Available online at www.sciencedirect.com

ScienceDirect

The Journal of Finance and Data Science 4 (2018) 101−119
http://www.keaipublishing.com/en/journals/jfds/

Twitter as a tool for forecasting stock market movements: A short-window event study

Tahir M. Nisar*, Man Yeung

Southampton Business School, University of Southampton, Highfield, Southampton SO17 1BJ UK

Received 19 May 2017; revised 6 October 2017; accepted 6 November 2017
Available online 9 February 2018

Abstract

In order to explore the relationship between politics-related sentiment and FTSE 100 movements, we conducted a short-window event study of a UK based political event. We collected a sample of over 60,000 tweets using 3 key hashtags during the period of 6 days including before, during and after the 2016 local elections. The study involved performing a collection of correlation and regression analyses to compare daily mood with daily changes in the price of the FTSE 100 at the market level. The findings suggest that there is evidence of correlation between the general mood of the public and investment behavior in the short term; however, the relationship is not yet determined as statistically significant. There is also evidence of causation between public sentiment and the stock market movements, in terms of the relationship between MOOD and the daily closing price, and the time lag findings of MOOD and PRICE. Overall, these results show promise for using sentiment analytics on Twitter data for forecasting market movements.

JEL classification: G14; G17; G41

Keywords: Twitter; Political discussions; Stock market prediction; Sentiment analysis; FTSE 100

1. Introduction

An individual can broadcast a brief statement in real time to some or all members of the sender's social network through Twitter. Twitter, which has large audience potential, currently attracts an estimated average of 271 million users every month.[32] Recent research has explored whether the 'Twitter effect' is economically substantial.[12,16] Twitter effect has been shown to be particularly relevant to experiential media products (e.g., movies, music, and electronic games); these are generally the products for which 'instant' success is essential. For instance, the failure of Brüno, a multimillion-dollar movie, has been attributed to the negative sentiments about the movie expressed on Twitter. Positive sentiments, on the other hand, have been the perceived cause of the unexpected opening success of the remake of Karate Kid.[16] Twitter-based models can then be built to aggregate the opinions of the collective

* Corresponding author.
E-mail addresses: t.m.nisar@soton.ac.uk (T.M. Nisar), manyeung17@gmail.com (M. Yeung).
Peer review under responsibility of China Science Publishing & Media Ltd.

https://doi.org/10.1016/j.jfds.2017.11.002

population. They can be used to predict future trends while gaining useful insights into individual behavior.[2] Social networks offering microblogging services enable the rapid spread of user generated content (UGC) from a handful of individuals to millions of people around the world in the form of short text, images or videos. Microblogging platforms have grown so exponentially, that they are now perceived as indispensable sources of information[25] and are fast gaining popularity amongst users, organizations and researchers in various disciplines.

The popularity of microblogging can be explained by its distinctive features such as convenience and accessibility, which allows users to instantly respond and disseminate information with limited or no restrictions on content.[3] It can be argued that social media has significantly impacted our daily lives and has changed the way that individuals and businesses perform, create awareness and seek advice.[3] It is therefore not surprising that the rich and continuous mass of data made available by these platforms is being harnessed with the purpose of studying individual and group behavior as well as global patterns especially in regards to sentiment towards brands, products, events, recent news and social and political issues. Twitter is currently the 10th most popular website globally with over 300 million active monthly users.[47] Twitter is updated hundreds of millions of times a day with content varying from individual daily life updates to worldwide news and events. Twitter allows users to create personal profiles that others may subscribe to or 'follow', publish status updates known as 'tweets' limited to 140 characters and to communicate with others through 'replies'. 'Retweeting' is a common practice whereby a user can choose to forward a tweet they find interesting to their followers whilst crediting the original author, allowing popular posts to travel well beyond the network of the original creator. It is therefore perceived that highly retweeted posts reflect the views of the global Twitter community.[50] Twitter also encourages the use of hashtags which allow tweets to be collated in a thread that can be used for following specific events and topics. Twitter tracks the most mentioned phrases and hashtags and posts them under a list of 'trending topics', which is updated regularly, and allows users to keep track of what is most popular at any given time.

Originally, Twitter was set up as a type of communication platform designed to allow friends to keep tabs on one another. Thanks to the availability of an application programming interface (API), which stores tweets that may be accessed by researchers, and its convenient features such as filtering by variables like location and keywords, Twitter has encouraged researchers to take an interest and explore its potential beyond that of a social network.[3] Since the conception of Twitter in 2006 studies into properties of Twitter have grown in popularity and can be classified in one of the following streams; structural, content or sentiment. When evaluating the structural properties of Twitter as a social network studies have focused on user influentially.[5,10] Content analysis studies have focused on analyzing the content, virality and motivations of tweets. Java et al[24] categorized the motivations of tweets into the following categories; daily chatter, conversations, sharing information and reporting news. Sentiment analysis studies have focused on using Twitter chatter sentiment for predicting behavior. Bollen et al[8] arguing that although each tweet represents individual opinion, an aggregate sample should provide an accurate representation of public mood.

Motivated by the prospect of a forecasting relationship between politics and market performance,[28,37] the EU Referendum debates inspired an investigation into the short term impact of user-generated content defined as the collection of UK based tweets recording user sentiment towards politics and the government during the 2016 local elections, on financial market movements specifically looking at the FTSE 100 at the stock market level. The research questions we are aiming to answer are:

1. Does correlation exist between public sentiment regarding the local elections, and stock market movements in the FTSE?
2. Does the relationship of causation exist between public sentiment regarding the local elections and stock market movements in the FTSE?

In order to explore the potential relationship between politics-related sentiment and FTSE 100 movements a UK based political event has been identified and will be the main focus of the data mining in this investigation. On the 5th of May 2016 UK saw the biggest set of elections in years outside of the general election.[6] Voters were deciding on a variety of issues including but not limited to; local councils, the police commissioner and the Mayor of London. During the days before and after the election highly charged debates took place over Twitter with users weighing in and sharing their sentiment towards candidates and political issues of concern. The debates on Twitter reached a high number of posts with many politics-related hashtags achieving the 'Trending Topic' status, including: #toryelectionfraud, #mayoralelections, #VoteConservative, #iVoted, #PollingDay and #LondonElects.[44]

The rest of the paper is organized as follows: Section 2 will look at existing literature in order to illustrate the motivation behind the study. Section 3 will describe the method used and Section 4 will illustrate the empirical results. Section 5 will form a discussion of the main findings in context of the existing literature and original aims and objectives. Finally, Section 6 will conclude with a summary of the findings, key limitations and future research implications.

2. Literature review

Twitter is characterized by the real-time transmission of product quality information and reviews, and thus it enables feedback. The receiver of such information can potentially be a very large group, and not just an individual or a small group.[32] Furthermore, the brevity of Twitter content is a unique element that is not typical for other types of word-of-mouth (WOM) communication, but nevertheless contributes to concise evaluations that are perceived as unequivocal. Microblogging is also recognized to hold huge potential for the successful implementation of many other related organization and management practices.[32] Microblogs are a form of social communication whereby users can express their interests and attitudes in short posts, which are instantly distributed to other users via mobile phones, instant message, and the web.[12] Chen and Xie[12] argue that it is microblogging's instant, real-time access to consumers that makes it ideal for managing wider market relationships. At its most basic, Twitter is a communication tool that allows users to post 140-character messages (tweets) to all those who have opted to follow them.[16] However, regardless of its simple exterior, with its numerous features this straightforward platform has proved itself to be incredibly valuable to businesses.[12,16] In order to determine whether there is a relationship between public mood and stock market performance in the context of this study, prior literature will be used to determine the following:

1. The existence of political debates on Twitter and their implications
2. The assumptions about the existence of a relationship between political events and stock market performance
3. The ability to use Twitter data to predict stock market movements

