Twitter Sentiment Analysis to Predict Bitcoin Exchange Rate

Ciaran McAteer

A dissertation submitted to the University of Dublin in partial fulfilment of the requirements for the degree of MSc in Management of Information Systems

2014

Declaration

I declare that the work described in this dissertation is, except where otherwise stated, entirely my own work, and has not been submitted as an exercise for a degree at this or any other university. I further declare that this research has been carried out in full compliance with the ethical research requirements of the School of Computer Science and Statistics.

Signed: ________________________ Date: _________________

Ciaran McAteer

Sept 2014

Permission to Lend or Copy

I agree that the School of Computer Science and Statistics, Trinity College Dublin may lend or copy this dissertation upon request.

Signed: ________________________   Date: _________________

Ciaran McAteer

Sept 2014

Acknowledgements

I would like to thank my supervisor Susan Leavy for her support and advice throughout this dissertation.

Thanks also to the lecturers and staff of Trintiy College Dublin.

And finally to my wife Muireann and sons Manus, Conall and Senan for their ongoing support.

Abstract

The microblogging platform Twitter has become a valuable source of user sentiment. This paper presents an evaluation of Twitter sentiment as a useful metric for predicting financial markets, specifically the bitcoin exchange rate. The tweets associated with the bitcoin digital currency are tracked in order to determine if the user sentiment contained within those tweets reflects the exchange rate of the currency. The sentiment of users' tweets is categorised as having a positive, negative or neutral opinion of the virtual currency using machine learning techniques. Time series analysis is performed which reveals that there is a positive correlation between the Twitter sentiment and the bitcoin exchange rate, and that sentiment is reflected in price after a time delay of 24 hours. Other aspects of Twitter, such as volume of tweets related to the subject, and a separate analysis of retweets, also observe a relationship to the bitcoin digital currency.

Table of Contents

TABLES

FIGURES

1 Introduction

1.1 Introduction

The purpose of this chapter is to provide background information related to the research question selected for this paper. The research topic is introduced, as are the main research question and sub-questions. This chapter also provides background on the topic and the reasons why this research question was selected. The scope of the research, its importance and the beneficiaries are discussed.

1.2 Research Background

Sentiment can be defined in its simplest terms as "*a view or opinion that is held or expressed"* (OxfordEnglishDictionary, 2014). In terms of financial markets, sentiment can be viewed as being positive (bullish), negative (bearish) or neutral about a certain investment (Brown and Cliff, 2004). Harvesting sentiment has long been used as a mechanism for predicting economic trends, surveys of sentiment such as the Consumer Sentiment Index and Purchasing Managers' Index being two examples of this. With the advent of the information age the ability to identify and categorise this sentiment has become increasingly important for businesses and researchers alike. Businesses want to know consumer opinions about their products and services (Liu, 2012). Potential customers want to know the opinions of existing users before they purchase a product (Pang and Lee, 2008). As the information posted by users online covers a broad set of topics, researchers can use online sentiment not only in field of computer science but also in the fields of social sciences and management sciences (Liu, 2012). Advances in machine learning and processing power allow computers to perform analysis of this sentiment in real time and on a very large scale.

The term sentiment analysis (or opinion mining) broadly refers to the computational treatment of sentiment, opinion and subjectivity from text (Pang and Lee, 2008). This paper uses the technique of classification to categorise Twitter messages according to their sentiment. Classification is the task of identifying which category a value belongs to. In the context of text classification it means labelling natural language texts with categories from a predefined set (Sebastiani, 2002). Classification is a type of supervised learning, that is, correctly categorised items of text are made available to train the classifier. Researchers can take advantage of sites that provide ratings along with customer reviews to build corpuses of automatically categorised data from sites such as Amazon and Rotten Tomatoes in order create this training data (Pang and Lee, 2008).

There are many different sources of sentiment online including websites, blogs, and social networking sites like Facebook and Twitter. The use of natural language processing, text analysis and computational linguistics enables computers to identify subjective human communication and classify it. This practice is common place amongst large organisations, with many software providers (such as IBM and SAS) now offering solutions to allow corporate customers to perform analysis of customers' views in relation to their brand or product. Social networking sites offer opportunities as a new source of information to harvest user sentiment in real time and on a much larger scale than was previously possible. The volumes of data being produced by social networking sites on a daily basis far exceeds what would be practical with human users classifying this data. Thus the explosion of use of social networking sites has seen a parallel explosion in research using sentiment analysis (Liu, 2012). Pang and Lee (2008) suggest that 2001 was the year that research into sentiment analysis became widespread, as researchers became aware of the opportunities of online data, and that it has been increasing since.

Twitter recently announced the results of their 'Twitter Data Grant', an initiative to allow researchers access to the full Twitter live and historical data set. They received 1,300 proposals from research institutions, finally selecting 6 institutions to be allocated access to the data (Twitter, 2014b). The 6 research proposals cover health care (2), sports science, disaster and flood analysis (2) and human happiness. The fact that the areas being researched are so diverse is an indication of the information that can be extracted from these sites both directly, in the form of user's own opinion and thoughts, and indirectly in the form of who follows whom and what they retweet. Previously researchers have used Twitter as a source of sentiment and opinion across multiple topics: finance (Bollen et al., 2011, Sprenger et al., 2013), politics (Conover et al., 2011, Wang et al., 2012), and geopolitical topics (Huang, 2011, Howard et al., 2011). Users of services like Twitter speak openly about how they feel about the brands, products or services they use. The opinions spread quickly through the network magnifying the word of mouth effect (Hennig-Thurau et al., 2012). In one sense social networking sites like Twitter and Facebook have become a huge pool of consumer sentiment and public opinion (Pak and Paroubek, 2010).

1.2.1 Bitcoin – A currency for a digital age

Bitcoin originated from a white paper (Nakamoto, 2008) and subsequent open source software implementation from a person going by the name Satoshi Nakamoto. The real identity of Satoshi Nakamoto is unknown. Whether or not this name is the pseudonym of an individual or a group is also unknown. His involvement with the project ended in 2010 but the bitcoin community has grown with many developers contributing to it (bitcoin.org, 2014). It is the first example of a crypto-currency (a digital currency that uses cryptography to control its creation and transactions) and provides decentralised peer-to-peer financial transactions without going through a financial institution.

Bitcoin is an implementation of a crypto-currency based on the concept described by the cryptographer Wei Dai in 1998. One of the main problems with a digital currency is the concept of double spending - if the currency unit can be represented as a text in a file (as opposed to physical paper or coin), then what stops the holder of the currency spending it multiple times. The conventional answer to this problem was to have a central ledger to track all transactions, and a trusted central authority to administer it. The Satoshi solution was to remove the dependency on a central authority and publicly distribute the ledger, in what is known as the 'block chain'. This makes Bitcoin a distributed and peer-to-peer digital currency with no one point of failure, or point of weakness, for attack. Despite this, there have been numerous attacks on the surrounding ecosystem that have rocked the bitcoin community. Particularly the rumoured hack of the largest exchange Mt Gox in February 2014, when the exchange lost bitcoin to the value of 409 million US dollars and went bankrupt (Forbes, 2014).

New bitcoins can only be created through a process known as 'mining'. Miners run a dedicated piece of software to try to solve a puzzle. When a puzzle is solved, a new block is added to the block chain. All miners are notified that a new block has been found and the process starts over trying to solve a new puzzle to add another block to the chain. Miners typically use dedicated hardware (in the form of specially designed integrated circuits) to solve the puzzles. The difficulty of each puzzle increases as the number of miners (or mining power) on the network increases, the difficulty factor of the puzzle is calculated every 2016 blocks and is based upon the time taken to generate the previous 2016 blocks. This keeps production at a steady rate and currently one block is mined roughly every 10 minutes. In addition, the size of each block reward given to the miner that discovers it is halved every 210,000 blocks - first from 50 bitcoins to 25 (as of November 2012 it is now 25 bitcoins reward), then from 25 to 12.5, and so on. Bitcoin is designed to be finite, with a limit of 21 million bitcoins, this is expected to be reached by the year 2140. In this way bitcoin is more similar to gold than a fiat[1] currency where a government can decide to print new money, as recently occurred in the rounds of

[1] fiat currency is being used in this context as a government backed currency not linked to a commodity such as gold, as all of the main currencies such as the US dollar are.

quantitative easing undertaken by the central banks of Japan, US and UK in response to the recession brought about by the financial crisis.

Although the technical workings of Bitcoin are complicated and beyond the scope of this paper, using it to actually purchase products is straightforward, once a supplier supports it as a payment method. It is becoming more commonly used and has been receiving widespread media coverage in the last number of years. More and more retailers are accepting it as payment. Virgin Galactic now accepts bitcoin as payment for their commercial space flights (Galactic, 2013). Expedia has recently become the largest online brand to accept payment in bitcoin. The currency has garnered much attention as a potential alternative to traditional fiat currencies. Forbes recently published a book detailing the efforts of their online editor to live for a week on bitcoin (Hill, 2014). Since its inception bitcoin has been associated with the purchase of illegal substances on sites such as Silk Road, an online marketplace operated as a Tor hidden service (sometimes called the eBay for drugs (Barratt, 2012)), primarily due to its anonymous nature. When the FBI closed the Silk Road site, the bitcoin exchange rate dropped dramatically, only to recover its price again in the weeks that followed. The currency has achieved much more widespread adoption in the last 2 years. Its use is growing with regular businesses now accepting it and with dedicated ATMs in place in a number of countries (BitcoinATMMap, 2014). There are also now a number of hedge funds that trade in bitcoin with new funds appearing all the time (Newsweek, 2014).

1.3 Research Question

This paper asks the research question (RQ):

(RQ1) Can the sentiment on Twitter predict bitcoin exchange rate?

Sub questions that are relevant within this research are:

(RQ2) Does the volume of Twitter messages relate to bitcoin price movement?

(RQ3) Does sentiment merely reflect bitcoin price movements or cause them?

(RQ4) Are retweets a better gauge of sentiment and are they more closely linked to bitcoin price changes?

1.4 Research Scope

This work focuses exclusively on Twitter. Twitter is a microblogging platform that allows users to post their thoughts and opinions to a public forum in the form of 140 character messages known as ‘tweets’. These tweets are publicly accessible and can be searched for or followed in real time. Twitter has 255 million monthly users, over 500 million tweets are sent a day (Twitter, 2014a). The Twitter platform has been shown to offer unique insight into consumer opinion and sentiment (Pak and Paroubek, 2010). The open and honest nature of the users’ messages, or ‘tweets’, offers an immediate view on their opinions, likes and dislikes. Consumer sentiment, either on an individual basis or aggregated across a user group, can be extracted from these tweets using specific tools and techniques. This information has been shown to be as accurate as traditional models of capturing user sentiment such as surveys. One such study has shown the use of user tweets to predict election results (Tumasjan et al., 2010). As well as offering a forum for expressing opinions, many users use Twitter to keep track of information or to follow other users. Up to 40% of users merely follow others (News, 2013). Users can also ‘retweet’, which is essentially forwarding someone else’s message to their followers. This results in data being disseminated very quickly across the twitter network. In this way Twitter has become similar to a news network or instant bulletin board, with research showing that 85% of the topics that are trending on Twitter are related to current news events (Kwak et al., 2010). Recent events such as the Arab Spring have illustrated the wide reach of Twitter and its importance in spreading information and shaping popular opinion. Several studies have shown the prominent role of Twitter in the Arab Spring (Howard et al., 2011, Khondker, 2011, Lotan et al., 2011, Huang, 2011).

1.4.1 Why bitcoin and not some other Forex?

The global foreign exchange trading market (or Forex) is not a market that receives exposure outside of financial institutions. The market for currency trading is enormous and dwarfs all other financial markets, for example the stock exchange. The foreign exchange market is on average $5.3 trillion worth of trades a day (GRAHAM, 2014). The transactions are between banks and have a low profit margin but, given the size of the market, offer an enormous reward. Several banks in Switzerland, the UK and the US are currently under investigation for the illegal fixing of exchange rates. As this market is essentially controlled by large institutions, there is little to be gained by analysing publicly available sentiment in relation to established currencies.

Since its inception, and particularly since it has seen a large increase in value, bitcoin is often viewed as a speculative investment and is actively traded (Yermack, 2013) Bitcoin was selected for this research as it offers the potential for a more democratic trading platform. Its users are actively engaged with its success and hence are more likely to publicly state their opinions and share information on a service like Twitter. Twitter can be seen as being analogous to the Bloomberg terminals in this context. Whereas the Bloomberg terminals are used by traders to get the latest financial information and to exchange information with other traders for a price that is prohibitive for most users, Twitter can be used for free. Bitcoin users and traders can express their opinions and feelings on the currency on a public platform. Bitcoin users by definition will tend to be technology savvy and hence are more likely to be active users of Twitter. These users could be either active tweeters or users that simply follow the topic to view other users' tweets on the subject. As stated previously, Twitter is often used to follow news events, and bitcoin users can use Twitter to keep up to date with the latest bitcoin news and exchange rates. This information is regularly tweeted from the official Twitter accounts for the various exchange platforms.

Another reason for selecting the bitcoin exchange rate is that it is difficult to assign a fundamental value to it (Gomez et al., 2014), its value is subjective and should be more prone to the influence of sentiment on its investors[2] (support for this statement will be shown in the literature review). Thus sentiment should correlate to price movements.

1.5 Importance of this Research and Beneficiaries

When it comes to financial markets, there are distinct advantages in harnessing this publicly available data over a traditional method like an investor survey. Firstly, the scale is well beyond what can be done through traditional methods, and secondly, the data can be captured in near real time. In the modern financial market this second factor is crucial. The Purchasing Managers Index takes weeks to collect; by the time the survey results are available the data may be stale or rendered irrelevant by socio-political changes. Given the real time nature of Twitter, it offers the ideal source of public data. Companies like StockTwits.com have formed by providing this information in a convenient manner, and Twitter introduced the concept of 'cashtags' (for example $APPL) to allow users to specifically track stock symbols they are interested in.

This research will be of benefit to both those interested in the field of sentiment analysis of online data and those with an interest in the bitcoin digital currency. This paper builds on

[2] in this context investors can be seen as users of the currency, as they have invested in its future by purchasing it

many research activities in recent years that show that sentiment can be used as a predictor for financial markets.

1.6 Guide to Dissertation

The structure of this dissertation is divided into the following chapters.

Chapter 1: Introduction – This chapter outlines the context, rationale and background to the research question.

Chapter 2: Literature Review – This chapter reviews the history of sentiment research with financial markets, moving to later day sentiment analysis of online data. The literature review shows why the research question was selected.

Chapter 3: Methodology and Fieldwork – This chapter explores the methodologies considered for this research and the reason for choosing the selected methodology. Details are given of how the research was carried out, the data collected and analysed.

Chapter 4: Findings and Analysis – This chapter states the findings of the research and analyses and reflects on these findings.

Chapter 5: Conclusions and Future Work – This chapter will show if the research has answered the research query, found any new or interesting results, and indicate any possible future research in that could come from this work.

2 Literature Review

2.1 Introduction

The overall aim of this research is to investigate whether the exchange rate of bitcoin will reflect the prevailing sentiment related to the digital currency. To explore this it is important to establish whether research has shown that investors are affected by sentiment in relation to more traditional investments like stocks. It is then necessary to review the methods that have been used to measure sentiment and if any are applicable to this study. It is also pertinent to examine the characteristics of investments more prone to sentiment, and if those same characteristics apply to bitcoin. Also, as this paper uses publicly available data from the internet, it is important to establish the reliability of that source, where it has been used previously, particularly in relation to financial markets. There follows a literature review of existing works in this field.

