J. V. & Parker, G., 1967. The random-walk theory: an empirical test. *Financial Analysts Journal,* pp. 87-92.

Malkiel, B. G., 1989. Efficient market hypothesis. *Finance. Palgrave Macmillan,* pp. 127-134.

Malkiel, B. G., 2003. The efficient market hypothesis and its critics. *Journal of economic perspectives,* pp. 59-82.

MonkeyLearn, 2019. Sentiment Analysis
Available at: https://app.monkeylearn.com/main/classifiers/cl_pi3C7JiL/

Naseer, M. & Tariq, Y. b., 2015. The efficient market hypothesis: A critical review of the literature. *IUP Journal of Financial Risk Management,* pp. 48-63.

Rayarel, K., 2018. *The Impact of Donald Trump's Tweets on Financial Markets.* [Online]
Available at: https://www.nottingham.ac.uk/economics/documents/research-first/krishan-rayarel.pdf

Rechenthin, M., Street, W. & Srinivasan, P., 2013. Stock chatter: Using stock sentiment to predict price direction. *Algorithmic Finance,* pp. 169-196.

Shantha, K., 2018. Shifts in Herd Mentality of Investors in Uncertain Market Conditions: New Evidence in the Context of a Frontier Stock Market. *Journal of Economics and Behavioral Studies,* pp. 203-219.

Spitzer, F., 2013. Principles of random walk. *Springer Science & Business Media,* Volume 34.

Statista, 2019. *Number of Monthly Active Twitter Users.* [Online]
Available at: https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/

TIME, 2019. *100- Most Influential People.* [Online]
Available at: https://time.com/collection/100-most-influential-people-2019/5567754/donald-trump-3/?fbclid=IwAR14vwViRoz6VL1aOH6aRvF-fB8XedFLUvH-H8qWNo2pbQBYRJbwrPYYdko

TrumpTwitterArchive, 2019 [Online]
Available at: http://www.trumptwitterarchive.com

Twitter, 2019. *@RealDonaldTrump.* [Online]
Available at: https://twitter.com/realdonaldtrump

Wilhelm, E., 2011. *Agent-based Modeling of Herd Mentality in the Stock Market.* s.l.:Department of Computer Science, Faculty of Engineering, LTH.

Yen, G. & Lee, C.-f., 2008. Efficient Market Hypothesis (EMH): Past, Present and Future. *Review of Pacific Basin Financial Markets and Policies,* pp. 305-329.