2.1. Twitter as a communication platform

Several studies have identified Twitter as a social media platform used primarily for communication and spreading information. Whilst studying user motivations Java et al[24] observed that users participate in communities which share similar interests, some users taking on the role of information providers and others information seekers. In contrast, Miller[34] argued that communication on Twitter is mainly of a 'phatic' nature and users are seeking to build connections rather than spread information. Nonetheless, a more recent study by Smith et al[42] supports the findings of Java et al[24] concluding that users of Twitter, compared to Facebook and YouTube, are the least likely to use the platform for self-promotion, instead most use it to engage in discussions and disseminate news. Cha et al[10] found that news stories consistently receive a high level of retweets over a range of topics, however not all events or news obtain the same amount of attention or level of virality.[7] One explanation for the variance in popularity is a difference in emotional arousal[21,23,25] or the content of the post.[39] Authors found that out of all sampled categories political hashtags and therefore discussions were the most persistent on Twitter but, this does not necessarily signify that the increase in the number of tweets is due to new users joining the discussion.[29]

2.2. Twitter as a communication platform- political discussions

A growing research trend is looking at the presence of political discussions and propaganda within Twitter by politicians, campaigners and the public. The focus of these studies have been in a variety of countries and over a plethora of political issues but the relative newness of the discipline and use of mixed methodology has meant that, patterns of behavior are yet to be established.[26] Hanna et al[22] show that the volume of Twitter messages discussing politics rises in reaction to a political event. In an extension to this Shamma et al[40] find that the number of tweets varied depending on the stage of the event with the bulk of messages posted in the hours directly after an event. Other studies have focused on evaluating whether Twitter is used for political propaganda or to stir up political debates. Tumasjan et al[46] investigated Twitter chatter in regards to the German Federal Elections in 2009 and found that Twitter is used as a platform for political discussions based on a large number (over a third) of their sample tweets consisting of replies. In contradiction, Small[41] found that only 7% of the tweets collected, represented conversation in the Canadian general election discussions. At first glance this contradicts the findings of Tumasjan et al[46]; however, the sample size used in Small[41] was considerably small (a hundredth), which could imply that the sample of data collected was not a fair representation of the political chatter on Twitter.

Finally, some studies have focused on further implications of political debates on Twitter. Tumsjan et al[46] find that simply the number of tweets collected per party is a good reflection of the number of votes they receive in the election and concluded that Twitter can be considered a good indication of real-time political sentiment. Another application of political debates on Twitter is the organization of protests as observed in the aftermath of the 2009 Iranian elections.

2.3. Relationship between political news and events, and market performance

Although it is generally accepted that stock market prices are largely driven by new information and follow a random pattern, many studies have tried to predict stock market behavior using external stimuli on the basis of behavioral economics that emphasizes the important role of emotion in decision making.[4,31,35] In general, Zach[49] suggests that there is limited evidence of a direct link between political events and market performance due to difficulty in quantifying political changes. In an extension to this Klibanoff et al[28] conjecture that the problem with this thread of research is that usually studies are retrospective and, therefore, determinants of investor behavior such as investment sentiment are no longer observable.

Niederhoffer[36] observes that world news had a 'discernible influence' on the movement of the market basing their conclusion on larger returns (S&P 500) following world events, as opposed to normal days. In contradiction to this Cutler et al[17] found no significant relationship between market movements (S&P 500 largest price changes) and the release of political news. Chan and Wei[11] tested the impact of political news on Hong Kong stocks and found that favorable political news cause positive returns and vice versa, however this was only true for blue chip stocks and red chip stocks were considered safe from political influence. Klibanoff et al[28] show that absolute values of returns and net assets are significantly higher in weeks containing major news suggesting that investors change their behavior as a reaction to dramatic news. This could be explained by the results of De Bondt an Thaler[18] who found that most people overreact (negatively) to unexpected and dramatic news events, so prior 'losers' are found to outperform prior 'winners' in the future. More recent studies seem to confirm the existence of a relationship between political news and events and stock market movements. Zach[49] discovers that stock market returns (Israeli) are more extreme following political events, while Mei and Guo[33] find that there is a significant relationship between political uncertainty and financial crises after controlling their investigation for factors such as market contagion and differences in economic conditions of the sampled nations. Finally, Parker[37] finds that the relationship between public mood and trust in the government is more significant for stock market investors than non-participants, suggesting that public mood in relation to politics is a factor of investment decision making.

2.4. Mining twitter to predict market movements

Volume and sentiment analysis of UGC is an increasingly popular area of research with the potential of many applications. UGC has been used to predict outcomes and successes of books,[15] movies,[19] music albums (Dhar and Chang, 2009) and even political elections.[46] A more vibrant area of research has been the use of UGC to predict financial market movements. Studies looking at discussion forums discovered correlations between high activity levels and abnormal market returns[1,45] but found no predictive value. Similar studies by Mao et al, Kaminski and Gloor, Rao and Srivastava, Mittal and Goel, Zhang et al, Zheludev et al and Bollen et al[8,27,31,35,38,49,50,52] have found Twitter chatter more successful in establishing a predictive relationship in terms of financial market performance. These studies can be broken down into two main categories: volume or sentiment driven. For instance, Mao et al[31] using volume analysis show that the daily number of Tweets mentioning S&P 500 are significantly correlated with S&P 500 stock indicators, and can be used to predict movements at market and sector level with 68% accuracy. In contrast, Rao and Srivastava[38] find no significant causal relationship between volume and stock market predictions and Zheludev et al[52] through a comparison of methods suggest that sentiment is a consistently stronger forecaster of stock market movements than volume.

Studies such as Rao and Srivastava[38] focus on general levels of emotions (negative or positive) and uncovered that there is a predictive relationship between twitter mood and stock movements but only up to a certain point. Kaminski and Gloor[27] employed a similar method and found that positive correlation exists between volume of positive tweets and close prices, but in a time lag of 48 h, negative tweets are better indicators. Bollen et al[8] employed a more sophisticated sentiment collection model and as a result were able to predict daily changes in DIJA closing price with 87.6% accuracy. They observed that specific emotions of tweets (calm) were more correlated with stock movements than generalized ‘positive’ tweets, confirming a superiority of this method over volume and other types of sentiment classification methods of prediction.

3. Hypotheses

Although there has been significant progress made in determining the relationship between volume and sentiment of UGC and market performance, arguably the possibility of a relationship between content of postings and market performance has been neglected.[20] Focusing on political Tweets specifically, allows this study to explore the relationship between content and market performance using the mixed methodology method as employed by Zheludev et al[52] Based on the relationships identified by the literature and the expectation of seeing similar results in this study the hypotheses for the research questions being asked are established below.

3.1. Hypothesis for Q1

Hypothesis 1: There is significant correlation between Twitter sentiment and volume statistics and stock market indicators for a given day.
Null Hypothesis 1: There is no significant correlation between Twitter sentiment and volume statistics and stock market indicators for a given day.

3.2. Hypothesis for Q2

Hypothesis 2: There is a statistically significant predictive relationship between Twitter sentiment statistics and stock market indicators.
Null Hypothesis 2: There is no statistically significant predictive relationship between Twitter sentiment statistics and stock market indicators.

4. Methodology

In this section the data collection methods will be outlined including the sample selection, method of computing sentiment and statistical analysis methods.

4.1. Twitter data collection

4.1.1. Choice of sample days

Using a desktop program called Tweetcatcher[9] with access to the Twitter API, a sum of 60,944 English language tweets have been collected in the sample window of May 4th 2016−May 9th 2016. May 4th was chosen as a sample day to measure the election mood prior to the Election Day. May 5th was chosen to measure the mood during the day of the election itself. The post-election period was extended to May 6th−9th in order to measure mood whilst election results were revealed and in case of a lag in user reactions. The rationale for selecting a small sample window was the short-window event study nature of the investigation revolving around the 2016 local elections specifically.

4.1.2. Choice of sample, extraction and the corresponding limitations

Although the exact method of how Twitter mine the trending topics is not known, there is a general assumption that the list is a good representation, if not complete, of the most popular issues.[29] The method used for collecting relevant tweets is adapted from Cheong and Lee[13] where they analyzed the ‘trending topics’ list to identify topics related to breaking news stories of terror events in order to narrow down their data collection. Trendogate[44] independently of Twitter records the trending topics by country for any given day and allows location, date and hashtag queries. In order to find the main hashtags discussing the local elections in the UK, Trendogate[44] was perused using UK and the specified date as parameters. As Twitter does not group related trending topics, in the time period of 4th−9th May, election-related hashtags trending in the UK included; #toryelectionfraud, #mayoralelections, #VoteConservative, #iVoted, #PollingDay, #Election2016, #LP2016, #SP2016 and #LondonElects. For the purpose of this study only the following hashtags were selected and mined; #PollingDay, #mayoralelections and #LondonElects. The rationale for selecting certain hashtags relates back to the original aim of measuring the general mood of the public in relation to the election. By selecting general hashtags the aim is to avoid bias in sentiment towards specific parties that may exist due to over representation and stronger social media presence and propaganda. #Election2016 was general enough to be sampled, however, as there are several elections taking place in 2016 in a variety of countries it would be unclear in the tweets which specific election was being discussed and so it was excluded.