2.2 How Sentiment Relates to Market Prices

In examining how sentiment can predict or affect real world events the financial markets are often used. They provide a price over time that can be used to compare with sentiment to see if there is a correlation between the two. There is also, of course, considerable financial reward in trying to predict what the financial markets will do. Financial markets should, according to efficient-market hypothesis (often abbreviated as EMH) (Fama, 1970), follow a pattern based on sound economic data and not something as intangible as sentiment.

The concept of investor sentiment can be traced back to Keynes. He used the term "animal spirits" to describe the force that takes over the market, "*a spontaneous urge to action rather than inaction*" (1936, pp. 161-162). The irrational takes over from the logical. The wild market swings seen throughout the last 100 years cannot be attributed to the rational market forces, where, as Baker and Wurgler (2007, p. 3) put it, "*unemotional investors always force capital market prices to equal to the rational present value of expected future cash flow".* Events such as the boom of the 1920s that led to the Wall Street Crash of 1929, Black Monday in 1987, and the latest financial crisis when the Dow Jones Industrial Average lost 54% of its value from October 2007 through to March 2009, cannot be explained by the rational market behaviour predicted by EMH. Some have stated after the latest financial crisis that EMH should be abandoned as it discourages regulation in the belief that the market will look after itself and bubbles won't form, see Justin Fox (2011) and former Chairman of the Federal Reverse Paul Vockler (2011).

Interestingly in relation to this study the economist John Quiggin has written that the bitcoin bubble represents a clear refutation of EMH (Quiggin, 2013).

The bear or bull market runs have shown that the prevailing mood becomes contagious, driving the market higher or lower, defying what should be the rational price of the stock. Indeed the returns on stock have been a few percent higher than government bonds in the last century despite confounding what economists would predict based on arbitrage opportunities. Arbitrage can be defined as the practice of taking advantage of the difference in price between the same or similar securities in different markets for a profit. Arbitrage is a fundamental concept in finance which should bring prices to their fundamental value. It is the basis for the main argument against sentiment as a factor in price, which is that mispricing based on sentiment would be eliminated by rational traders seeking to exploit the profit opportunities created by non-fundamental prices. However what we see with stock returns being higher than government backed securities is that the magnitude of the risk premium (the return earned by a risky assets in excess of the return from a relatively riskless asset such as government bonds) is greater than would be expected by economic modelling. This has become known as the Mehra-Prescott equity premium puzzle (Mehra and Prescott, 1985). Sentiment has been proposed to explain this puzzle.

A model has been presented by De Long et al. (1990) and Sheifer (1997) based on 'noise traders' as defined by Kyle (1985) to help explain a number of financial anomalies, including the excess volatility of asset prices and the Mehra-Prescott equity premium puzzle. Their model is based on the assumption that investors are subject to sentiment and betting against a sentimental investor is risky. These noise traders can be more influential in setting the price than rational traders or arbitrageurs. Much of the work around investor sentiment and how it relates to price has been built on the work of Black (1986, p. 532) who contends – "*Noise trading is trading on noise as if it were information…. The more noise trading there is, the more liquid the markets will be, in the sense of having frequent trades that allow us to observe prices. But noise trading actually puts noise into the prices. The price of a stock reflects both the information that information traders trade on and the noise that noise traders trade on."*

The work of De Long et al. (1990) has demonstrated that this noise in the market will influence investor sentiment and that investors are subject to sentiment. The noise of Black can be viewed as sentiment and the noise traders as trading in sentiment as opposed to market fundamentals and facts. Shleifer and Vishny later expanded on this (1997) showing the limits of arbitrage where high volatility created by noise trader sentiment can deter arbitrage activity. Baker and Wurgler (2006) have built on this work to show the important role investor sentiment can play in setting market values. As does the model developed by Barberis et al. (1998) based on empirical evidence that predicts stock prices overreact to consistent patterns of good or bad news. This helps to explain the irrational or runaway behaviour of financial markets during a bull or bear market. Investor sentiment spreading and thus influencing market prices, the investor trading based on the sentiment and not the fact or fundamentals.

It has even been shown that external factors affecting the mood of investors as a whole can affect the market prices. One recent study linked a loss for a nation or team in a major sporting event such as a world cup match to a slump in the market the following day (Edmans et al., 2007). The collective mood of a nation reflected by the investors and traders and their depression reflected in the stock price. It seems that Keynes's animal spirits are at work. With investor sentiment being shown to be an important factor influencing market prices, the process of measuring sentiment becomes of great importance.

2.3 How to Measure Sentiment

Based on the knowledge that sentiment exists and affects markets, a key question is how to measure this sentiment or, more particularly in the case of financial markets, investor sentiment. This is of course a difficult task, and much of the existing work on measuring sentiment involves measuring proxies for sentiment. In the absence of a direct measure of investor sentiment, like a survey, the sentiment is inferred through a proxy. Baker and Wurgler (2007) provide a list of investor sentiment proxies that have been used previously by researchers: *investor surveys, investor mood proxies, retail investor trades, mutual fund flows, trading volume, dividend premia, closed-end fund discounts, option implied volatility, first-day returns on initial public offerings, volume of initial public offerings, new equity issues, and insider trading.* Of note is the fact that they have listed investor surveys as a proxy. The American Association of Individual Investors (AAII) example as used by Brown and Cliff (2004), is used a direct measure of investor sentiment, as discussed later.

However Baker and Wurgler selected the proxies from their earlier paper (2006) to do their analysis, those being: *the closed-end fund discount, NYSE share turnover, the number and average first-day returns on IPOs, the equity share in new issues, and the dividend premium*. As with the other sources of data listed previously, these are proxies through which sentiment can be inferred and measured, as example, high first-day IPO returns are used as a measure of positive investor sentiment.

Brown and Cliff (2004) use both direct data and proxy data. The direct data is in the form of a survey that directly measures the sentiment of market participants. This is a survey conducted by the American Association of Individual Investors (AAII). The survey asks each participant where they think the stock market will be in 6 months: up, down, or the same. AAII then labels these responses as bullish, bearish, or neutral, respectively. The second survey, Investors Intelligence (II), compiles another weekly bull-bear spread by categorizing approximately 150 market newsletters. They interpret the Investors Intelligence data as a proxy for institutional sentiment as many of the authors of these newsletters are current or retired market professionals and may not be directly reflecting the sentiment of the firm.

Both studies create a composite sentiment index grouping the proxy sentiment measures, as Baker and Wurgler (2007, p. 12) put it "*the practical approach is to combine several imperfect measures".* The approach although thorough seems somewhat unsatisfactory, useful for proving the theory of market sentiment affects prices but not useful as an approach for prediction. Using sentiment proxies is the primary method used by other researchers in how sentiment influences investors. Other prominent work which uses proxies include: Baker and Stein who use trading volume (2004), Lee et al. use the closed-end fund discount (1991), and Baker and Wurgler using equity issues as a fraction of total capital issuance (2000).

A more straightforward approach is used by Edelen et al. (2010) by looking at actual actions of institutional and retail investors in a historical context. However this approach would only work for past events and not as a predictor. For a predictive and simpler approach the work of Tetlock (2007) is of interest, he looked at the impact of the *Wall Street Journal*'s (*WSJ*) 'Abreast of the Market' column on U.S. stock market returns. He found that pessimism reflected downward market trends, and when pessimism was high or low trading volumes were higher, which tallies with other studies findings that sentiment affects trading volumes. This study also shows the importance of certain publications in shaping and setting opinion.

Tetlock's approach also uses a proxy for sentiment, the paper not being a direct source of investor sentiment but merely a bellwether for it. The study uses only one proxy and not a composite. It is also an example of how a media outlet which investors actively follow can shape sentiment. This paper will use a similar approach to Tetlock, it will use one source of data with Twitter, which, as seen, has similar characteristics to a news outlet in terms of disseminating news stories. Where this study differs from Tetlock is that the source of data can be seen as both a proxy, in the sense that is used to disseminate news related to

Bitcoin, and as a direct source of sentiment, in the sense that it should also directly reflect investors in bitcoin's opinion and mood. The source is also different in that it is a more immediate source of content, and focused solely on one investment by directly tracking bitcoin related tweets from Twitter.

Existing research into online sources of sentiment will be looked at shortly. First the historical findings of the main investor sentiment studies will be examined to assess whether results show the power of sentiment and if there are any key findings that can be applied to this study.

2.4 Empirical Evidence – Is Sentiment a Factor?

Looking in more detail at some of the key papers it can be seen that the effects of sentiment on stock prices have been shown time and again. Stock prices have been shown to overreact to patterns of good or bad news, good earnings announcements having a disproportionate effect on price (Barberis et al., 1998). Baker and Wurgler found that investor sentiment, broadly defined, has significant cross sectional effects. They found that "*When sentiment is estimated to be high, stocks that are attractive to optimists and speculators and at the same time unattractive to arbitrageurs—younger stocks, small stocks, unprofitable stocks, non-dividend paying stocks, high volatility stocks, extreme growth stocks, and distressed stocks—tend to earn relatively low subsequent returns. Conditional on low sentiment, however, these cross-sectional patterns attenuate or completely reverse*." (2006, p. 33)

Often in studies of sentiment the proof of sentiment's influence on price is if the stock or asset affected by the positive or negative sentiment returns to its fundamental value. The process involves tracking the correlation between positive sentiment and overvaluation and tracking the subsequent return to fundamentals. This is often used as it proves that it is sentiment, rather than a change in fundamentals, that is driving the price change in the first place. Tetlock (2007) noted that the price impact of pessimism appears especially large and slow to reverse itself in small stocks. Thus its impact is greater and seen for longer. Moreover that study linked stocks traded by individual investors (small stocks in this case) as those most susceptible to sentiment. This will be applicable to bitcoin as although bitcoin funds and investment products are emerging it is certainly not a traditional investment. Edelen et al. (2010) have shown that fluctuations in relative retail sentiment are positively associated with contemporaneous stock market returns and negatively associated with future stock market returns. This pattern is consistent with the hypothesis that retail sentiment is more variable than institutional sentiment and retail investors move prices as they update their asset allocations to reflect their shifting sentiment. Again as bitcoin is currently traded by individual investors more than institutional ones this is also a relevant finding for this study.

More latterly Baker and Wurgler (2007) examined the empirical effects of sentiment. They show that it is possible to measure investor sentiment, and that waves of sentiment have clearly discernible, important, and regular effects on individual firms and on the stock market as a whole. In particular they find that stocks that are difficult to arbitrage or to value are most affected by sentiment, a common finding across the research. Figure 2.1 neatly illustrates that point.

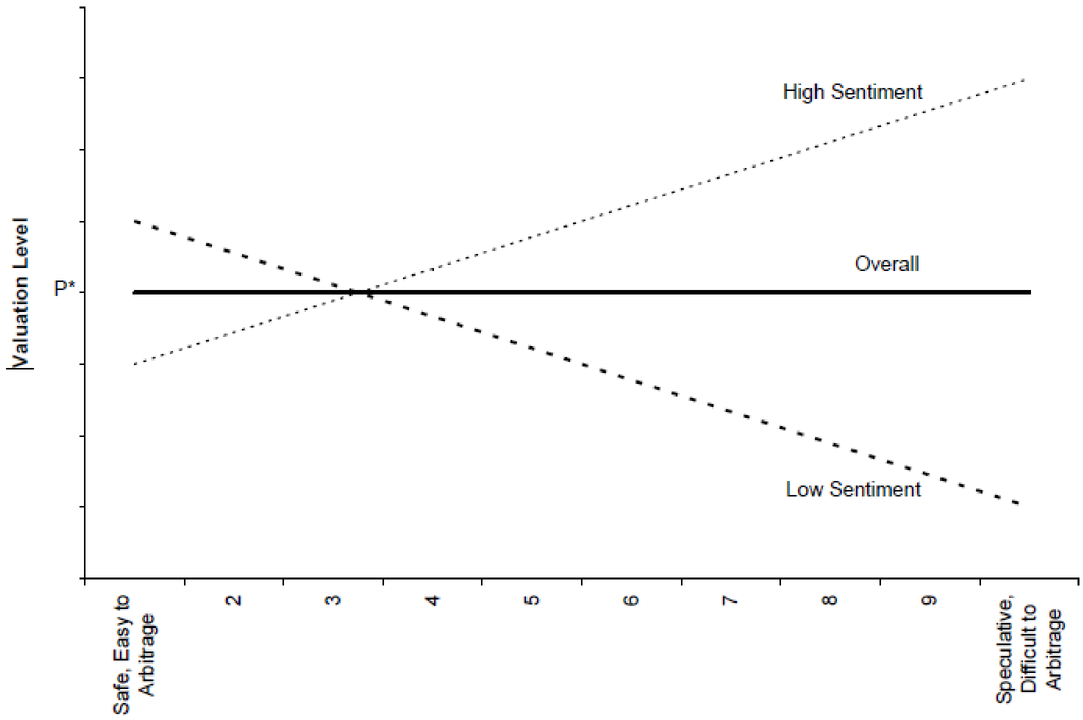


FIGURE 2.1 *(Baker and Wurgler, 2007) Cross-sectional effects of investor sentiment. Stocks that are speculative and difficult to value and arbitrage will have higher relative valuations when sentiment is high.*

There are a number of common findings that are pertinent to this study. One, the effects of sentiment have a greater impact on stocks that are difficult to put a fundamental value on or are volatile. Two, investments that are difficult to arbitrage are more prone to the effects of sentiment. Three, sentiment has a greater impact on stock that are more likely to be traded by individual investors rather than institutional investors, this can be due to a number of factor such as like the stocks being young, highly volatile, distressed etc.

Now that it has been shown that sentiment influences investors and that it can be measured and used to predict market returns, the next section will assess a more recent source of sentiment. Information available on the internet, in particular the data available on social network sites.

2.5 Using Online Data

The link between sentiment and market price has been established by others, but the imperfect proxies used to measure sentiment are unsuitable for this study. Looking beyond the imperfect proxies listed earlier to other potential sources of sentiment data one quickly turns to the publicly available data online. The internet offers researchers new possibilities in data collection. Granello and Wheaton (2004) have documented some of the benefits in using online surveys - these include reduced response time, lower cost, ease of data entry, flexibility of and control over format, advances in technology, recipient acceptance of the format, and the ability to obtain additional response-set information. As well as methods of collecting data the internet also offers huge publicly accessible data pools that researchers can use. The internet opens exceptional possibilities for researchers in both increasing the amount of information available and in lowering the cost of collecting this data (Edelman, 2012).