To collect the relevant data for the sample window of the study, keyword queries were carried out using Tweetcatcher[9] for the selected hashtags using May 4th−9th as the time parameters. One limitation of sampling our data using hashtags is that tweets related to the election without a hashtag would be ignored. Another limitation stemming from the method of only looking at trending hashtags is that other relevant hashtags may exist but as they have not reached enough mentions to make the top list of the day they will be ignored and therefore potentially so will relevant data. One limitation of collecting data using independent hashtag queries is that if a tweet contains more than one hashtag as sampled by this study then that tweet will be recorded twice and therefore will lead to an overrepresentation of that tweet and its sentiment in the study. For instance, if a negative tweet is tagged with more than one sampled hashtags and if this tweet is retweeted or there are several negative double or tripled tagged tweets, this may lead to an overrepresentation and bias towards the negative sentiment in the data with the multiplier effect unknown.

4.1.3. Data cleaning method and the corresponding limitations

As the aim of this paper is to investigate UK Twitter buzz only, sampled tweets needed to be filtered by location. For each tweet extracted by the queries some of the key information provided included; date and time of submission, tweet content including whether it is a retweet and location. The availability of these variables allowed the data to be split by day and filtered by location. Although the query results provided the location variable, the variable provided was the location as displayed in the user profile, which can be manually entered and therefore can present itself as anything. This meant the filtering process was met with several issues:

1. Location variable was left blank.
2. Location variable had an unknown location or a random word e.g. wouldn't you like to know, somewhere over the rainbow and here
3. Location variable was not specific enough e.g. earth, global and EU
4. Location variable had multiple locations e.g. sometimes London sometimes Paris

Tweets with issue 1−3 were excluded from the final sample of tweets as a location could not be determined or narrowed down to the UK. Tweets with issue 4 remained in the sample as the assumption was made that users are likely to be located in the UK at the time. One way to improve filtering by location could be to adopt the method used by[50,51] mapping the longitude and latitude of the tweets to narrow down to a specific location. This is a more sophisticated and certain way to ensure only UK based data would be collected, however, this option is limited to only those users that have agreed to share their location with Twitter. In this study both retweets and tweets were included in the sample. The rationale behind including retweets in the sample is that the more a post is retweeted the more it is accepted and widely recognized[50,51] so retweets are more likely to represent the common views of the public than actual tweets.

Other considerations to have in mind during data cleaning should be the existence of spam, unrelated noise and fake accounts within Twitter[13,14,29] that can pollute the relevance of the data and add bias. A study by Cheong and Lee[14] found that there are certain characteristics that are likely to be exhibited by spam accounts such as newness of account, low level of profile customizations and exclusion of certain biographic information. Kwak et al[29] characterized spam tweets as those containing 3 or more trending hashtags and tweets containing shortened URLs recommending their removal. It should be noted that the URL spam characteristic is more appropriate to data sets that contain general tweets rather than those related to a particular event. An observation of a sample of URL containing tweets in this data set revealed that most URLs referenced relevant news articles in relation to the election and therefore would not necessarily be classed as spam.

4.1.4. Sentiment analysis and the corresponding limitations

In order to extract the sentiment of every tweet a lexicon based sentiment classifier, Umigon,[30] was used. The app uses a 4 step process to identify sentiment (positive, negative or neutral) with special attention paid to the use of smileys and the use of hashtags. In this sample, 21.44% of tweets are positive, 14.30% are negative and 64.27% are neutral. The similarity of negative and positive percentage suggests that the election discussions are relatively balanced but the majority of opinions in relation to the local election are neutral. It should also be noted that due to the different sample sizes collected every day, the daily volume of each emotion is highly variable. See Appendix 1 for a full break down of the sentiment results for each sampled hashtag.

There are several limitations to the method used of extracting and classifying sentiment within the study. One limitation of using Umigon[30] as the sentiment classifier is its poor precision when it comes to finding negative sentiment. In a formal accuracy test its ability to precisely identify negative tweets was below 50%, with positive and neutral sentiment classification performing on a more accurate basis. One possible reason for this is the definition adopted to classify negative sentiment. This could potentially explain why the negative sentiment in the data sample was the lowest scoring and suggest an inaccuracy with the sentiment analysis which may potentially impair the results of this study.

4.2. FTSE100 data collection

4.2.1. Choice of sample days and the corresponding implications

FTSE does not operate over weekends or holidays so the data for 7th− 8th of May were missing. This would be a significant limitation of this study as it would impair the ability of like for like comparison and therefore the ability to calculate the relationship between public mood and stock market movements on a given day would be ompromised. In order to bypass this issue a method of approximating the missing FTSE100 values using a concave function proposed by Mittal and Goel[35] was adopted. If a FTSE100 value on a given day is x and the next available value is y with n days missing in between, the first missing value of x_{+1} can be approximated using$(y+x)/2$ and repeated for other missing values. Mittal and Goel[35] argue that this works for stock data as it usually follows a concave relationship, unless an incident of a sudden anomalous rise or fall occurs.

4.2.2. Choice of market indicators

In order to calculate the daily stock movement the opening and closing values of FTSE100 were obtained from Yahoo! Finance[48] as well as the volume of daily trades for each day in the given period. Table 1 displays the financial data collected. It should be noted that missing values are highlighted in red.

4.2.3. Volume as a basis of analysis

The volume based analysis of correlation compares the volume of daily tweets ‘VTWEETS’ against the volume of daily trades ’VFTSE′ and the daily closing price ’CLOSE′ as the dependent variables. The independent variables for this test can be defined as: volume of daily tweets ‘VTWEETS’, volume of daily positive tweets ‘VPOS’ and volume of daily negative tweets ‘VNEG’. However, before the relationship can be explored further each variable must be normalized using z-scores with the local mean and standard deviation as the basis. The formula for calculating the z-score of x for a dataset, where $\mu(X)$ and $\sigma(X)$ represent the mean and standard deviation respectively is defined as:

$$Z(x_i) = \frac{x_i - \mu(X)}{\sigma(X)}$$

4.2.4. Percentage change as a basis of analysis

The percentage change analysis aims to test for correlation between the independent variable of average daily mood ‘MOOD’ and the daily change in price ‘CHANGE’ as the dependent variable.

4.2.5. Independent variable − MOOD

In order to calculate MOOD, an assumption must be made about tweets collected containing neutral sentiment. For the purpose of this study their mood value is assumed as 0 and therefore they may be ignored when calculating MOOD. It should be noted that

Table 1
FTSE 100 market indicators.

Date	Opening price	Closing price	Change in price	Volume of daily trades
May-04	6185.6	6112	−0.01190	752,296,300
May-05	6112	6117.3	0.00087	725,515,650
May-06	6117.3	6125.7	0.00137	698,735,000
May-07	6121.5	6120.25	−0.00020	674,072,550
May-08	6123.6	6138.45	0.00243	658,873,425
May-09	6125.7	6114.8	−0.00178	649,410,100
May-10	6114.8	6156.65	0.00684	643,674,300
May-11	6156.65	6162.49	0.00095	555,870,400

the assumption that neutral sentiment has no implications is not strictly true, as Tang et al[43] found that mixed-neutral UGC amplifies the effect of positive and negative sentiment but indifferent-neutral UGC has the opposite effect. Nonetheless, the MOOD for a given day t is defined as:

$$MOOD_t = \frac{(VPOS_t - VNEG_t)}{VTOTAL_t}$$

The value of $MOOD_t$ can be anything between −1 and 1 where 1 represents a 100% positive mood in the sentiment of tweets for a given day t and −1 would define the inverse.

4.2.6. Dependent variable − CHANGE

The stock movement at a day t is defined as the normalized change in CLOSE from the previous day, which can be expressed as

$$CHANGE_t = \frac{CLOSE_t - CLOSE_{t-1}}{CLOSE_{t-1}}$$

This has made the parameters of MOOD equal to those of CHANGE and so the data can now be likened on a comparable scale.

Pearson's correlation will be the key method employed to establish the level of association between Twitter variables and FTSE variables for a given day$t$. Correlation is denoted by r and if found between variables it signifies that when a systematic change occurs in$x$, there is also a systematic change in$y$. The value of r may vary between −1 and 1 where a score of 1 is indicative of the strongest positive correlation possible, and −1 indicates the strongest negative. The closer the correlation is to zero the weaker the relationship between the variables. A relationship is said to be statistically significant if $p < 0.05$ p being the probability of obtaining the particular r value by chance. In order to successfully apply the method to specific sets of variables there are certain key assumptions that must be met, such as linearity and homoscedasticity. If a linear relationship does not exist between the variables then Pearson's cannot be used to find the correlation.

For the purpose of the study the multiple regression will also be used for volume based analysis which includes the independent variables of VTWEETS, VPOS and VNEG. In differentiation to a single regression model the output of the multiple regression explains the relationship that each individual independent variable has, on the dependent variable, when the other independent variables have also been taken into consideration. An assumption of linearity has been made when using the multiple regression analysis in this study.

The multiple regression models for this study can therefore be defined as follows:

$$VFTSE = a + \beta 1 VTWEETS + \beta 2 VPOS + \beta 3 VNEG + \varepsilon_t$$

$$CLOSE = a + \beta 1 VTWEETS + \beta 2 VPOS + \beta 3 VNEG + \varepsilon_t$$

where a is the intercept, β are the coefficients of the independent variables and ε_t is a random error term for day t.

For variables that have non-linear relationships, the predictors in the regression model can be transformed to account for curvature in the relationship. This is achieved by creating new (cubic or quadratic) variables representing the non-linear functions in the data. If the curvature is successfully modeled the new variable can express the curve function of the original variables as a linear function.

There is implied consent as the tweets collected as part of this study are made publicly available by the authors and the FTSE100 data is made public by Yahoo! (2016).[48] No personal data were used or revealed as part of the study which limits the risk of damage to the reputation of a participant. The tweets were extracted straight from the API rather than an existing database or collection which should reduce the possibility of bias.

5. Results

In order to investigate the existence of relationship between political discussions on Twitter, in terms of both volume and sentiment, and the stock market indicators for a given day, a variety of correlation and regression tests were carried out. Having collected the data, the variable VFTSE was missing the value for May 5th (Election Day) so for the purpose of this research the method of approximation was used.

Hypothesis 1: There is significant correlation between Twitter sentiment and volume statistics and stock market indicators for a given day.

5.1. Correlation between VTOTAL, VPOS, VNEG and VFTSE

Based on the findings of (Mao et al, 2012)[31] it is expected that positive correlation will exist between the variables VTOTAL and VFTSE. Another expectation is that the correlation between VPOS and VFTSE would be positive and the correlation between VFTSE and VNEG will be negative. To obtain an initial indication of existence and type of correlation between the defined variables, the z-score of VTOTAL, VPOS and VNEG was compared to the z-score of VFTSE using a Pearson's correlation test. It is assumed that the relationship between the variables is linear.

Before the correlation coefficients are analyzed it should be noted that the ability to draw many conclusions is limited as the *r* values do not allow for the effects of other independent variables. Table 2 confirms the expectation of positive correlation between VTOTAL and VPOS in relation to VFTSE and rejects the expectation of a negative correlation with VNEG. This implies that as volume of political discussions grows so does the volume of trades in the stock market, regardless of sentiment. As the correlations observed (0.284, 0.263 and 0.163) are significantly less than 0.5 it can be concluded that the relationship between volume of discussion and volume of trades, regardless of sentiment, is relatively weak and statistically insignificant as the significance values observed are above the threshold of $p < 0.05$. This means that the Hypothesis is rejected and the null Hypothesis is accepted as it has not been statistically disproved. The relationship between these variables is further tested using a multiple regression, on the assumption that the data are normally distributed.

The mid-range *r* value of 0.503 indicates a fairly significant linear relationship between the variables (see Table 3). The r^2 score of 0.249 indicates that 24.9% of the variation of VFTSE can be explained using the independent variables. The adjusted r^2, which is considered a better measure as it accounts for bias, shows a high negative result, which indicates that the regression model does not follow the trend of the data and ‘fits’ worse than a horizontal line would. In other words, this implies that the dataset violates the assumption of linearity required by this type of regression.

The *F* ratio in the ANOVA table (Table 4) measures whether the overall regression model is a good fit for the data set. The small value of 0.234 indicates that the regression model has a bad predictive capability which is confirmed by the insignificant Sig. value of 0.865. We also have information on which independent variables in the regression model are the best predictors of the VFTSE but as previous results have already indicated that the linear regression model is a bad fit for this dataset it is not necessary for evaluating the relationship of these variables. It can be concluded that the relationship between VFTSE and the independent variables cannot be computed using a linear regression model.

Table 2
Correlation coefficients for VTOTAL, VPOS, VNEG and VFTSE.

	VFTSE	VTOTAL	VPOS	VNEG
VFTSE				
Pearson Correlation	1	.276	.263	.158
Sig. (2-tailed)		.577	.618	.762
VTOTAL				
Pearson Correlation	.284	1	.947[a]	.958[a]
Sig. (2-tailed)	.576		.003	.001
VPOS				
Pearson Correlation	.263	.947[a]	1	.951[a]
Sig. (2-tailed)	.618	.003		.003
VNEG				
Pearson Correlation	.167	.958[a]	.951[a]	1
Sig. (2-tailed)	.761	.001	.003	

[a] Correlation is significant at the 0.01 level (2-tailed).

Table 3
Regression results for VTOTAL, VPOS, VNEG and VFTSE.

Model	R	R square	Adjusted R square	Std. error of the estimate
1	.509[a]	.249	−.872	1.37746

Predictors: (Constant), Zscore: VNEG, Zscore: VPOS tweets, Zscore: VTOTAL.
[a] Sig. (2-tailed).

Table 4
ANOVA[a] for VTOTAL, VPOS, VNEG and VFTSE.

Model	Sum of squares	df	Mean square	F	Sig.
1 Regression	1.354	3	.387	.234	.865[b]
Residual	2.894	2	1.793		
Total	4.000	6			

[a] Dependent Variable: Zscore: VFTSE.
[b] Predictors: (Constant), Zscore: VNEG, Zscore: VPOS, Zscore: VTOTAL.

5.2. Correlation between VTOTAL, VPOS, VNEG and CLOSE

In order to ascertain the strength of the relationship between the following set of variables the same set of steps were followed as previously. It is expected that there will be a positive correlation between VPOS and CLOSE and a negative correlation between VNEG and CLOSE. This is based on the assumption that the happier the general public mood is towards politics and the government, the higher the closing price of the market. Based on the findings of Mao et al,[31] it is expected that the relationship between VTOTAL and CLOSE is positive.

Table 5 confirms the expectation of negative correlation between CLOSE and VNEG, which suggests that a day with a low volume of negative tweets will most likely end with a higher close market price. The expectation is rejected for VPOS and for VTOTAL, which suggests a general trend that an increase in political discussions over Twitter leads to a lower closing price. As r values are around the 0.5 level it can be concluded that the correlation is relatively high but as the significance values are observed above the threshold of $p < 0.05$ the relationships are deemed as not statistically significant. This means that the Hypothesis is rejected and the null Hypothesis is accepted. The relationships between these variables is further tested using a multiple regression model, the output of which can be seen in Appendix 2. A negative adjusted R^2 and low value F ratio, 0.347, indicated that the linear model was a bad fit for this dataset also and therefore no conclusive assertions can be made due to the inappropriateness of the statistical analysis.

5.3. Correlation between MOOD and CLOSE

It is expected that there is a positive relationship between MOOD and CLOSE based on the findings of Rao and Srivastava.[38]

Table 6 confirms the expectation of a positive correlation between MOOD and CLOSE but as the r values are below the 0.5 level it can be concluded that the correlation is relatively low. The significance values are observed as above the threshold of $p < 0.05$ and therefore the relationship is not statistically significant. This means that the Hypothesis is rejected and the null Hypothesis is accepted. Pearson's correlation test is not truly conclusive for this data set as Fig. 1 shows a clear violation of the linearity rule which is a key assumption of the test. Applying the line of best fit/regression line identifies the relationship between these variables as non-linear with a quadratic curve function. A quadratic curvilinear regression is run to explore the relationship further.