There are also services that allow researchers easy access to this data. For example Google Trends (Trends) provides reports on frequency of google searches. There have been a number of studies that have used this data, Choi and Varian (2012) used the data to predict a number of economic indicators including automobile sales, unemployment claims, travel destination planning and consumer confidence. Wu and Brynjolfsson (2013) showed that the search data can be used as a predictor of the housing market, showing that prior to the housing collapse in Florida searches related to real estate plummeted. There have been a number of studies that used search data to detect epidemics and disease (Ginsberg et al., 2008, Pelat et al., 2009, Seifter et al., 2010). The data provided by Google Trends is easily accessed and can provide a quick insight on a topic, as example see Appendix A for a comparison of the bitcoin search results on google and the historical exchange rate. As can be seen there is a clear correlation. A study that uses this approach for bitcoin will be reviewed in section 2.7

As well as a source of raw data, the internet offers a vast well of information to mine for consumer sentiment and opinion. The increase in internet users and users of social networking sites, blogging and microblogging platforms has opened up a huge data pool to collect and analyse. This has led to much research in recent years, as Bing Liu (2012, p. iv) states, "*For the first time in human history, we now have a huge volume of opinionated data recorded in digital form for analysis".*

Harvesting customer sentiment and opinion is becoming a vital tool for companies looking to understand their consumers and tailor products to them. With the advent of 'big data'[3] companies are using new tools and techniques gain new insight into their customers' likes and dislikes. Moreover, from a consumer perspective, the opinions of others (and what is new here is that the 'others' are complete strangers) have become increasingly important. Customer reviews and ratings have become common place on the websites of retailers and have been shown by a number of studies to influence potential purchases (Zhu and Zhang, 2010, Gretzel and Yoo, 2008)..

Pang and Lee (2008) present a comprehensive overview of the topic and related work in sentiment analysis and opinion mining and latterly Liu (2012) presented the latest developments and papers on the topic. The common approach from most of the research is to use machine learning techniques to automatically perform the classification. Deriving overall sentiment from a piece of text is a difficult problem to solve. It is easier to classify text into categories (such as sports related, politics related etc). One of the reasons it is so difficult is to derive sentiment from text is that human communication can be difficult to understand. Pang et al. (2002, p. 7) noted the problem in relation to movie reviews, they noted what they describe as the "thwarted expectation" in reviews, one example they gave was –

"This film should be brilliant. It sounds like a great plot, the actors are first grade, and the supporting cast is good as well, and Stallone is attempting to deliver a good performance. However, it can't hold up".

Examples such as this and sarcastic language present a problem for machine learning tools. Though it's easy for a human to interpret the sentiment. Most machine learning approaches for classification use training data to learn how to interpret sentiment. This involves the researcher manually classifying training data which can be time consuming.

Notwithstanding those problems research has continued with great success. Other sentiment analysis of online systems include the work of Liu el al. (2007), in which a sentiment model was proposed to predict sales performance. Hong and Skiena (2010) studied the relationship betting and public opinion in blogs and Twitter in the NFL. Similarly Sinha et al. (2013) looked at NFL tweets as a means to predict future match

---

[3] The term 'big data' is all encompassing term that normally refers to the 3 Vs.
Volume – bigger that can be processed and analysed efficiently with traditional approaches
Velocity – Data streaming in real time from online or through
Variety – structured (in existing databases) and unstructured data from social media, email etc

results. Predicting box office returns based on the sentiment of Twitter and other Social Media sites has been researched a number of times: Asur and Huberman (2010), Sadikov et al. (2009) and Mishne and Glance (2006) to name a few. Twitter is a very common data source for user sentiment based research, and will be looked it in more detail in this paper. A common theme in the research is to use a time-series variable with which to measure and compare the sentiment analysis against. Opinion Polls, box-office taking, and sales of a product all offer a useful real life comparison. Of course so do the financial markets, as will be examined next.

2.6 Public Sentiment and Trading

As shown earlier Tetlock (2007) showed the interactions between media and the stock market. They showed how the Wall Street Journal can act as a proxy for investor sentiment and how it can influence prices. A number of studies have looked at the sentiment of online data and how it relates to stocks. This is moving closer to the core of this study.

Antweiler and Frank (2004) performed a study of online posts to Yahoo finance and Raging Bull message boards. They studied 1.5 million messages posted on these platforms about the 45 companies in the Dow Jones Industrial Average and the Dow Jones Internet Index. Their study is analogous to this research as they used machine learning techniques and the training set and data tested was of similar volumes (1000 messages were manually classified). They aggregate sentiment over multiple time periods, 15 minutes, 1 hour and 1 day, in order to test the sentiment ('bullishness' is the term used). A similar approach will be adopted for this research, however the aggregation of sentiment is over longer time periods. They tested two algorithms, Naïve Bayes and Support Vector Machine, and had similar findings for both so only reported on the Naïve Bayes results. Naïve Bayes is one of the oldest and most widely used algorithms and is the one selected for this paper. They found find that stock messages help predict market volatility. That their effect on stock returns is statistically significant but economically small, consistent with previous findings in the field. That paper also introduced a measure of bullishness that will be used to test the results in Chapter 4.

A similar study was performed by Das et al. (2005) using message boards. They measured the intensity and dispersion of sentiment for over 170,000 messages posted about four stocks. They found that there is a close relationship between sentiment levels, stock prices, and trading volume. They explore the usefulness of expressed investor sentiment to predict stock returns. Their study failed to find a predictive link, the message board sentiment reflects the sentiment but does not influence the price.

In a later paper by Das and Chen (2007) which again traces the sentiment of message boards, they found that technology sector's aggregate sentiment can predict the level of the sector's aggregate index but not of individual stocks. They used several different machine learning algorithms to classify each message. Gu et al. (2006) present a study on the predictive power of message board sentiments over abnormal stock returns. Their findings show a link between sentiment and trading, which they say are consistent with psychological theories. They suggest investors overreact to news, and those who happen to predict correctly in the past are more likely to overreact. They devised a trading strategy that involves buying stocks with low sentiments while selling stocks with high sentiments was implemented. The results indicated weekly returns ranging from 0.44% to 0.66%.

Sabherwal et al. (2011) performed a study of sentiment related to small firms with weak financials. They found that a two day pump and dump strategy existed among online traders, suggesting that message boards can be used to temporarily drive up prices. This is an important finding for this study, financial message boards can be viewed as analogous to modern day Twitter conversations about a stock or asset such as bitcoin. They conclude that message board sentiment is an important predictor of trading-related activities. Their work tallies with findings in financial research that say stocks prices for volatile, small firms or ones that are difficult to value are more subject to the effects of sentiment. This is a useful finding for this study as will be elaborated on in the final section of the literature review.

Moving from the older stock message boards to social networking sites a connection can be drawn to Twitter. Many of the social networking sites that we have become used to for sharing information are similar to message boards or messaging on Bloomberg terminals used by traders. The Twitter related service StockTwits.com can be seen as a challenger to Bloomberg terminals (Bloomberg, 2014b), and are likely to be used by individual investors (for whom the 20,000 dollar a year price tag of Bloomberg subscription might be too much). Twitter's place in the field of research will now be examined, looking particularly at the studies that have linked Twitter sentiment to stock trades.

2.7 Twitter and Trades

Millions of users share their opinions and thoughts on Twitter on a daily basis. Consumers increasingly use these communication technologies for trusted sources of information and opinions (Jansen et al., 2009). The messages are limited to 140 characters in length and hence tend to be concise and to the point. The Twitter API allows researchers to mine the data for a particular topic, thus getting a focused view of Twitter users that are actively engaged in this topic.

The limited character sets and extensive use of Emoticons – graphic representations of facial expressions common in emails and text messages (the smiley or sad faces) – makes Twitter an excellent and easy source to harness consumer sentiment. The use of smiley or sad faces in a tweet can allow for the categorisation of tweets according to emoticon used, and avoid the manual and troublesome effort of categorising the training and test data as mentioned earlier. Several studies have used this approach and shown that emoticons increase the success rate in classifying text based data. Go et al. (2009) and Davidov et al. (2010) and have shown the use of emoticons for automatically building a sentiment corpus avoiding the manual process of classifying data. Pak and Paroubek *(2010)* have done the same thing, in their work they showed that the Naïve Bayes classifier worked best for analysing tweets.

There has been much research using Twitter as a barometer of public opinion. In relation to political matters, O'Connor et al. (2010) compared the sentiment of Twitter messages with opinion polls from America. Tumasjan et al. (2010) carried out similar research in the lead up to the German Federal elections. They compared Twitter sentiment with opinion polls and found that Twitter sentiment can be used when predicting elections. The latter used a simplistic approach to text analysis but still showed that the number for tweets related to a particular party reflected the election results.

Naturally the financial sector offers a rich area to compare social media sentiment with real life market trends. Indeed a number of such studies have appeared in recent years. Vincent and Armstrong (2010) assess high-frequency trading strategies grounded in messages on Twitter, finding a profit opportunity in fast-breaking Twitter discussions. Bollen et al. (2011) used Twitter moods to predict the stock market. Using large scale Twitter feeds they found a correlation between changes in the public mood and that shifts in the Dow Jones Industrial Average (DJIA) values that occur 3–4 days later. Oh and Sheng (2011) showed that Twitter can predict future stock price moves. Their study showed that stock micro blog sentiment do have predictive power for simple and market-adjusted returns. Their study used StockTwits.com and Yahoo Finance as sources. Promisingly for this research they find that *"irrational investor conversations and such distinct features of microblogging as succinctness, high volume and real-time contribute to the predictive value of micro blog sentiments*." (2011, p. 13).

Sprenger et al. (2013) also had similar findings. Finding that tweets were a valuable proxy of investor behaviour and belief formation. They also performed an analysis of Twitter message volumes and trading volumes, finding that messages volumes predict trading volumes one to two days later. This study will perform a similar analysis to answer research question (RQ2). Recently Sul et al. (2014) analysed Twitter messages related to stocks in the S&P and rated their sentiment. Their results show that the cumulative emotional valence (positive or negative) of Twitter tweets about a specific firm was significantly related to that firm's stock returns.

2.8 Bitcoin as an Investment affected by Sentiment

This paper looks at using the sentiment of tweets as a way to measure the exchange rate of the bitcoin digital currency. Bitcoin price is not connected to the performance of a country or socio-political changes as other currencies are. Bitcoin is not traded by large institutions in the same way that other foreign exchange is. One of the key findings in the research into sentiment effects on market prices is that their influence is most felt for stocks or assets that are difficult to put a fundamental value on, are volatile, or are difficult to arbitrage. As Baker and Wurgler (2006) found, some firms are more likely to be disproportionately sensitive to broad waves of sentiment. The characteristics they defined are: stocks with low market capitalisation, young, unprofitable, highly volatile, non-dividend paying, growth companies or stocks of firms in distress. Considering bitcoin as an asset rather than a stock some of these characteristics apply to it: young, highly volatile, low market cap and a growing asset.

In a later paper Baker and Wugler describe what makes stocks more speculative than others (2007, p. 7): "*the crucial characteristic is the difficulty and subjectivity of determining their true values. For instance, in the case of a young, currently unprofitable but potentially extremely profitable growth firm, the combination of no earnings history and a highly uncertain future allows investors to defend valuations ranging from much too low to much too high, as befits their prevailing sentiment.*" This statement can certainly be applied to bitcoin.

During the initial research for this paper to check content on Twitter related to bitcoin, the following two tweets were repeatedly retweeted:

> *Winklevoss twins: bitcoin could hit market cap of $400bn*
>
> *#bitcoin Tulipmania of our times*

This clearly shows the different extremes of the valuations amongst Twitter users in relation to bitcoin. Baker and Wurgler (2006) predict that investor sentiment has larger effects on securities whose valuations are highly subjective and difficult to arbitrage. Another study with similar findings from Ali et al. (2003), show that the book-to market effect is strongest in stocks that are difficult to arbitrage, which is consistent with the effect arising from mispricing rather than missing risk factors. As Shleifer and Vishny note (1997), professional arbitrageurs may try to avoid extremely volatile arbitrage positions, although potentially very rewarding they run the risk of big losses should they need to liquidate quickly for a client. This is applicable to bitcoin as bitcoin exchanges are volatile and arbitrage opportunities have not yet emerged (Gandal and Halaburda, 2014).

It is difficult to apply a fundamental value to bitcoin, it is very young in the context of other currencies and younger still when compared to finite commodities such as gold, silver or platinum. Since it first launched its price has been highly volatile, the price has fluctuated wildly in the last number of years, reaching a valuation of over 1,000 dollars for one bitcoin (the current value is roughly 500 dollars), although the price fluctuations have settled down since the beginning of 2014. For a comparison of fluctuation in price since launch please see Appendix B.

Bitcoin can also be said to be difficult to arbitrage for the reasons listed above, although there are reasons why it should provide arbitrage opportunities. It is traded on multiple exchanges at different rates. An investor could trade on the differences between these markets, which is classic arbitrage. Although based on the instability of some of the markets, this would still be a risky endeavour, an arbitrageur could see their investment disappear. Moore and Christin (2013) have presented work that tries to quantify the risk of using certain exchanges over others. Of interest is a company called Bitcoins Reserve (Reserve, 2014) that recently formed, claiming to trade on arbitrage opportunities available between the different market places. As they state on their website: "*one such investment vehicle is our Arbitrage fund, which performs automated simultaneous trades across multiple exchanges with price differentials, to correct market inefficiencies and bring liquidity, all in the while netting profitable trades*".

Should more such companies appear the price of bitcoin should start to stabilise. However as things currently stand bitcoin does not offer arbitrage opportunities, a working paper from the Bank of Canada (Gandal and Halaburda, 2014) provided a comprehensive analysis of different bitcoin exchanges over several months and found that there were little if any arbitrage opportunities between bitcoin exchanges, and what little opportunities there were have dissipated.

As shown bitcoin would seem to have many of the characteristics that would make it prone to the effects of sentiment. Kristoufek (2013) linked the price of bitcoin to the Google Trends and visits to Wikipedia. They analysed the dynamic relationship between the bitcoin price and the interest in the currency measured by search queries on Google Trends and frequency of visits on the Wikipedia page on bitcoin. They found a strong correlation between price level of the digital currency and both the Internet engines, they also find a strong causal relationships between the prices and searched terms. Of note, they found that this relationship is bidirectional, i.e. not only do the search queries influence the prices but also the prices influence the search queries. They found that while the prices are high (above trend), the increasing interest pushes the prices further up. From the opposite side, if the prices are below their trend, the growing interest pushes the prices even deeper. They pointed to the fact that bitcoin is interesting to study from a bubble-burst perspective. They believe that their paper will serve as a starting point for research into the statistical properties, dynamics and bubble-burst behaviour of the digital currencies as these provide a unique environment for studying a purely speculative financial market.

The results of that paper are promising for this research. However the results crossed over a time of great volatility for the currency, when it first entered the public consciousness and saw enormous gains in its price followed by a rapid deprecation. A high level view of the swings in the currency would have been easier to predict through search alone. As mentioned earlier and shown in Appendix A, a coarse view also shows a correlation between the price of bitcoin and searches related to it. In another study Glaser et al. (2014) that found that bitcoin price volatility is significantly influenced by media coverage and positive sentiment.

The work presented here differs in that it occurred over a period of relative stability for the bitcoin currency compared to what has gone before. Whether the same evidence will exist as bitcoin becomes more mature and the price stabilises remains to be seen. Although whether or not the price will remain stable for long is open to debate. A recent poll conducted by Bloomberg (2014a) showed that a majority of investors felt that bitcoin was overvalued. The results of that poll are interesting in themselves. The surveyed 562 investors who are Bloomberg subscribers: 55 percent of those surveyed said the virtual currency trades at unsustainable, bubble-like prices. 14 percent said it’s on the verge of a bubble. 6 percent of respondents said a bubble isn’t forming. The remaining 25 percent were unsure. The lack of a clear consensus seems to reinforce the point of the difficulty in setting a fundamental value for bitcoin. Though Bloomberg themselves must have some confidence in the digital currency as they recently started providing bitcoin pricing to their subscribers.

Having no clear fundamental value does throw up a problem for a researcher trying to prove that it is sentiment that is causing the market swings. One of the key approaches used in research into sentiment is to track the value from overpriced back to its fundamental level. Thus proving that it is sentiment, rather than a change in fundamentals, that is driving the price change. Of course this may not be possible where the fundamental value is not well known as is the case with bitcoin.

2.9 Conclusion

In summary – previous research has shown that sentiment is a real factor in influencing investors and thus setting prices. There has also been a clear link found between stocks or assets that are difficult to arbitrage or without a fundamental value and the influence of the effects of sentiment. The act of measuring sentiment online has been demonstrated and how these techniques are being used to measure sentiment related to financial markets. As a source of data Twitter has been shown to be an excellent source of consumer sentiment and a disseminator of news. Therefore, based on this knowledge bitcoin should provide an excellent investment to analyse for this study, and Twitter the perfect mechanism to monitor sentiment.

3 Methodology and Fieldwork

3.1 Introduction

Research methodology refers to the various steps a researcher uses in order to answer a question or address a problem with a particular objective in mind. The research methodology used in this paper can be traced through the layers of the 'Research Onion' as defined by Saunders et al (2012). The concept of a Research Onion encourages a researcher to resist the temptation to chase the data to answer a particular research question, instead it encourages the researcher to step through the layers to build a systematic approach to their research. The 'Research Onion' graphic is shown in Figure 3.1.

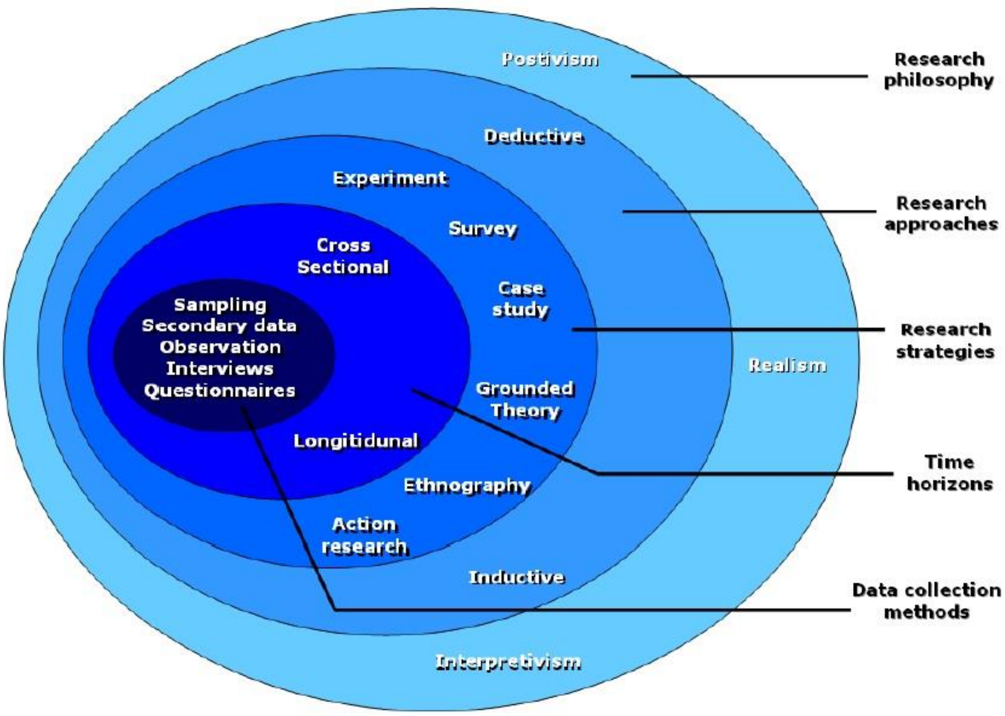


FIGURE 3.1 Research Onion (Saunders, 2012)

The main layers in the research onion will be discussed now and how they relate to this research. The main layers are: research philosophy, research approaches, strategy, choices, time horizon, and techniques and methods of data collection.

3.2 Research Philosophy

A research philosophy is a belief or an idea regarding the collection, interpretation, and analysis of data collected. There are various philosophies explained in Saunders' research onion. The most significant among them are: Positivism, Realism, Interpretative and Pragmatism. The philosophies are outlined in Table 3.1.

TABLE 3.1 Comparison of four research philosophies (Saunders, 2012)

	Positivism	**Realism**	**Interpretivism**	**Pragmatism**
Ontology: *the researcher's view of the nature of reality or being*	External, objective and independent of social actors	Is objective. Exists independently of human thoughts and beliefs or knowledge of their existence (realist), but is interpreted through social conditioning	Socially constructed, subjective, may change, multiple	External, multiple, view chosen to best enable answering of research question
Epistemology: *the researcher's view regarding what constitutes acceptable knowledge*	Only observable phenomena can provide credible data, facts. Focus on causality and law like generalisations, reducing phenomena to simplest elements	Observable phenomena provide credible data, facts. Insufficient data means inaccuracies in sensations (direct realism).	Subjective meanings and social phenomena. Focus upon the details of situation, a reality behind these details, subjective meanings motivating actions	Focus on practical applied research, integrating different perspectives to help interpret the data
Axiology: *the researcher's view of the role of values in research*	Research is undertaken in a value-free way, the researcher is independent of the data and maintains an objective stance	Research is value laden; the researcher is biased by world views, cultural experiences and upbringing. These will impact on the research	Research is value bound, the researcher is part of what is being researched, cannot be separated and so will be subjective	Values play a large role in interpreting results, the researcher adopting both objective and subjective points

The philosophy adopted for this study is positivism. Positivism is grounded in the theoretical belief that there is an objective reality that can be known to the researcher, by applying the correct methods in a correct manner. The research is external and objective of that being observed. The research aims to answer a specific research question using quantitative data, it is highly structured and uses a large sample (over 500,000). The results are applicable to others.

3.3 Research Approach

The second last layer of the research onion is the research approaches of which there are two described by Saunders: Deductive and Inductive

Deductive Approach: This comes from scientific principles. In general it is the journey from a theory to data results. A characteristic of the deductive approach is it seeks to explain causal relationships between variables. The researcher will be separate from that they are researching.

Inductive Approach: This approach is used if a clearly defined theoretical framework is not used. It typically involves collecting data, identifying relationships and patterns, and developing questions and hypotheses or propositions to test these patterns. The theory emerges from the process of data collection and analysis. The inductive approach may involve a lengthy period of time and prove to be resource intensive. Often used with elements of a deductive approach to develop a theoretical position and then test its applicability through subsequent data collection and analysis.

This study uses the deductive method, the data is collected with a specific research question and approach in mind. This is more suitable for a study of this nature as the study is limited by time. Quantitative data will be generated and analysed to seek to prove whether the research question is true or false. The inductive approach could be suitable for other research using a social networking site such as Twitter, as the volume of data may reveal interesting patterns leading to research questions. However, as this study is time limited, the deductive method is used.

3.4 Research Strategy

The next important layer in the research onion is research strategy. There are various strategies that researchers adopt for a particular research study. In Saunders' research onion various research strategies are explained. The main strategies are: experiment, survey, action research, case study, grounded theory, ethnography and archival research.

For this study Experimental research is the only research strategy suitable. Had there be an existing source of data then archival research would also have been a possibility. As stated by Saunders (2012, p. 173) "*The simplest experiments are concerned with whether there is a link between two variables. More complex experiments also consider the size of the change and the relative importance of two or more independent variables.*"

A link between two variables is precisely what this study is trying to establish (sentiment and exchange rate). In order to do so, a machine learning experiment needs to be conducted. This paper uses experimental research to answer a specific question.

3.5 Research Choices

The next layer in the research onion is Choice. The choice types are: Mono Method, Mixed Method and Multi method refer to the data collection techniques. Which often go with corresponding data analysis procedures, whether they are qualitative or quantitative. As Saunders state (2012, p. 182) "*In choosing your research methods you will therefore either use a single data collection technique and corresponding analysis procedures (mono method) or use more than one data collection technique and analysis procedures to answer your research question (multiple methods)*".

This paper uses the Mono Method. All the data is collected in the same way.

3.6 Research Time Horizons

Time Horizons refer to the time limit which is imposed on the research. There are two types of time horizons, longitudinal and cross sectional. In the longitudinal study the researcher observes the phenomena for an extended period of time, whereas in a cross-sectional study the time is limited or fixed.

As the time frame for this research is limited, and historical tweets are not available, a cross-sectional time horizon will be used.

3.7 Research Data Collection and analysis

The most important elements in a research study are data collection and data analysis. Data collected and analysed in a systematic manner will allow a research question to be answered. Two types of data can be collected for a systematic analysis for any research: Primary Data and Secondary Data.

Primary Data

Primary Data refers to that information that is generated for the first time, or that is generated to meet the specific requirements of the investigation at hand. Primary data is collected directly from the respondents or the subjects of experiment. A major drawback of using primary data is the fact that it can be time consuming to collect, and it can be difficult to obtain large amounts of data. Examples of sources of primary data include: surveys, questionnaires, interview schedules and interviews, focus groups, case studies, experiments and observations.

Secondary Data

Secondary data is not collected directly from the respondents. Instead, the data has been collected by others. The collection of secondary data can be faster to complete, and it can be easier to obtain large amounts of data. For data comparison between large existing datasets, secondary data can be very effective. Yet the secondary data can be outdated and subjective as it has already evolved in the mind of somebody else.

There are various sources of secondary data: journals, newspapers, books, articles in magazines and websites, government statistics, company or organisation statistics or more latterly the internet.

In this study secondary data from the internet in the form of tweets from Twitter are used, secondary data in relation to the latest bitcoin exchange rate is also used, as provided by a third party website Coindesk (2014).

3.8 Population & Samples

A research population is the total number of individuals or objects that are the main focus of this study. The population in this study are all Twitter users that tweet about bitcoin. A sample is a smaller representation of the population from which it is taken. It is a subset of the population selected in such a way that they are the representative.

The sample size used in this study is all available tweets on a subject over a 3 week period, circa 700,000 tweets[4]. It is a sample in the sense that it is limited in time, as Twitter does not allow access to historical tweets via the Streaming API. Therefore there may be multiple users that are not engaged with Twitter during the period of the study. As the sample proportion of the whole is not known this is called Non-Probability sampling.

[4] The number of Bitcoin related tweets per month was benchmarked at 180,000 per month at the beginning of the project. That number has now risen to over 900,000 per month.

3.9 Twitter Data Capture – Building the Model

As the volume of data collected was much too large for a human to classify, machine learning techniques were used to perform the classification. The first step is to define the variables with which the tweets will be categorised. In this case it is simply positive, negative or neutral. The process of analysing text and assigning it to a category is known as Classification in machine learning. Classification is a 3 step process.

1. Train/build a classification model
2. Test the model
3. Use the model in production

Classification is a type of supervised learning, which means training data needs to be provided to build the classifier. The first step in building a classification model is to capture data with which to train it. Two approaches were adopted to build a classification model. The first was to build a custom model using tweets specifically related to bitcoin, which were collected and manually classified for this research. The second was to use a publicly available set of tweets (a Twitter corpus), as provided by the work of Go, Bhayani et al. (2009).

3.9.1 Building a custom model

Twitter exposes an API for collecting tweets based on particular search criteria. This API was used to collect a total of 29,511 tweets, based on several separate runs each collecting roughly 10,000 bitcoin related tweets. The data was collected in two time frames, December 2013 and May 2014. In this time period there was much coverage of bitcoin in the media both positive and negative.

The selected tweets were filtered to remove non English tweets and duplicates. In Twitter duplicates would be accounted for by re-tweets, although this information will be useful for viewing sentiment on the production run, it is not useful for training data. A subset (756 tweets) of the most useful data was used for training and testing. The data was manually classified according to three target variables: positive, negative or neutral. Table 3.2 contains sample of the data used to train the classifier.

TABLE *3.2 Sample of Training Data*

Negative	Bitcoin burglar bags a million bucks
Negative	$5 million worth of bitcoin vanish in China
Neutral	Bitcoin Couple Travels the World Using Virtual Cash: It was a three-month odyssey that spanned the globe
Neutral	I bet some where some one is gonna buy a #PS4 for #bitcoin.
Positive	A gold platform admits a humble defeat and shutting down because bitcoin is the better choice for their customers.
Positive	Bitcoin Price Hits New Record High

Some tweets can be difficult to classify. For example, for the following tweet the sentiment is unclear:

9 Alternative Currencies That Are Even Crazier Than Bitcoin

Due to Twitter's limit character length it can often be the case that there is no context for a particular tweet. Crazy could be good or bad depending on the point of view of the tweeter. This particular tweet was marked as Neutral.

3.9.2 Model based on Twitter Corpus

The process of manually classifying data can be laborious and prone to errors, due to the subjective nature of human input. Another issue with Twitter is that given its length restrictions abbreviations and slang can often be used. Thus manually selecting a representative data set can be difficult[5]. An existing Twitter sentiment corpus whose accuracy has been tested can eliminate some of the issues with manual classification. One such corpus was produced by the work of Go, Bhayani et al. (2009), and is available to download at the website Sentiment140 (2014). They used Twitter emoticons to automatically categorise 1.6 million tweets. The presence of smiley or sad faces was taken as a signal of positive or negative sentiment. A similar approach was attempted with this paper but with a more targeted approach. Tweets with positive and negative emoticons and with the term 'bitcoin' were collected with a view to building a model of domain specific sentiment. However, after one week of continuous polling of Twitter API less than 20 tweets with emoticons were collected and the activity was abandoned. Though that activity was abandoned the 1.6 million tweet corpus was used to build a second model to test against the test data from the custom model.

[5] This is without mentioning the use of symbols like hashtags (#) denoting subjects, @ to indicate usernames of other twitter users, and retweets as symbolised by RT

3.10 Classifying Tweets

The Mahout Project from Apache Software Foundation was selected as the machine learning framework to use. Its ability to scale to multi-million pieces of information (tweets in this case) was seen as beneficial for any future work with Twitter.

3.10.1 Algorithm Selection

Some of the algorithms commonly used in classification include

- Naïve Bayes
- Complimentary Naïve Bayes
- Stochastic Gradient Descent (SGD)
- Support Vector Machine (SVM)
- Random Forests

As implemented by the Mahout Machine Learning Software the algorithms have the following characteristics summarised in Table 3.3.

TABLE 3.3 Summary of Machine Learning Algorithms in Mahout

Algorithm	Execution Model	Data Set Size	Characteristics
SGD	Sequential	Small to Medium <10million training samples	Efficient with smaller dataset
SVM	Sequential	Small to Medium <10million training samples	
Naïve Bayes	Parallel	Medium to Large Millions to hundreds of millions training sample	Good for text like data, useful for large datasets
Complimentary Naïve Bayes	Parallel	Medium to Large Millions to hundreds of millions training sample	
Random Forests	Parallel	Small to Medium <10million training samples	

When selecting a machine learning algorithm for use with Twitter messages Naïve Bayes has been shown to be very accurate, as stated earlier from Pak and Paroubek's work *(2010)*. For this reason Naïve Bayes was chosen as the algorithm to use for this research.