Table 5
Correlation coefficients for VTOTAL, VPOS, VNEG and CLOSE.

	CLOSE	VTOTAL	VPOS	VNEG
CLOSE				
Pearson Correlation	1	.549	−.479	−.548
Sig. (2-tailed)		.263	.326	.241
VTOTAL				
Pearson Correlation	−.549	1	.947[a]	.958[a]
Sig. (2-tailed)	.263		.003	.001
VPOS				
Pearson Correlation	.479	.921[a]	1	.951[a]
Sig. (2-tailed)	.326	.003		.003
VNEG				
Pearson Correlation	−.548	.958[a]	.951[a]	1
Sig. (2-tailed)	.241	.001	.003	

[a] Correlation is significant at the 0.01 level (2-tailed).

Table 6
Correlation coefficient for MOOD and CLOSE.

	Closing price
Average change in price	
Pearson Correlation	.381
Sig. (2-tailed)	.457

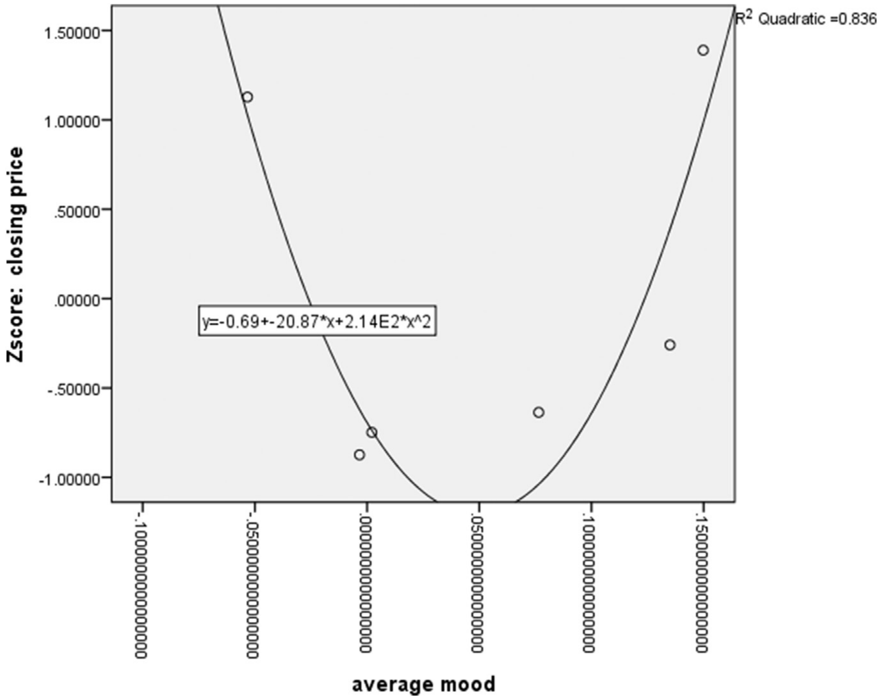

Fig. 1. Scatter plot for MOOD and CLOSE.

A high positive r value, 0.923, confirms the initial expectation of a positive correlation and suggests a strong relationship between the variables. Using the quadratic variable as part of the function improved r significantly, from 0.162 to 0.923, suggesting that the quadratic function is optimal for this data set. This is further confirmed by the adjusted r^2 going from a negative to positive value (see Table 7). The new adjusted r^2 indicates that 74.4% of the variance in CLOSE can be explained by MOOD making this a successful relationship model.

The F-ratio in the ANOVA examines how well the model predicts the dependent variable and based on the score of 7.482 there is an indication that the mood has a good predictive capability (see Table 8). Nonetheless, as the p falls above the threshold of $p < 0.05$ the relationship between MOOD and CLOSE is found statistically insignificant. Therefore, the Hypothesis is rejected and the null Hypothesis is expected. However, the p value of 0.069 is only marginally out of the significance zone, so although we are unable to conclude with certainty that the change in CLOSE is a direct result of change in MOOD based on this specific sample, we can conclude that a trend exists between this set of variables. A lack of significance in times of high correlation can usually be explained by a too small sample size, a retest with a bigger sample period would either confirm the relationship and its significance, or reject the trend as an issue of sample/technique error in these tests.

5.4. Correlation between MOOD and CHANGE

It is expected that there is a positive relationship between MOOD and CHANGE based on the assumption that the general mood on the day would signal the change movements in the stock market.

Table 7
Regression results for MOOD and CLOSE.

Model	R	R square	Adjusted R square	Std. error of the estimate	R square change	F change	df1	df2	Sig. F change
1	.162[a]	.031	−.234	1.267385	.031	.112	1	4	.784
2	.923[b]	.857	.744	.5638327	.821	16.439	1	3	.045

[a] Predictors: (Constant), MOOD.
[b] Predictors: (Constant), MOOD, MOOD_squared.

Table 8
ANOVA[a] for MOOD and CLOSE.

Model	Sum of squares	df	Mean square	F	Sig.
1 Regression	.143	1	.143	.107	.794[b]
Residual	4.764	4	1.356		
Total	4.000	5			
1 Regression	4.285	2	2.174	7.482	.069[c]
Residual	.843	3	.254		
Total	4.000	5			

[a] Dependent Variable: Zscore: CLOSE.
[b] Predictors: (Constant), MOOD.
[c] Predictors: (Constant), MOOD, MOOD_squared.

Table 9 confirms the expectation of a positive correlation between MOOD and CHANGE. As r values are close to 0 it can be concluded that the correlation is relatively low and as the significance values are observed above the threshold of $p < 0.05$ that the correlations are not statistically significant and therefore no relationship is conclusively identified. This means that the Hypothesis can be rejected and the null Hypothesis is accepted. Pearson's correlation test is not truly conclusive for this data set as Fig. 2 shows a clear violation of the linearity rule which is an assumption of the test. Inserting the line of best fit/regression line identifies the relationship between these variables as non-linear with a cubic curve function.

In order to test the slight relationship identified previously a linear regression was modeled using a cubic curve function.

The low R value, 0.179, confirms the weak positive correlation between variables initially identified by Pearson's (see Table 10). Using the cubic variable as part of the function improved R marginally – from 0.173 to 0.179 – but the negative increase in the adjusted r^2 suggests that the cubic function linear regression is not the optimal fit for modeling the relationship of this data. A negative adjusted r^2 and low F value of 0.039 indicate that the linear model was a bad fit for the dataset and therefore no conclusive assertions can be made about the relationship of MOOD and CHANGE without increasing the sample to get a better picture of the relationship between variables and the use of a better suited statistical model (see Table 11).

Hypothesis 2: There is a statistically significant predictive relationship between Twitter sentiment statistics and stock market indicators.

5.5. Correlation between CHANGE and MOOD 1–3 days before

To obtain an indication of whether CHANGE could be predicted based on MOOD for a time lag of 1, 2 and 3 days a Pearson's correlation coefficient was calculated (see Table 12). For a time lag of 1 day there is a low positive

Table 9
Correlation coefficient for MOOD and CHANGE.