3.10.2 Naïve Bayes

The Naïve Bayes algorithm counts the number of times each word appears in a document in the class and divides that by the number of words appearing in that class. This is referred to as a conditional probability. In this case, the probability that a word will appear in a particular category. This can be written as P(Word | Category)

Naïve Bayes assumes that the occurrence of all words in a document are independent of each other. It treats the document as what is known as a 'bag of words' treating each word as independent from the other. Though the approach is simplistic, it is well proven technique that has shown to be effective when compared to more sophisticated algorithms (Pang et al., 2002).

3.10.3 Testing the Models

With the algorithm selected and the training data prepared, the models were built with Mahout. The process is iterative, i.e. testing of the model occurred after each attempt to improve accuracy. With each iteration new tweets were added to the training data. The testing process involved using 20% of the previously classified tweets held back from the training data (140 tweets) to verify the accuracy of the models. The custom model proved to be more accurate than the Twitter corpus model. The custom model had a score of 78% accuracy as opposed to 52% for the twitter corpus. The confusion matrix from these tests can be found in Appendix B. A confusion matrix displays the number of correct and incorrect predictions made by the model compared with the actual classifications in the test data. The matrix is n-by-n, where n is the number of classes. As the training and testing set was focused on bitcoin and market related terms, it is not surprising that the custom model performed better. As bitcoin matures and the tweets related to it are less focused on price changes, than the model based on the corpus could be more useful.

3.11 Twitter Data Capture – Live Data Capture

As bitcoin exchanges are 24/7 data is captured continuously for a 3 week period. In that time circa 700,000 tweets related to bitcoin are captured. Twitter provides a streaming API which allows a researcher, or end-user, or business to programmatically download tweets. For practical reasons Twitter limits the number of tweets that can be downloaded via the Streaming API[6]. For search queries with millions of related tweets a day only a fraction of

[6] They provide a pay service called Twitter Firehose (through 3rd parties) that guarantees 100% of all tweets.

these will be returned. Early benchmarking as part of this project showed that for a term such as ‘bitcoin’ all of the tweets were captured, as could practically be observed.

3.12 Bitcoin Price data

Price data is taken from the website Coindesk (2014). This site provides a price index based on an aggregate from a number of exchanges, called the Bitcoin Price Index (refer to Appendix B for how this value is calculated).

The Bitcoin Price Index (BPI) represents an average of bitcoin prices across leading global exchanges that meet criteria specified by the BPI. The criteria for an exchange to be included are:

1. USD exchanges must serve an international customer base.
2. Exchange must provide a bid-offer spread for an immediate sale (offer) and an immediate purchase (bid).
3. Minimum trade size must be less than 1,500 USD (9,000 CNY) or equivalent.
4. Daily trading volume must meet minimum acceptable levels as determined by CoinDesk.
5. Exchange must represent at least 2% of the total 30-day cumulative volume for all of the exchanges included in the BPI.
6. Fiat currency and bitcoin transfers in or out of the exchange must be completed within seven business days and 24 hours, respectively.

At the time of the research the following bitcoin exchanges were included in the US dollar BPI calculation:

- Bitfinex – Hong Kong based
- Bitstamp – UK based
- BTC-e – Bulgaria based
- LakeBTC – Shanghai based

CoinDesk provides a simple API to make its Bitcoin Price Index (BPI) data programmatically available to others. This service is updated with the latest value every 60 seconds. For a sample response and how to query the service, refer to Appendix B.

3.13 How Sentiment is Measured

Each tweet is evaluated against the model to determine its sentiment. Each tweet will be given a score, -1 for Negative, 0 for Neutral and 1 for Positive. The scores will be aggregated over three time frames: 1 hour, 8 hours, and 24 hours. The 8 hour timeframe was selected to represent a notional trading day, as bitcoin market is 24/7. As the Bitcoin Price Index used is based on the four exchanges in geographically dispersed locales trading occurs throughout the 24 hour period. Bitcoin prices are measured for each tweet but for analysis purposes the point to point value across each time frame, representing opening and closing values for the time frame concerned (or opening and closing price), are captured. The total number of tweets 24 hour, per 8 hour and per 1 hour period are also recorded. As are the number of tweets for each sentiment category, which will be used for calculating the *bullishness* value as described in Chapter 4. The trading volumes of bitcoin for each day are taken from the website Coindesk (2014), who provide the transaction volumes as a downloadable csv file.

Time series analysis is performed on this data to assess whether there is a correlation between these variables. In order to discover if there is a lead-lag relationship between the two variables, cross-correlation analysis is used to calculate the cross-correlation function, or CCF. The cross-correlation function shows the correlation between two series at the same time, and with each series leading by one or more lags. By inspecting the CCF between two series, the lag when they are most highly correlated can be determined. The bitcoin prices are transformed to a stationary process in order to perform cross correlation. This is done by differencing, subtracting the previous value to calculate the change in price between the time periods, 1 hour, 8 hour and 24 hours. One disadvantage of differencing is that one time observation is lost, the first, as no previous value exists. This can be mitigated for the 24 hour time period, as the previous days price is publicly available.

When the CCF value with the strongest value is calculated, the lag will be applied to the data and Pearson's *r* will be used to measure the correlation at that point in time. Pearson's *r* is a measure of the linear correlation between two variables *X* and *Y*, giving a value between +1 and −1 inclusive, where 1 is total positive correlation, 0 is no correlation, and −1 is total negative correlation. Pearson's *r* will also be used for testing the relationship between the Twitter message volumes and the bitcoin transaction volumes.

3.14 Missing Data

There are two timeframes when the data collection stopped running. Firstly a 4 hour period when the Twitter authentication failed. Secondly, for a 14 hour period when the cloud based server that hosted the application collecting the tweets had an unscheduled outage. The application was not set to start on server startup so the outage was magnified to 14 hours. The gaps in data are being handled as follows. Overall sentiment is set to 0 for these hours. For bitcoin prices the missing hours are filled in with the average between the last two collected values. For Twitter message volumes the average from the previous and subsequent day's volumes for the same hours are used for the missing volume data.

3.15 Conclusion

The path through the research onion is complete and the methodology and process has been outlined for this research. This study is based on a philosophy of positivism. A deductive approach that creates machine learning based experiments. The experiments use a custom classification model based on the Naïve Bayes algorithm. The experiments will generate secondary data for a cross-sectional timeframe. Quantitate analysis in the form of cross-sectional and correlation analysis is performed on the data that is produced. The specific approach used to capture the data has been outlined. The sources and the sample size have also been described.

4 Findings and Analysis

4.1 Findings and Analysis Introduction

This chapter outlines the findings and analysis based on the data collected for this research. Firstly, an overview of the data collected in the time frame that will help to frame the analysis. Then each of the research questions will be addressed by performing quantitative and statistical analysis of the data collected. The data will be examined in the context of the main research question and the sub questions.

Tweets containing the word 'bitcoin' were collected from the Twitter streaming API for a continuous 3 week period. This resulted in the collection of 741,434 bitcoin related tweets. The exchange rate of bitcoin was collected continuously for the same period. The period of data collection is noteworthy as for the first two weeks the price was stable, at just below 600 dollars for 1 bitcoin. In the third week the price dropped to below 500 dollars. The linear chart below shows the price variation for the research period.

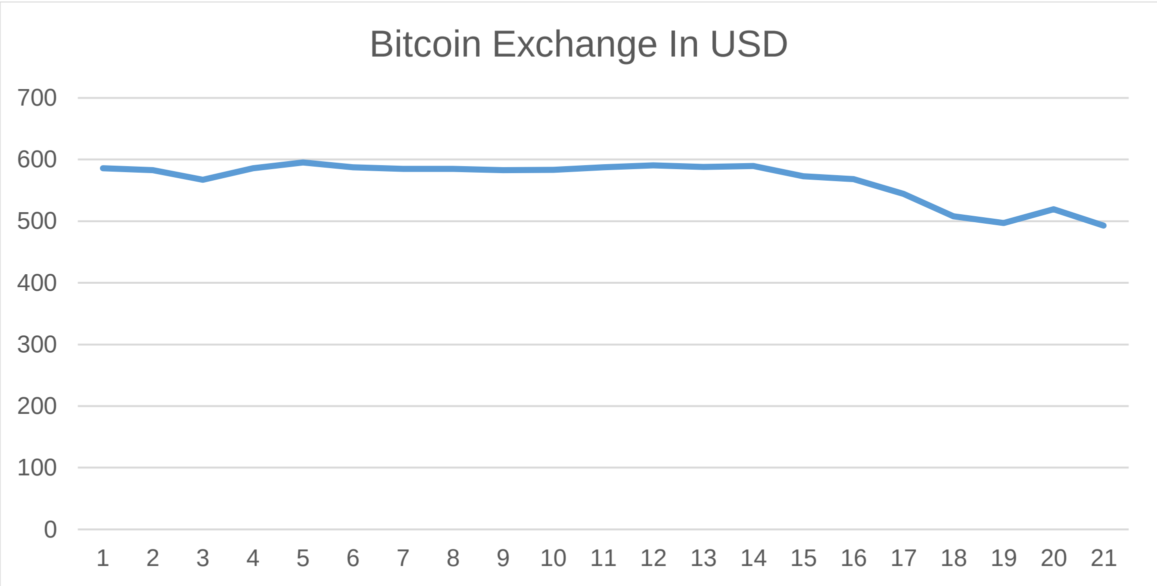


FIGURE 4.1 Bitcoin exchange price over 21 day period

In examining whether or not the bitcoin price can be correlated to information from Twitter, there are a number of factors that can be looked at. The first is message volume and whether and how that relates to bitcoin transaction volume and bitcoin price fluctuation. The main research question is then examined by looking at whether the sentiment contained within tweets can be used to predict the future exchange rate of bitcoin.

4.2 Twitter Message Volume

Trading volume has been shown previously to be a proxy of sentiment (Baker and Wurgler, 2007)(see list of sentiment proxies as listed in Chapter 2). High trading volumes can be connected to periods of excessive buying or selling of stock and subsequent rises or falls (Jones et al., 1994). The total number of tweets related to bitcoin per day are examined to assess whether there is a correlation to transaction volumes and to price fluctuations. The number of tweets related to bitcoin is collected, and the value is recorded on an hourly, 8-hourly and 24 hour basis. All of the transactions ever carried out on the Bitcoin system are available on the internet (the data on transactions is available but the users are anonymous). An important point for the subsequent analysis is the fact that the transactions data will represent both the purchases of products and services with bitcoin, and trading on exchanges. The bitcoin daily transaction volumes are taken from the website coindesk.com (2014). The volume of transactions is only available on a per day basis and is compared against the Twitter message volume for 24 hours.

In order to compare number of tweets to bitcoin price change the amount of change per day as a percent is calculated as follows.

$$C = (|P_t - P_{t-1}|/P_{t-1}) * 100$$

Where *C* is the percentage of change. This is the absolute value of the difference in price between days where *P* is the closing price for the day, and *t* is the day. This is divided by the previous days' value and multiplied by 100 for percentage change.

Before running the analysis it is important to ensure that the bitcoin digital currency has historically followed the trend seen in other financial markets, namely market volumes correlate to price change. Using data from the previous calendar year, a correlation test is performed. The results are displayed in Table 4.1.

TABLE 4.1 Correlation of Bitcoin transaction volume and Bitcoin price fluctuation for the year from July 1st 2013 to June 30th 2014

		Transaction Volume	Price Fluctuation
Transaction Volume	Pearson Correlation	1	
	Sig. (2-tailed)		
	N	365	
Price Fluctuation	Pearson Correlation	.274**	1
	Sig. (2-tailed)	.000	
	N	365	365

**. Correlation is significant at the 0.01 level (2-tailed).

A Pearson's *r* data analysis revealed a modest positive correlation, *r=.27* indicating that there is a modest correlation between Transaction Volume and Price Fluctuation for a sample size of 365 days. As the sample size is large the result is significant. Thus, it can be stated that the historical correlation between Transaction Volume and Price Fluctuation has been shown. When the Transaction Volume per day is high, so is the fluctuation in price.

Turning to data collected for the 3 week period as part of this study, a linear chart (Figure 4.2) shows that there appears to be a correlation between the number of tweets and Transaction Volume.

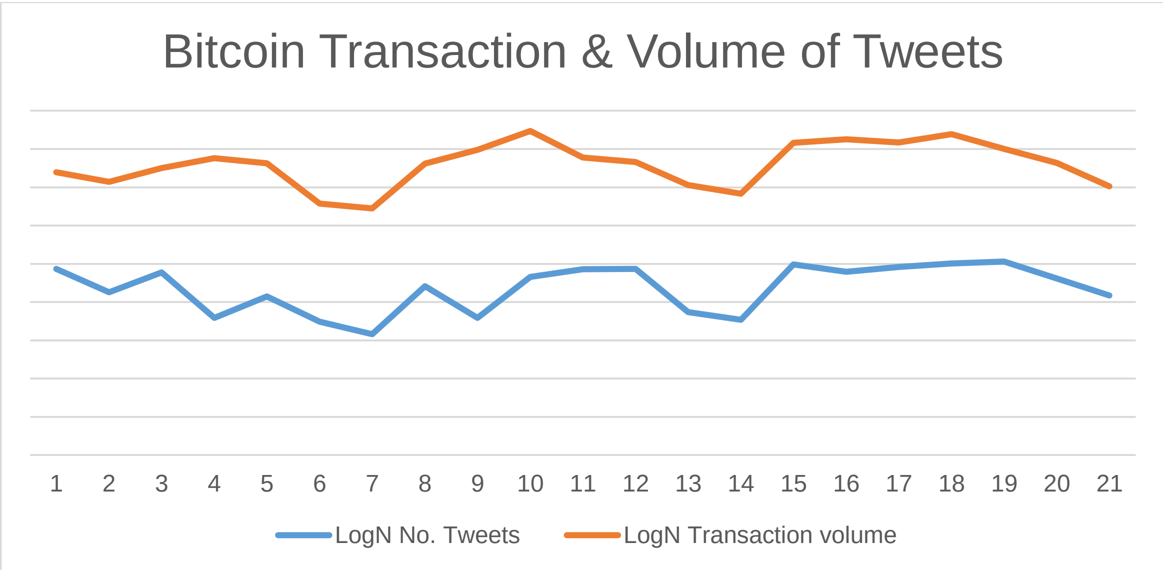


FIGURE 4.2 Natural log of daily volume of tweets and bitcoin transaction volumes.

To determine the strength of the correlation Pearson's r analysis is performed. The price fluctuation of bitcoin for the 3 week period under analysis is also included. Table 4.2 shows the results of a correlation analysis of the three variable Number of Tweets per day, Transaction Volumes of bitcoin and Price Fluctuation of bitcoin.

TABLE 4.2 Number of Tweets, Transaction Volume and Price Fluctuation Correlations

		Number of Tweets	Transaction Volume	Price Fluctuation
Number of Tweets	Pearson Correlation	1		
	Sig. (2-tailed)			
	N	21		
Transaction Volume	Pearson Correlation	.690**	1	.
	Sig. (2-tailed)	.001		
	N	21	21	
Price Fluctuation	Pearson Correlation	.340	.282	1
	Sig. (2-tailed)	.132	.215	
	N	21	21	21

**. Correlation is significant at the 0.01 level (2-tailed).

The total number of Twitter messages and the bitcoin transactions per day have a strong Pearson *r* value, $r = .69$. There is a strong correlation between the two variables. Further, the number of transactions have a modest correlation to bitcoin price fluctuations with a result of (r =.282). Significantly the number of Twitter messages per day related to bitcoin price fluctuation has a higher Pearson *r* value of $r = .34$. In summary when the number of tweets related to bitcoin is low/high the transaction volume is low/high and price fluctuation is low/high. The number of tweets has a stronger correlation to the price fluctuation of bitcoin than the transaction volume. This is a significant result that requires more analysis.