	Price
Mood	
Pearson Correlation	.171
Sig. (2-tailed)	.764

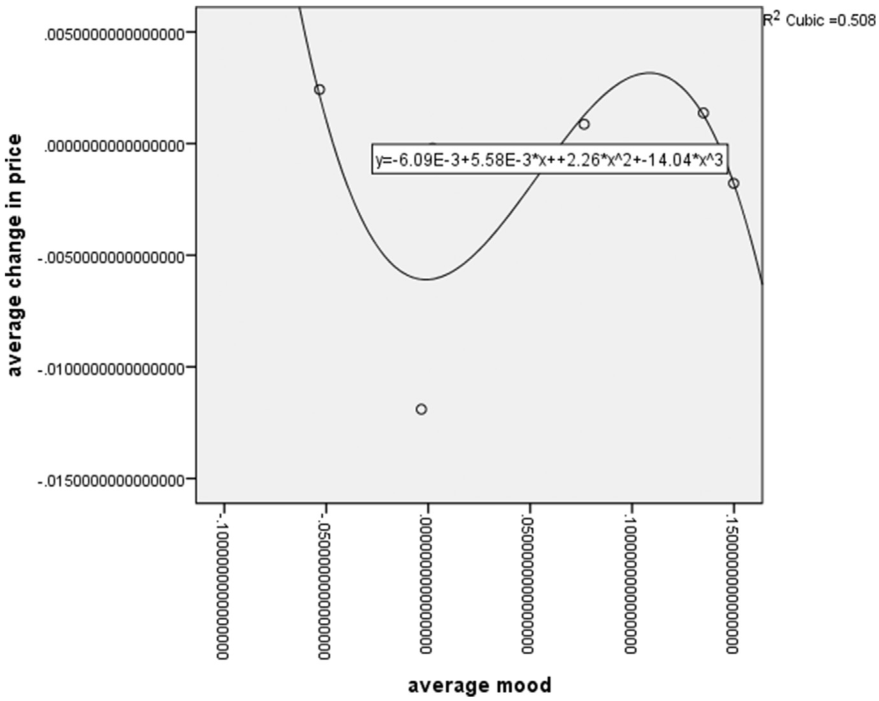

Fig. 2. Scatter plot for MOOD and CHANGE.

Table 10
Regression results for MOOD and CHANGE.

Model	R	R square	Adjusted R square	Std. error of the estimate
1	.173[a]	.031	−.243	.00673
2.	.179[b]	.028	−.587	.00598

[a] Predictors: (Constant), MOOD.
[b] Predictors: (Constant), MOOD, MOOD_cubed.

Table 11
ANOVA for MOOD and CHANGE ANOVA.[a]

Model	Sum of Squares	df	Mean Square	F	Sig.
1 Regression	.000	1	.000	.123	.754[b]
Residual	.000	4	.000		
Total	.000	5			
1 Regression	.000	2	.000	.039	.967[c]
Residual	.000	3	.000		
Total	.000	5			

[a] Dependent Variable: PRICE.
[b] Predictors: (Constant), MOOD.
[c] Predictors: (Constant), MOOD, MOOD_cubed.

correlation, 0.187, meaning that if the mood of politics-related discussions on Twitter increased in positivity, the following day the change in the closing price of the FTSE would be a marginal increase. For a time lag of 2 days there is a relatively high positive correlation, 0.587, meaning that if the mood of political discussions on Twitter increased in positivity, in 48 h the change in the closing price of the FTSE would be a significant increase. For a time lag of 3 days there is a relatively low negative correlation, −0.263, meaning that if the mood of politics-related discussions on Twitter became more negative, in 72 h the change in the closing price of the FTSE would be a marginal increase.

Table 12
Correlation coefficients for MOOD-1,-2 and -3 days and CHANGE.

	PRICE day 6-1
MOOD day before	
Pearson Correlation	.187
Sig. (2-tailed)	.769
MOOD 2 days before	
Pearson Correlation	.587
Sig. (2-tailed)	.442
MOOD 3 days before	
Pearson Correlation	−.263
Sig. (2-tailed	.824

However, the p values are above the significant threshold of $p < 0.05$ so no statistically significant relationship is found. The Hypothesis must be rejected and the null Hypothesis accepted as it has not been statistically disproved.

6. Discussion

This study endeavored to investigate the relationship between the mood of election related chatter on Twitter and FTSE 100 performance in the same period. Although, sample size prevented the observations from being confirmed as statistically significant, they did show some trends in relationships. The main findings and their implications are discussed in this section.

6.1. Volume-based analysis

It was expected before the commencement of this study that an increase in political discussions over Twitter (VTOTAL) would encourage the daily trading of shares on the stock market (VFTSE), based on the premise that emotional arousal leads to action (Debele et al, 2007). Although on the surface the positive relationship between volume of discussion, regardless of sentiment, and volume of trading as evidenced by Table 2 suggests just so, the low score of r and high score of p indicate otherwise. These findings directly contradict the findings of Mao et al[31] who established a strong correlation between volume of tweets and volume of trades among other financial performance indicators. A possible explanation of this is the dominance of neutral tweets as a proportion of the total sample of tweets collected (64%). Tang et al[43] found that one type of neutral UGC known as 'indifferent' reduces the impact of positive and negative UGC, hence it might be speculated that high levels of neutral tweets, failed to stir up enough emotion in the local election discussions to have a significant contagion effect on the volume of trades. In order to avoid this potential problem in the future it is suggested that samples of tweets should be selected using emotional phrases (such as happy, sad and I feel) as witnessed in successful research by Bollen et al, Kaminski and Gloor and Zhang et al.[8,51]

The r values observed in Table 5 suggest that an increase in political discussions (VTOTAL), regardless of sentiment, is likely to lead to a lower CLOSE price. This goes against the observations of Klibanoff et al, Chan and Wei and Mao et al,[11,28,31] all of whom established a positive relationship between close price and volume of chatter. A possible explanation for this is that an increase in political debates increases uncertainty in the government, leading to a decrease in confidence and as a result in investment. This observation is supported by Mei and Guo[33] who determined the existence of a significant relationship between financial crises (extremely bad stock performance) and political uncertainty and Parker[37] who found that trust in the government is a strong consideration for stock market participators. One possible reason why this study failed to identify a significant relationship between stock market indicators and the volume of political discussions, is that the studies that found the relationship such as Klibanoff et al and Niederhoffer[28,36] were based on 'major' and 'world' events, however, this study was limited to a local election which has restricted making more general investigations. It may be speculated that using a more influential political event such as the EU Referendum is likely to find a more significant relationship as leaving the EU potentially has stronger implications on UK businesses.

6.2. Sentiment based analysis

Using a curvilinear regression established that 71.6% of variance in CLOSE can be explained by changes in MOOD at an almost significant level of $p = 0.067$. This observation has implications for Question 2 of this study as it indicates that there is a potential that a model can be used to predict changes in CLOSE based on changes in MOOD for a given day as observed in Zhang et al, Rao and Srivastava and Bollen et al,[8,38,50,51] with high levels of accuracy. However the period of collection must be extended initially, in order to aid in the determination of a significant relationship. Unexpectedly, study results as per Table 9 showed a weak relationship between MOOD and CHANGE. It would be expected that changes in the mood in regards to politics would be reflected in the movements of the FTSE100 on a given day based on the findings of other studies such as Rao and Srivastava, Bollen et al and Kaminski and Gloor.[8,38] One potential reason why this observation was not achieved in this study is due to the large number of neutral tweets diluting the average mood calculation to an insignificant percentage. As the average mood was calculated using total tweets, a large proportion of neutral tweets lower the average mood, in comparison to the level of change in price, hence the lack of a significant relationship. It would be advised that a similar study was carried out in the future that neutral tweets should be excluded from the sample. In extension to this, this result could also be due to a nonlinear relationship between variables which would impair the correlation calculations. A linear and a cubic regression were both found inadequate for measuring this relationship and it is suggested that in the interest of future research a better suited model may find a stronger relationship similar to the relationship of MOOD and CLOSE.

Based on Table 12 there is a weak positive correlation between stock market movements and the MOOD of the previous day but a slightly higher negative correlation between the variables for a 3 day time lag. Meaning that if the public mood in relation to politics fell on day t, on day t+3 the CHANGE would be a marginal increase. This directly contradicts De Bondt and Thaler[18] who suggest that in revising their beliefs individuals assign more weighting to recent information and underweight prior data. The strongest predictor time lag for mood is 2 days, it is expected that if MOOD increased on day t, on day t +2 there will be a relatively high increase in CHANGE. Although these relationships were not statistically significant they show a promising start for the use of politic discussions on Twitter for forecasting future market performance.

7. Conclusion

Motivated by the prospect of a forecasting relationship between politics and market performance[28,37] and the prospect of evaluating Twitter as a social media tool in a political context, data mining methods using the Twitter API allowed for a preliminary study to be carried out investigating the relationship between political mood and stock market performance. Lexicon based sentiment analytics allowed for data sentiment to be classified in order to measure the general public mood in relation to the local elections and stock market indicators such as volume of trades, market closing price and the average daily change in price were collected to track the market movements around the same time as the election. A range of correlation and regression based tests determined the existence or lack thereof of relationships between the variables in context and allowed this investigation to explore the influence of general public opinions on investment decision making.

Although the sample size prevented this study from acquiring statistically significant results in regards to the relationship between Twitter chatter and stock market movements, the following trends can be observed:

1. MOOD and CLOSE show potential of a strong causation relationship in this sample, almost at the significance level.
2. There are promising correlations between CHANGE and MOOD with various time lags highlighting the possibility of using Twitter chatter as a forecasting tool for future FTSE performance.
3. Contradictory to past research, volume based predictors (VTOTAL, VPOS and VNEG) have weak ties with VFTSE and CLOSE, highlighting the superiority of sentiment analytics over volume analytics in stock market prediction studies.

In order to conclude, the original research questions must be reflected upon. In terms of Question 1, observations in this study have determined that correlation does exist between public sentiment in regards to the local elections and the

FTSE movements, in the form of correlation between MOOD and CLOSE; however, the relationship is not yet determined as statistically significant. In relation to Question 2, findings in this investigation suggest that there is evidence of causation between public sentiment and the stock market movements, in terms of the relationship between MOOD and CLOSE, and the time lag findings of MOOD and PRICE, however, they are also yet to be determined statistically significant.

We can identify some key limitations of this investigation: the first issue with the sample size is whether it is a good representation of the views of the UK Twitter population. The final tweet sample in this investigation contained over 60,000 tweets over a 6 day period. Considering Twitter is updated hundreds of millions times a day this seems an insignificantly sized sample to be a fair representation of the population mood. Having said that, the sample of this study should aim to be a good representation of the political discussions on Twitter by the UK population, and as the total number of them is unknown, it is difficult to say whether the sample of Tweets is a fair representation of the general mood of the UK public. The second and more detrimental issue to the study is the size of the specified sample period. Although the 6 day period was specifically chosen due to the aim of the study to measure the existence of a short term relationship only, based on the study of an event (Local Election). The size of the sample was not big enough for the relationships to be classed as significant at the 95% certainty, due to the fact that relationship trends are not prominent enough in a smaller sample. However, just because a relationship is found as not statistically significant, it does not mean that the relationship does not exist. Finally, it should be noted that the mood expressed in this study is reflective of the discussions specifically related to the issue of the local elections of 2016. General day-to-day political sentiment was not measured as a part of this study and relationship with the stock market may vary to the observations as per this paper.

This early study looking into the relationship between political discussions and market performance in the UK shows promise for using sentiment analytics on Twitter data for potentially forecasting market movements. The results highlight groups of variables with stronger cause and effect relationships that should be the focal point of similar studies in the future (MOOD and CLOSE). Further studies may wish to explore the relationship between Twitter based political discussions and stock market movements using a broader sample of Tweets or an extended period of observation. Based on the trends as indicated by the results it is believed that in an extended investigation looking at the EU Referendum/National Elections as examples there is a possibility that the mood of the Twitter debates at the time, can be compared to stock market movements in order to measure the influence of political uncertainty in Twitter chatter on UK businesses, and to predict the performance of UK businesses via investment indicators based on public perceptions on the economic conditions.

Acknowledgement

We wish to thank a research program participant for her excellent research assistance.

Appendix 1. Results of sentiment analysis per hashtag per day

#London Elects	Pre election-May 4		Election-May 5		Post election-May 6		Post election-May 7		Post election-May 8		Post election-May 9	
	Sum tweets	% of total	Sum tweets	% of total	Sum tweets	% of total	Sum tweets	% of total	Sum tweets	% of total	Sum tweets	% of total
Positive	13	18	53	14	4799	30	2422	17	418	24	126	31
Negative	10	14	53	14	2559	16	2422	17	522	30	59	14
Neutral	48	86	265	72	8637	54	9405	66	801	46	214	55
Total	71	100	371	100	15995	100	14249	100	1741	100	399	100
#Mayoral election												
Positive	59	12	255	16	111	20	26	17	3	9	1	3
Negative	70	14	186	11	98	18	13	8	2	6	1	3
Neutral	352	74	1188	73	321	62	109	75	27	86	28	94
Total	481	100	1629	100	530	100	148	100	32	100	30	100

(*continued*)

#London Elects	Pre election-May 4		Election-May 5		Post election-May 6		Post election-May 7		Post election-May 8		Post election-May 9	
	Sum tweets	% of total	Sum tweets	% of total	Sum tweets	% of total	Sum tweets	% of total	Sum tweets	% of total	Sum tweets	% of total
#Polling day												
Positive	8	19	4525	19	190	18	30	17	14	13	11	16
Negative	2	4	2620	11	68	6	12	6	11	10	4	6
Neutral	32	77	16672	70	812	76	130	77	77	77	50	78
Total	42	100	23817	100	1070	100	172	100	102	100	65	100
#London elects												
Sum positive	80	13.5%	4833	18.7%	5100	29.0%	2478	17.0%	435	23.2%	138	27.9%
Sum negative	82	13.8%	2859	11.1%	2725	15.5%	2447	16.8%	535	28.5%	46	13.0%
Sum neutral	432	72.7%	18125	70.2%	9770	55.5%	9644	66.2%	905	48.3%	292	59.1%
Sum total	594	100.0%	25817	100.0%	17595	100.0%	14569	100.0%	1875	100.0%	494	100.0%
Sum MOOD	−0.00337		0.07646		0.13498		0.00213		−0.05333		0.14980	

Appendix 2. Multiple regression results for VTOTAL, VPOS, VNEG and CLOSE

ANOVA[a]

Model		Sum of squares	Df	Mean square	F	Sig.
1	Regression	1.710	3	.570	.347	.800[b]
	Residual	3.290	2	1.645		
	Total	5.000	5			

[a] Dependent Variable: Zscore: closing price.
[b] . Predictors: (Constant), Zscore: volume of negative tweets, Zscore: volume of positive tweets, Zscore: volume of tweets.

Results[a]

Model		Unstandardized coefficients		Standardized coefficients	t	Sig.	95.0% Confidence interval for B	
		B	Std. error	Beta			Lower bound	Upper bound
1	(Constant)	6.615E-15	.524		.000	1.000	−2.253	2.253
	Zscore: volume of tweets	−.638	2.331	−.638	−.274	.810	−10.670	9.393
	Zscore: volume of positive tweets	.628	2.016	.628	.312	.785	−8.045	9.302
	Zscore: volume of negative tweets	−.532	2.298	−.532	−.231	.839	−10.420	9.356
R		.585[b]						
R Square		.342						
Adjusted R Square		−.645						

[a] Dependent Variable: Zscore: closing price.
[b] Predictors: (Constant), Zscore: volume of negative tweets, Zscore: volume of positive tweets, Zscore: volume of tweets.