On further examination there is a pronounced difference in the data covering a weekend. Trading volumes are consistently low for each weekend but on the final weekend there was a significant change in the price of bitcoin. A one day gain of 4.67 percent followed by a fall of 5.17 percent. For these days the volume of messages on Twitter related to bitcoin rises more significantly than trading volumes, as shown in Table 4.3.

TABLE 4.3 Sunday Twitter volumes and number of bitcoin transactions with Price Fluctuation (full table available in Appendix C)

Date	Number of Tweets	Transaction Volume	Price Fluctuation
02/08/2014	29667	54989	1.31
03/08/2014	27787	53621	0.30
09/08/2014	31180	60599	0.41
10/08/2014	29952	57913	0.23
16/08/2014	37174	67974	4.67
17/08/2014	34020	60223	5.17

A correlation analysis of the weekend value is shown in Table 4.4.

TABLE 4.4 Weekend Twitter Volumes, Transaction Volumes and Price Fluctuation Correlations***

		Weekend Tweets	Weekend Volume	Weekend Fluctuation
Weekend Tweets	Pearson Correlation	1	*	.
	Sig. (2-tailed)			
	N	6		
Weekend Volume	Pearson Correlation	.946**	1	
	Sig. (2-tailed)	.004		
	N	6	6	
Weekend Fluctuation	Pearson Correlation	.870*	.669	1
	Sig. (2-tailed)	.024	.146	
	N	6	6	6

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

***For comparison weekday correlation analysis is available in Appendix C.

The Number of Tweets outperforms the Transaction Volumes as a correlation of the price fluctuations more noticeably in this case. The Pearson $r = .87$ as opposed to $r = .69$. The data set is clearly too small to derive a long term prediction but it seems to suggest that Twitter messages have improved correlation on a weekend, particularly when there are major market swings. This would suggest Twitter is a better barometer of investor (or trader) sentiment than transaction volumes. Transaction volumes would cover both speculation and general transactions associated with the purchase of goods using bitcoin. It would appear then, that number of Twitter messages are more correlated with trading in bitcoin than with general bitcoin transactions. What these values show is that for bitcoin, Twitter volumes can be a better proxy for sentiment than volume of transactions.

In answer to research question (R2), “*Does the volume of Twitter messages relate to bitcoin price movement?”* there are three findings:

Finding 1. Number of Tweets per day is a proxy of investor sentiment for bitcoin.

Many studies have linked trading volumes with change in prices or volatility. It has been used repeatedly as a proxy of investor sentiment. The analysis performed shows that Twitter message volume also correlates to the price fluctuation of bitcoin. On days when the number of Twitter messages is high the price fluctuation is high, and vice versa.

Finding 1(a). Number of Tweets per day is strongly correlated to transaction volumes of bitcoin.

The number of tweets per day is strongly correlated to same day transaction volumes for the data in this study. This is in agreement with Antweiler and Frank (2004) and Sprenger et al. (2013) who have found a correlation between number of messages and trading volumes. This study differs from Sprenger et al. in that the correlation is for same day transaction volumes, whereas Sprenger et al. find that Twitter message volumes lead trading volumes by one to two days. However, Antweiler and Frank (2004) find both intraday and next day effects of message boards on trading volumes. Twitter message volume does not correlate to next day bitcoin transaction volumes. A correlation of the number of Twitter messages with next day trading volumes was performed as part of this study. The Pearson’s *r* value was negative at -.06. This can be explained by the nature of bitcoin exchanges. They trade 24/7, so there is no pause in trading. At weekends, number of Twitter messages and trading volumes decrease. For next day analysis this does not follow through to the Monday. Therefore the correlation is strongest on same day trading volumes as opposed to next day trading volumes.

Finding 1(b). Number of Tweets is more correlated to price fluctuation than transaction volumes.

Perhaps the most interesting finding is that the correlation between number of tweets and bitcoin price fluctuation is stronger than the correlation between transaction volumes and price fluctuation. Trading volume is a well-established barometer of price fluctuation and volatility as documented by numerous research (Jones et al., 1994). Data from the weekend analysis seems to suggest that Twitter volumes are a better barometer of pure trading than transaction volume. Perhaps, with bitcoin it can be explained by the fact that the transaction data used includes trading activity and normal purchases. If trading values alone were available they may perform better in relation to price fluctuations.

4.3 Sentiment of Tweets as a Predictor

The main research question asks if the sentiment on Twitter can predict bitcoin exchange rate. As the data set is much too large for manual classification, machine learning tools are used to automatically classify each tweet. Before classification, each tweet is checked, and some non-English tweets removed[7]. Then each tweet is classified using the custom model produced for this study. Each tweet is assigned a score of 0 for Neutral, 1 for Positive and -1 for Negative based on the results for the classifier. These values are aggregated over 1 hour, 8 hour, and 24 hour to give a sentiment score to each time period. Other variables are also tracked, an overall sentiment value and the number of tweets of each category in each of the three time periods. For bitcoin the current price, percent change and amount of change are calculated for the same time periods. The tables below summarise the data for the 1 day time frame.

TABLE 4.5 Bitcoin prices changes over 21 day period

Bitcoin Prices Each Day			
Date	Bitcoin Price	Price Change (Amount)	Price Change (Percent)
28/07/2014	584.69	-6.26*	-1.06*
29/07/2014	582.20	-2.49	-0.43
30/07/2014	564.37	-17.83	-3.06
31/07/2014	581.35	16.98	3.01
01/08/2014	595.08	13.73	2.36
02/08/2014	587.29	-7.79	-1.31
03/08/2014	585.51	-1.78	-0.30
04/08/2014	586.76	1.25	0.21
05/08/2014	583.11	-3.65	-0.62
06/08/2014	583.04	-0.07	-0.01
07/08/2014	587.40	4.36	0.75
08/08/2014	590.53	3.13	0.53
09/08/2014	588.09	-2.44	-0.41
10/08/2014	589.45	1.36	0.23

[7] Twitter streaming API does not have the ability to filter on language, non-English tweets were removed by checking the tweets for certain accented characters with the tweets, as such there were some non-English tweets that were not filtered out but they were negligible.

11/08/2014	573.31	-16.14	-2.74
12/08/2014	568.21	-5.1	-0.89
13/08/2014	544.57	-23.64	-4.16
14/08/2014	508.55	-36.02	-6.61
15/08/2014	496.62	-11.93	-2.35
16/08/2014	519.83	23.21	4.67
17/08/2014	492.95	-26.88	-5.17

*Calculated based on closing price from the 27/08/2014 taken from coindesk

TABLE 4.6 Twitter sentiment for each day in the time period.

Twitter Sentiment Each Day				
Date	No. Of Tweets*	Neutral	Negative	Positive
28/07/2014	38856	23057	7668	8131
29/07/2014	34278	18476	7832	7970
30/07/2014	38204	19865	8105	10234
31/07/2014	23027	13528	5020	4479
01/08/2014	33370	19710	6955	6705
02/08/2014	28375	17860	5425	5090
03/08/2014	26632	16732	5468	4432
04/08/2014	17152	10750	3132	3270
05/08/2014	29564	17396	6358	5810
06/08/2014	36309	19998	7380	8931
07/08/2014	38116	21275	8626	8215
08/08/2014	38524	20581	9567	8376
09/08/2014	30520	19854	5045	5621
10/08/2014	29433	18294	5152	5987
11/08/2014	38774	21516	8353	8905
12/08/2014	36282	19561	8598	8123
13/08/2014	38824	22381	7428	9015
14/08/2014	39519	22646	7316	9557
15/08/2014	39096	20808	8313	9975
16/08/2014	35871	22227	7386	6258
17/08/2014	31341	20505	6077	4759

* It should be noted that the number of tweets per day is less than in table 4.3. This is accounted for by the fact that non-English tweets are used for the overall count but not when classifying.

The model used to classify will only match for strong signals of sentiment for positive or negative towards bitcoin, hence there are more neutral values in each day as much of the content on twitter merely references bitcoin and does not actually relate to the currency, for example:

> *Shopping for Your Health Care: Can You Tell if the Price Is Right? http://t.co/aDp889F9Ch #money #dogecoin #bitcoin #news,Â #love*

In order to examine the research questions RQ1 and RQ3 of this paper, cross-correlation is performed. This is to determine if there is a lag time between sentiment, as observed on Twitter, and the time for that sentiment to filter into the market. If the sentiment of tweets does predict bitcoin prices it can be expected to lead the bitcoin exchange rate. If the change in price is driving the change in Twitter sentiment, the bitcoin price will lead the sentiment. In this case the Twitter sentiment is merely reflecting the price change. If neither is the case, there should be a negative or no correlation. One of the main difficulties in trying to assess bitcoin in this way is to establish the ideal timeframe to run cross-correlation tests. For traditional stocks, the market operates on an eight hour window. Sentiment expressed at the end of the trading day can be applied to next day prices as done in (Antweiler and Frank, 2004) and (Sprenger et al., 2013). With bitcoin, as the market is available on a 24/7 basis, it is difficult to predict when the sentiment will filter through. It could be an hour later in some cases or the following day in others. Figure 4.3 shows the aggregated sentiment value (*Positives – Negatives*) for each 24 hour period over the 21 days of the data capture on a linear scale.

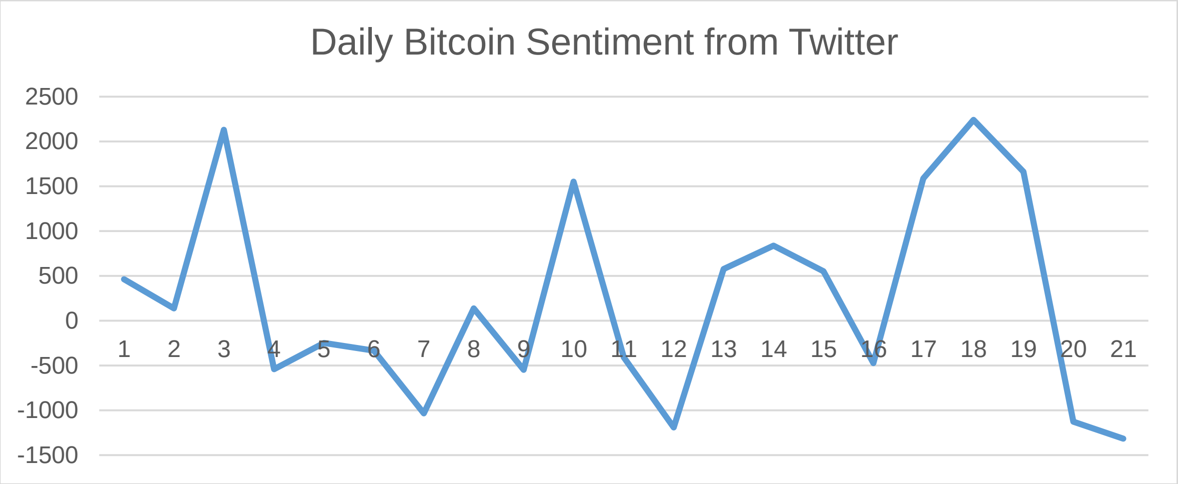


FIGURE 4.3 Daily Bitcoin Sentiment from Twitter as produced be automatic classification of Tweets

For comparison Figure 4.4 shows the bitcoin daily change as a percentage for the same time period.

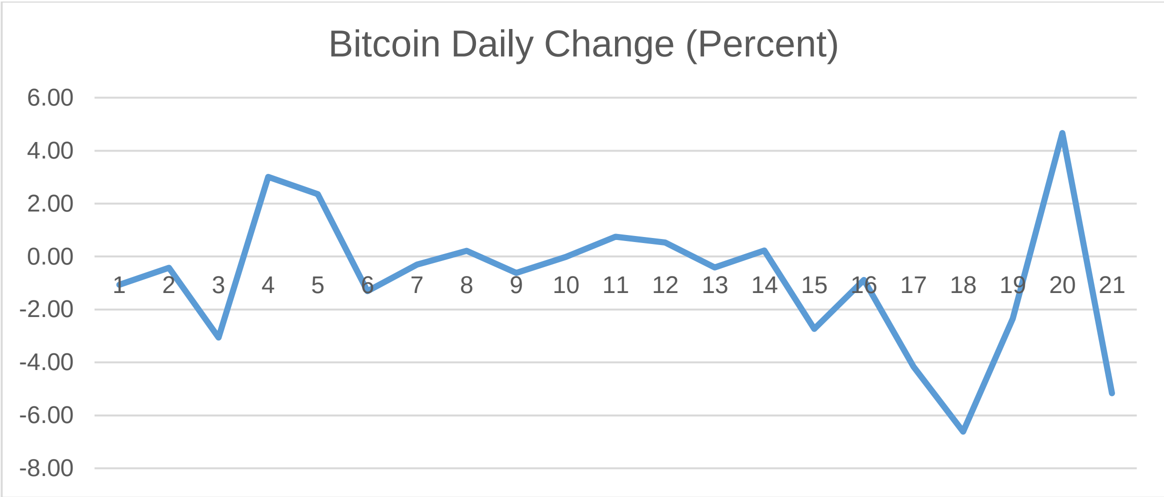


FIGURE 4.4 Bitcoin daily price change.

To examine whether or not the sentiment is a leading factor cross-correlation needs to be performed. If sentiment is a leading factor the results should show that the correlation with a lag (*l*) greater than 0 is stronger than at 0, or at a lag less than 0. Cross-correlation is run for all the time periods used for aggregation. The 1 hour results were not significant and will not be reported on, with the analysis provided in Appendix C. The cross-correlation of the aggregate sentiment for the 24 hour period is shown in Figure 4.5.

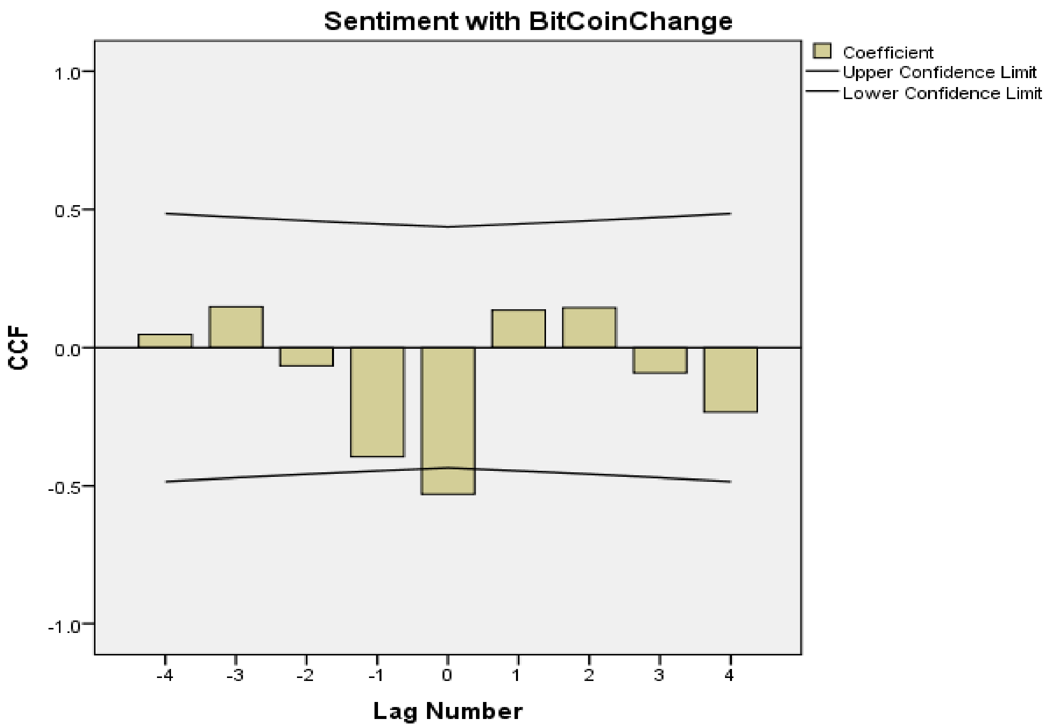


FIGURE 4.5 Cross correlation of Twitter Sentiment aggregated for 24 hours to Bitcoin price change in 24 hour period

The lag value (*l*) at 1 and 2 yield positive correlation of .135 and .144 respectively. That is, the strongest positive correlations are observed after 1 day and 2 days, i.e. the value of the Twitter aggregate sentiment is most closely correlated to bitcoin price after 1 and 2 days. The results of running the same test for the 8 hour timeframe of aggregated sentiment are shown in Figure 4.6.

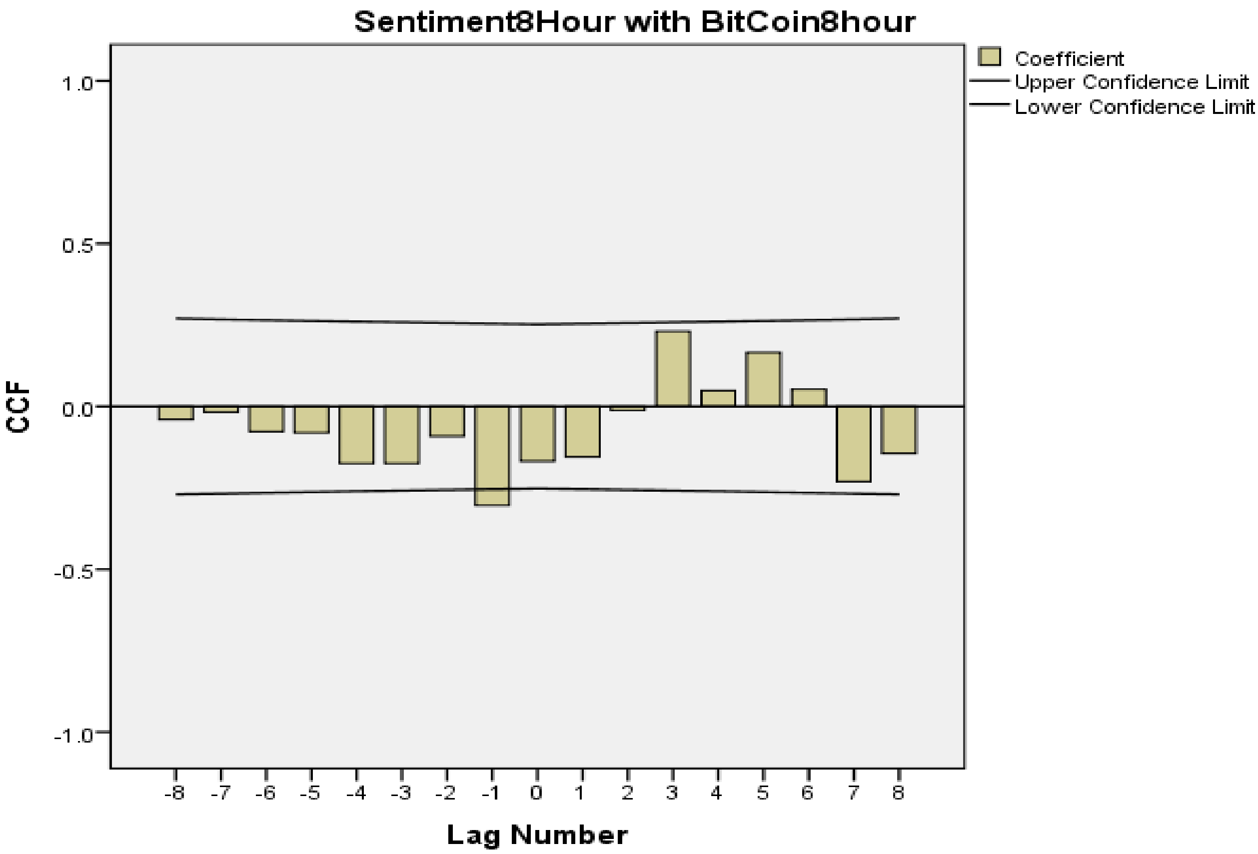


FIGURE 4.6 Cross correlation of Twitter Sentiment aggregated for each 8 hours to Bitcoin price change for each 8 hours

The cross-correlation of the 8 hour aggregate sentiment shows the strongest correlation at a lag value of *l* = 3, as shown in Table 4.7

TABLE 4.7 Strongest cross correlation

Lag	Cross Correlation	Std. Error[a]
3	.231	.129

A lag of 3 represents an elapsed time period of 24 hours. The 8 hour data will represent a more fine grained measurement reflecting fluctuations in three 8 hour periods that would be levelled when aggregated over 24 hours. Both sets of time frames report positive correlations when the bitcoin price leads the Twitter message sentiment. It appears that there is a positive correlation after 24 hours as best represented by 8 hour samples. The same data will be examined using a different measure of sentiment, namely bullishness.

4.3.1 Calculating bullishness value

Another measure as proposed by Antweiler and Frank (2004) and as used by Sprenger et al. (2013) is Bullishness. Bullishness value as defined as:

$$B_t = \ln (1+MBuy_t)/(1+MSell_t)$$

Where MBuy (MSell) represents the number of buy or sell signals in day. This measure reflects both the share of buy signals as well as the total number of messages giving greater weight to a larger number of messages expressing a particular sentiment. The Bullishness value is calculated for the data captured, using the same timeframes as before. The cross-correlation of Bullishness and bitcoin price change for the 24 hour timeframe is shown in Table 4.8.

TABLE 4.8 Cross Correlation of Bullishness value and bitcoin price change over the 24 hour time frame

Cross Correlations BullishnessDay with BitCoinChange

Lag	Cross Correlation	Std. Error[a]
-4	.128	.243
-3	.202	.236
-2	-.049	.229
-1	-.365	.224
0	-.450	.218
1	.068	.224
2	.112	.229
3	-.108	.236
4	-.287	.243

When calculating the cross-correlation for the 8 hour period, again the lag of 3 has the strongest correlation. This is shown in Figure 4.7

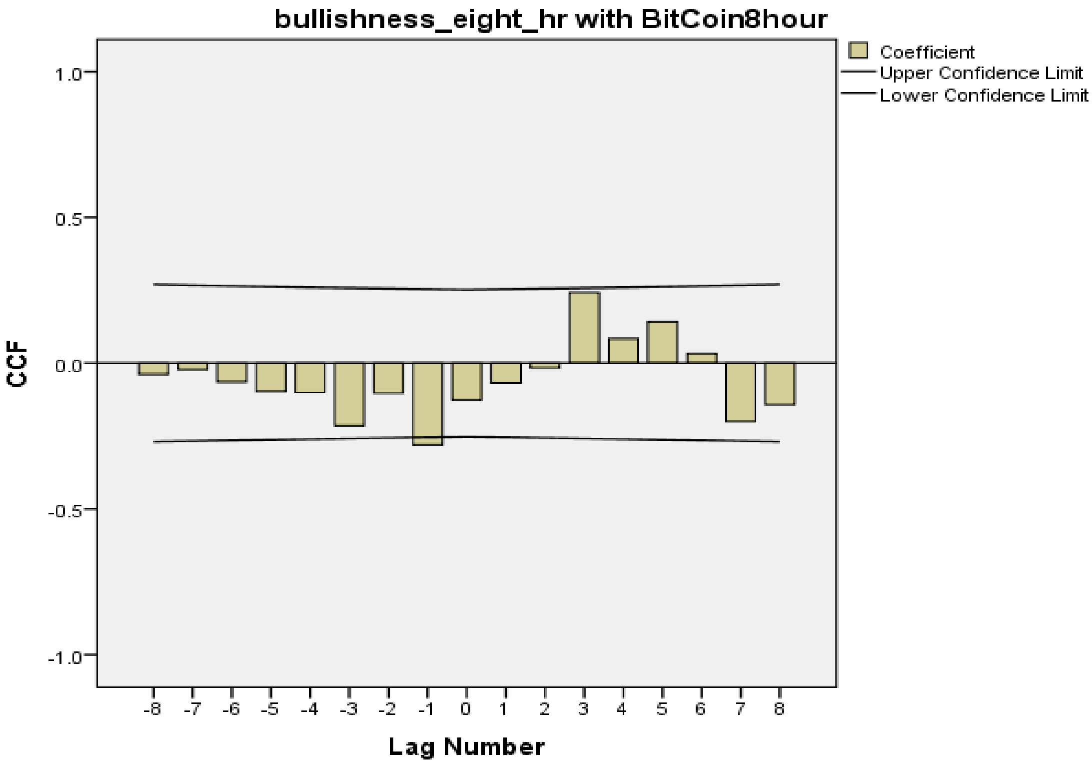


FIGURE 4.7 Cross correlation of Twitter bullishness for each 8 hours to Bitcoin price change for a day

At a lag of 3, the strongest correlation is observed. Using the Bullishness value the sentiment of Twitter is most correlated to the price of bitcoin after a time period of 24 hours. That is the price of bitcoin lags the sentiment by 24 hours. The Bullishness value is in agreement with the aggregated values used previously. Table 4.9 shows the cross-correlation value at a lag of 3.

TABLE 4.9 Strongest correlation for 8 hour time frame

Lag	Cross Correlation	Std. Error[a]
3	.242	.129

Table 4.10 contains a summary of the results obtained by using the aggregate sentiment value and the Bullishness value for the 8 hour and 24 hour time frame.

TABLE 4.10 Cross Correlation scores for 8 hour and 24 hour periods

Measurement	Time Period	Lag	Cross Correlation
Sentiment Aggregate	24 hour	2	.144
Bullishness	24 hours	2	.112
Sentiment Aggregate	8 hours	3	.231
Bullishness	8 hours	3	.242

This analysis provides answers to research questions (RQ1) and (RQ3):

(RQ1) Can the sentiment on Twitter predict bitcoin exchange rate?

For (RQ1) it can be seen that the sentiment on Twitter does predict the bitcoin exchange rate. The prediction is strongest for sentiment measured in 8 hour periods. The sentiment is reflected in a change of price after a 24 hour time delay.

Finding 2. Twitter sentiment analysis can be used to predict the currency exchange rate for bitcoin.

(RQ3) Does sentiment merely reflect bitcoin price movements or cause them?

For (RQ3) it is seen that the sentiment value is reflected in the price of bitcoin after an interval of 24 hours. Twitter sentiment leads the price of bitcoin.

Finding 3. Twitter sentiment related to bitcoin leads the change in bitcoin exchange rate

These findings will be revisited shortly. Firstly, in an effort to get a more accurate result, research question R4 will be addressed.

4.4 The Power of Retweets

Research question (R4) relates to the influence retweets have on sentiment. In Twitter a retweet is when a user rebroadcasts a tweet from another Twitter user. It can act as a powerful mechanism of disseminating messages over Twitter quickly and to a large audience (Kwak et al., 2010). It can also be a useful barometer for sentiment. It can be assumed that for the majority of the cases the person who retweets agrees with the original sentiment and hence decided to forward to their followers. By examining the data collected over the period it is clear that retweets[8] are very common in the Twitter dataset for this research. Table 4.11 shows the number of Twitter messages and their type.

Table 4.11 Number of tweets and retweets in data set.

Total Tweets	741432*
Retweet Messages	238982
Non Retweet	502450

*Non-English removed

Retweets can have a cascade effect, with the following tweet being retweeted over 1,000 times in the space of an hour.

We are proud to announce that @PlayerAuctions now accepts #bitcoin

This tweet was classified as positive thus the aggregate score for that hour was roughly +1000. To examine the influence of retweets on the bitcoin related dataset the same statistical analysis is performed on the retweet and non-retweet dataset. Table 2.14 shows the results of cross-correlation analysis of the aggregate sentiment value and the change in bitcoin price as before. The results for the full dataset with all messages is also included for reference.

TABLE 4.12 Cross correlation results of retweets only and no retweets 24 hour period

Aggregate Sentiment 24 hour

Lag	Cross Correlation No Retweets	Cross Correlation Retweet Only	Cross Correlation Total Dataset
-4	.057	.017	.048
-3	.133	.102	.147
-2	-.127	.035	-.066
-1	-.443	-.172	-.395
0	-.619	-.200	-.530
1	-.093	.352	.135
2	.095	.142	.144
3	.038	-.208	-.092
4	-.234	-.134	-.233

[8] Retweets are normally marked by a RT or @retweet signs, the data was filtered based on these criterion

As can be seen analysis of only retweeted data yields the strongest correlation with a lag of 1, meaning the correlation is moderate at .352 for the 24 hour period. That is, the sentiment on Twitter is most closely correlated to the value of bitcoin price change after a period of 24 hours.

The same analysis was performed for the 8 hour time period. Table 4.13 shows the results of the cross-correlation.

TABLE 4.13 Cross correlation results of retweets only and no retweets 8 hour period

Aggregate Sentiment 8 hour

Lag	Cross Correlation No Retweets	Cross Correlation Retweet Only	Cross Correlation Total Dataset
-8	-.050	-.011	-.040
-7	-.033	.006	-.017
-6	-.161	.045	-.077
-5	-.084	-.037	-.081
-4	-.189	-.076	-.175
-3	-.181	-.083	-.175
-2	-.170	.033	-.091
-1	-.425	-.033	-.304
0	-.179	-.074	-.168
1	-.103	-.131	-.155
2	-.206	.190	-.012
3	.189	.159	.231
4	-.016	.090	.049
5	.091	.159	.166
6	.090	-.011	.053
7	-.213	-.135	-.230
8	-.020	-.198	-.144

The retweet data returns the stronger correlation value of .190 at a lag of 2 (16 hours), however this is less than the total data set which had a correlation of .231 at lag 3. For both the 8 hour and 24 hour aggregate sentiment cross correlation, the retweets performed better than the dataset with retweets removed.

These results can now answer research question (RQ4):

(RQ4) Are retweets a better gauge of sentiment and are they more closely linked to bitcoin price changes?

Finding 4. Retweets are a better measure of sentiment than regular tweets.

The retweets outperform the regular tweet message on the cross-correlation analysis on this dataset. This could be explained by higher quality information being retweeted. Sprenger et al. (2013) found that above average investment advice received higher levels of retweets. In the case of bitcoin, positive and negative news stories are also likely to be retweeted. An example of one such retweet from the data set is:

RT @BitcoinAgile: Bitcoin Price Sharply Drops in Wake of US Government Report

This tweet was marked as negative by the classifier. This bad news story was retweeted several times. This finding may only hold true when there are prominent good or bad news stories being retweeted. It may not be as affective at revealing individual investor sentiment. Building a classification system solely based on retweets would have drawbacks. This will be examined in the Conclusions chapter.