References

1. Antweiler W, Frank MZ. Is all that talk just Noise? The information content of internet stock message boards. *J Finance*. 2004;59(3):1259—1294.
2. Asur S, Huberman BA. Predicting the future with social media. In: *Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*. 1. 2010:492—499.
3. Atefeh F, Khreich W. A survey of techniques for event detection in Twitter. *Comput Intell*. 2015;31(1):132—164.
4. Baker M, Wurgler J. Investor sentiment and the cross-section of returns. *J Finance*. 2006;61:1645—1680.
5. Bakshy E, Hofman JM, Watts JD, Mason WA. Everyone's an influencer: quantifying influence on twitter. In: *Proceedings of the 4th ACM International Conference on Web Search and Data Mining, Hong Kong , WSDM*. 2011:65—74.
6. BBC News. *Guide to May 2016 elections in Scotland, Wales, England and Northern Ireland*. BBC News; 2016, 5 May. Available from: http://www.bbc.co.uk/news/uk-politics-35813119. Accessed May 9, 2016.
7. Berger J, Milkman KL. What makes online content viral? *J Market Res*. 2012;49(2):192—205.
8. Bollen J, Mao H, Zeng X. Twitter mood predicts the stock market. *J Comput Sci*. 2011;2(1):1—8.
9. Brooker P, Barnett J, Cribbin T, Sharma S. Have we even solved the first 'big data challenge?' Practical issues concerning data collection and visual representation for social media analytics. In: Snee H, Hine C, Morey Y, Roberts S, Watson H, eds. *Digital Methods for Social Science: An Interdisciplinary Guide to Research Innovation*. Palgrave Macmillan; 2015. Available from: http://www.palgrave.com/page/detail/Digital-Methods-for-Social-Science/?sf1=barcode&st1=9781137453655. Accessed May 9, 2016.
10. Cha M, Haddadi H, Benevenuto F, Gummadi KP. Measuring user influence in twitter: the million follower fallacy. In: *Paper Presented at: Proceedings of the 4th International AAAI Conference on Weblogs and Social Media, Washington*. 2010:23—26.
11. Chan YC, Wei KCJ. Political risk and the stock price volatility: the case of Hong Kong. *Pac Basin Finance J*. 1996;4(2—3):259—275.
12. Chen Y, Xie J. Online consumer review: word-of-mouth as a new element of marketing communication mix. *Manag Sci*. 2008;54(3):477—491.
13. Cheong M, Lee VCS. A microblogging-based approach to terrorism informatics: exploration and chronicling civilian sentiment and response to terrorism events via Twitter. *Inf Syst Front*. 2011;13(1):45—59.
14. Cheong M, Lee VCS. A study on detecting patterns in twitter intra-topic user and message clustering. In: *Paper Presented at: 20th International Conference on Pattern Recognition, Turkey*. 2010:23—26.
15. Chevalier JA, Mayzlin D. The effect of word of mouth on sales: online book reviews. *J Market Res*. 2006;43:345—354.
16. Comm J. *Twitter Power 2.0: How to Dominate Your Market One Tweet at a Time*. New Jersey: Wiley and Sons; 2010.
17. Cutler MD, Poterba JM, Summers LH. What moves stock prices? *J Portfolio Manag*. 1989;15(3):4—12.
18. De Bondt WFM, Thaler R. Does the stock market overreact? *J Finance*. 1985;40(3):793—805.
19. Dellarocas C, Zhang X, Awad NF. Exploring the value of online product reviews in forecasting sales: the case of motion pictures. *J Interact Market*. 2007;21(4):23—45.
20. Dhar V, Chang E. Does chatter Matter? The impact of user-generated content on music sales. *J Interact Market*. 2007;23(4):285—340.
21. Dobele A, Lindgreen A, Beverland M, Vanhamme J, Van Wijk R. Why pass on viral messages? Because they connect emotionally. *Bus Horiz*. 2007;50:291—304.
22. Hanna A, Wells C, Maurer P, Shah DV, Friedland L, Matthews J. Partisan alignments and political polarization online: a computational approach to understanding the French and US presidential elections. In: Weber I, Popescu A, Pennacchiotti M, eds. *PLEAD 2013: Proceedings of the 2nd Workshop Politics, Elections and Data*. New York, NY: ACM; 2013:15—21.
23. Hansen LK, Arvidsson A, Nielsen FA, Colleoni E, Etter M. Good friends, bad news — affect and virality in twitter. *Future Inf Technol*. 2011;185:34—43.
24. Java A, Song X, Finin T, Tseng B. Why we Twitter: understanding microblogging usage and communities. In: *Procedings of the Joint 9th WEBKDD and 1st SNA-KDD Workshop 2007*. San Jose: Springer; 2007:55—65, 12 August.
25. Jenders M, Kasneci G, F Naumann. Analyzing and predicting viral tweets. In: *WWW '13 Companion Proceedings of the 22nd International Conference on World Wide Web*. New York: ACM; 2013:657—664.
26. Jungherr A. *Twitter in Politics: A Comprehensive Literature Review*; 2014. Available from: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2402443. Accessed May 9, 2016.
27. Kaminski J, Gloor PA. *Nowcasting the Bitcoin Market with Twitter Signals*; 2014. Available from: http://jermainkaminski.com/papers/BTC_Nowcasting_Abstract.pdf. Accessed May 9, 2016.
28. Klibanoff P, Lamont O, Wizman T. Investor reaction to salient news in closed-end country funds. *J Finance*. 1998;53(2):673—699.
29. Kwak H, Lee C, Park H, Moon S. What is twitter, a social network or a news media?. In: *WWW 2010: Proceedings of the 19th International Conference on World Wide Web*. New York: ACM; 2010:26—30. April, 591-600.
30. Levallois C. Umigon: sentiment analysis on Tweets based on terms lists and heuristics. In: *Paper Presented at: Proceedings of the 7th International Workshop on Semantic Evaluation, Georgia*. 2013.
31. Mao Y, Wang B, Wei W, Liu B. Correlating S&P 500 stocks with twitter data. In: *Proceedings of the First ACM International Workshop on Hot Topics on Interdisciplinary Social Networks Research, New York*. 2012:12—16, 69-72.
32. McGee M. *Twitter: 60% of Users Access via Mobile*; 2012. http://marketingland.com/twitter-60-percentof-users-access-via-mobile-13626. Accessed March 22, 2016.
33. Mei J, Guo L. Political uncertainty, financial crisis, and market volatility. *Eur Financ Manag*. 2004;10(4):639—657.

34. Miller V. New media, networking and phatic culture. *Convergence: The International Journal of Research into New Media Technologies*. 2008;14(4):387−400.
35. Mittal A, Goel A. *Stock Prediction Using Twitter Sentiment Analysis*; 2009. Available from: https://pdfs.semanticscholar.org/4ecc/55e1c3ff1cee41f21e5b0a3b22c58d04c9d6.pdf. Accessed May 9, 2016.
36. Niederhoffer V. The analysis of world events and stock prices. *J Bus*. 1971;44(2):193−219.
37. Parker WD. *The Socionomic Perspective on Social Mood and Voting: Report on New Mood Measures in the 2006 ANES Pilot Study*; 2007. ANES Pilot Study Report. Available from: http://www.electionstudies.org/resources/papers/Pilot2006/nes011884.pdf. Accessed May 9, 2016.
38. Rao T, Srivastava S. Analyzing stock market movements using twitter sentiment analysis. In: *ASONAM '12 Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining, Washington, 26 August*. Washington: IEEE Computer Society; 2012:119−123.
39. Romero DM, Meeder B, Kleinberg J. Differences in the mechanics of information diffusion across topics: idioms, political hashtags and complex contagion on Twitter. In: *Paper presented at: Proceedings of the 20th International conference on WWW, India, March 28- April 1*. 2011.
40. Shamma DA, Kennedy L, Churchill EF. Tweet the debates: understanding community annotation of uncollected 86 sources. In: Boll S, Hoi SCH, Luo J, Jin R, King I, Xu D, eds. *WSM 2009: Proceedings of the 1st SIGMM Workshop on Social Media, Ed. Susanne Boll*. New York, NY: ACM; 2009:3−10.
41. Small TA. What the hashtag? *Inf Commun Soc*. 2011;14(6):872−895.
42. Smith AN, Fischer E, Yongjian C. How does brand-related user-generated content differ across Youtube, Facebook and twitter? *J Interact Market*. 2012;26:102−113.
43. Tang T, Fang E, Wang F. Is neutral really Neutral? The effects of neutral user-generated content on product sales. *J Market*. 2014;78(4):41−58.
44. Trendogate. *Trends in United Kingdom*; 2016. Available from: https://trendogate.com/placebydate/23424975/2016-05-06. Accessed May 9, 2016.
45. Tumarkin R, Whitelaw R. New or noise? Internet postings and stock prices. *Financ Anal J*. 2001;57(3):41−51.
46. Tumasjan A, Sprenger TO, Sandner PG, Welpe IM. *Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment*. 2010. IN.
47. Twitter. *Company: About*; 2016. Available from: https://about.twitter.com/company. Accessed May 9, 2016.
48. Yahoo. *FTSE Historical Prices*; 2016. Available from: https://finance.yahoo.com/q/hp?s=%5EFTSE+Historical+Prices. Accessed May 9, 2016.
49. Zach T. Political events and the stock market: evidence from Israel. *Int J Bus*. 2003;8(3):243−266.
50. Zhang X, Fuehres H, Gloor PA. Predicting asset value through twitter buzz. *Adv Collective Intell*. 2011a;113:23−34.
51. Zhang X, Fuehres H, Gloor PA. Predicting Stock Market Indicators through Twitter “I hope it is not as bad as I fear”. *Soc Behav Sci*. 2011b;26:55−62.
52. Zheludev I, Smith R, Aste T. When can social media lead financial markets? *Sci Rep*. 2014;4:1−11.