4.5 Confirming Correlation with Lag Applied

With the retweet analysis and data complete, the main research question is revisited. The cross correlation analysis consistently confirmed that there is a correlation at the lag of 1 for the 24 hr time frame, and a lag of 3 for the 8 hour time frame. The results are positive, as they are in agreement. The question remains which time frame is most suited for predicting bitcoin, and how strong the correlation is when the lag value is applied. When the lag is applied the sentiment data is shifted forward, meaning we lose one observation. To mitigate against this, the bitcoin prices for the subsequent days following the time period under test have also been captured. This enables the lag value to be applied and tested against these new values. The two most significant cross-correlation results will be evaluated, i.e. the 8 hour Bullishness value and the 24 hour aggregate of sentiment from the retweet dataset.

Bullishness - 8 Hour Time Period

The cross-correlation analysis revealed that for the Bullishness measure of 8 hour sentiment data, the strongest correlation was at the lag of 3. By shifting the data[9] for Bullishness forward by 3 values for the 8 hour period the correlation and Pearson's *r* can be calculated. The shifted data will represent the optimum point of correlation. This is Twitter sentiment with next day bitcoin price change. The time shifted data is shown in Figure 4.8 and Figure 4.9 for reference.

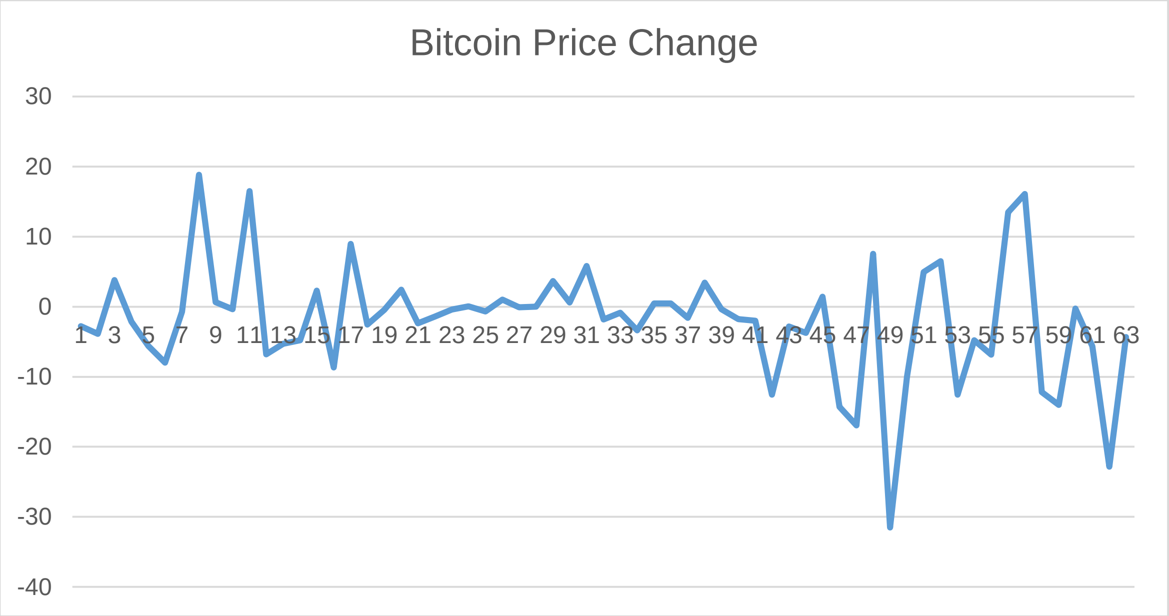


FIGURE 4.8 Bitcoin Price Change intervals of 8 hours

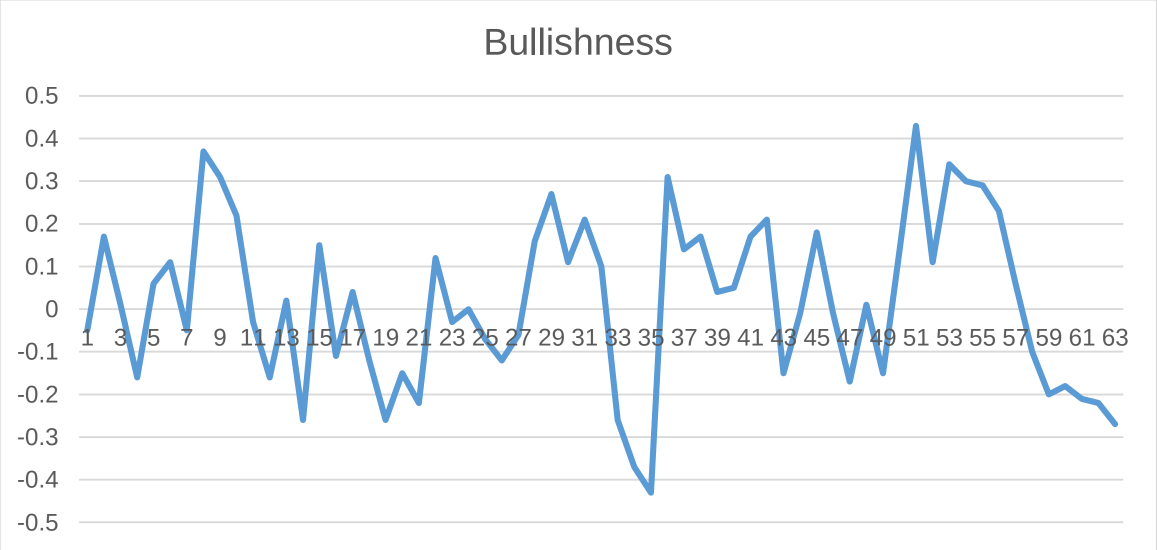


FIGURE 4.9 Bullishness value aggregated over 8 hour period

[9] Time shifted data available in Appendix C

The correlation analysis with the lag applied is shown in Table 4.14.

TABLE 4.14 Correlation of Bullishness and Bitcoin price for 8 hour aggregate with lag of 3 applied

		Bullishness 8hour	Bitcoin Price Change 8hour
Bullishness 8hour	Pearson Correlation	1	
	Sig. (2-tailed)		
	N	63	
Bitcoin Price Change 8hour	Pearson Correlation	.298*	1
	Sig. (2-tailed)	.018	
	N	63	63

*. Correlation is significant at the 0.05 level (2-tailed).

A Pearson's *r* data analysis revealed a positive correlation *r*=.298 indicating that there is a modest correlation between sentiment and the price of bitcoin using the bullishness measure. Rounding this value to 2 decimal places as is typically performed give *r* = .3.

Aggregate Sentiment Retweets - 24 Hour Time Period

The cross-correlation analysis revealed that, for aggregated sentiment of retweets for the 24 hour time frame, the strongest correlation was at a lag of 1. By applying the lag of 1, i.e. shifting the sentiment value forward by one day for the 24 hour time period, the correlation can be tested and a resultant measure for Pearson's *r* calculated. The time shifted data is shown in Figure 4.10 and Figure 4.11 for reference.

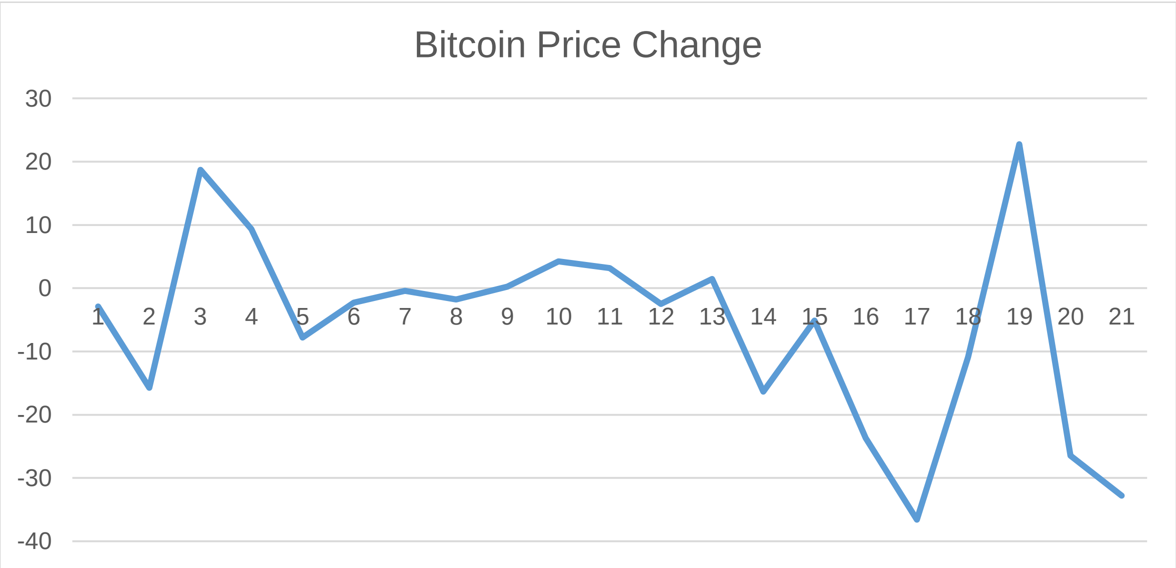


FIGURE 4.10 Bitcoin Price Change intervals of 24 hours

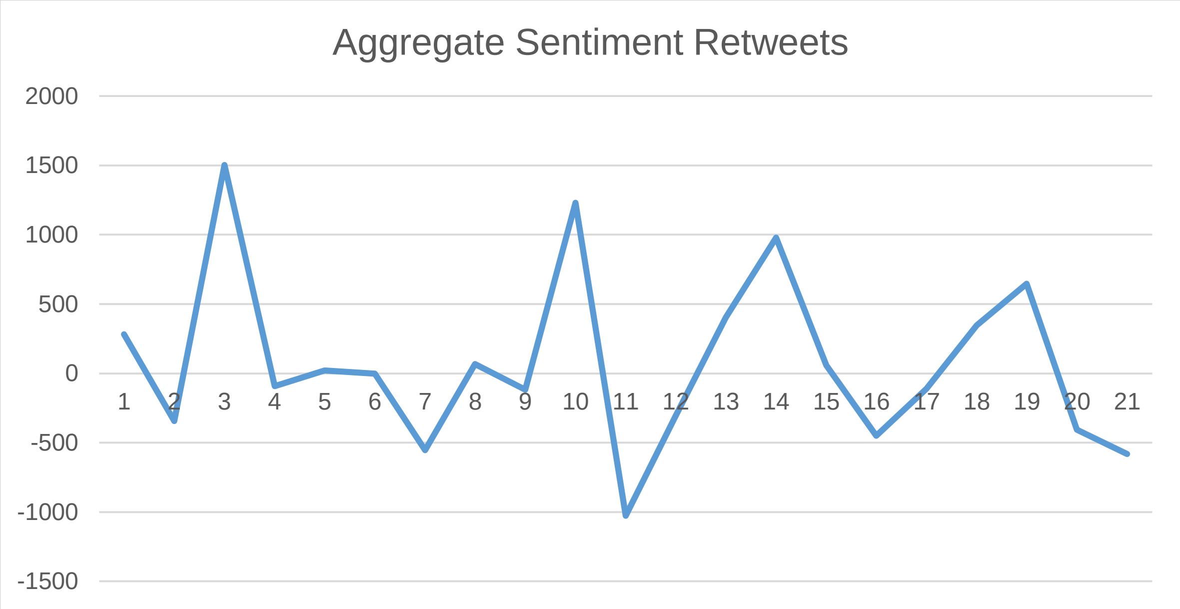


FIGURE 4.11 Aggregate sentiment of retweets intervals of 24 hours

The correlation analysis with the lag applied is shown in Table 4.15.

TABLE 4.15 Correlation results of sentiment and retweets only for 24 hour period

		Sentiment 24hr Retweet Only	Bitcoin Price Change
Sentiment 24hr Retweet Only	Pearson Correlation	1	
	Sig. (2-tailed)		
	N	22	
Bitcoin Price Change	Pearson Correlation	.440*	1
	Sig. (2-tailed)	.040	
	N	22	22

*. Correlation is significant at the 0.05 level (2-tailed)

A Pearson's *r* data analysis revealed a positive correlation *r=.44* indicating that there is a strong correlation between sentiment and the price of bitcoin. This is a significant result, both measures are in agreement and show a positive correlation. The sampling and aggregation time periods differ, but they both agree on the time frame when bitcoin price will reflect the sentiment value, namely after 24 hours. Based on these correlations a model based on the sentiment of Twitter can be used to predict the price of the bitcoin exchange rate 24 hours in advance. Appendix C shows the use of Twitter sentiment to predict next day movement. The model is correct for 12 days of the 21 in the test data set. The correlation is strongest for this study when retweets and bullishness are used to calculate sentiment.

This reinforces the previous two findings, finding 2 and finding 3.

Finding 2. Twitter sentiment analysis can be used to predict the currency exchange rate for bitcoin.

This finding is in alignment with Antweiler and Frank (2004), Oh and Sheng (2011) and Sprenger et al. (2013) who have shown how message boards, micro blogs, and Twitter, respectively, can be used to predict market movements. In showing that the price of bitcoin correlates to publicly available data, this study aligns with Kristoufek (2013) and study Glaser et al. (2014). This finding also relates to the work of the behavioural economists by showing that sentiment has an effect on market prices, similar to the work of De Long et al. (1990), Baker and Wurgler (2006), and others.

Finding 3. Twitter sentiment related to bitcoin leads the change in bitcoin exchange rate.

A consistent finding from all the analysis is that the sentiment of messages on Twitter leads the change in bitcoin price by 24 hours. This timeframe is consistent with previous work Antweiler and Frank (2004) and Sprenger et al. (2013) who find sentiment in messages is reflected in next day price.

5 Conclusions and Future Work

5.1 Introduction

This chapter aims to present the conclusions from the research carried out and demonstrate that the research questions have been answered. In addition the chapter includes a discussion of the research results alongside recommendations that may add to future work in the area. Finally a view of the limitations of this research, advances in the current knowledge and outline possible future directions for research in this area.

5.2 Conclusions

There are 4 important findings from this analysis that have come from the examination of the research questions. The research questions will be recapped and a summary outlined of the findings that the research has discovered.

To recap, the primary research question:

(RQ1). Can the sentiment on Twitter predict bitcoin exchange rate?

Finding 2. Twitter sentiment analysis can be used to predict the currency exchange rate for bitcoin.
It has been shown that there is a correlation between Twitter sentiment and the exchange rate of bitcoin. The correlation is consistent for the different time frames and measures of sentiment used. Twitter sentiment leads bitcoin price, the sentiment is reflected in price after 24 hours. This finding indicates that bitcoin investors are prone to sentiment and are reacting to changing in sentiment. When sentiment is low/negative, bitcoins are sold off. When it is high/positive bitcoins are bought. A trading strategy based on Twitter sentiment could be devised to take advantage of this.

Sub questions that are relevant to this research are:

(RQ2). Does the volume of Twitter messages reveal information on bitcoin price?

This research question was answered with the following 3 findings:

Finding 1. Number of Tweets per day is a proxy of investor sentiment for bitcoin.

Finding 1(a). Number of Tweets per day is strongly correlated to transaction volumes of bitcoin.

Finding 1(b). Number of Tweets is more correlated to price fluctuation than transaction volumes.

All three findings point to the fact that the number of tweets related to a topic can reveal useful information about the topic. By measuring the number of tweets related to bitcoin, useful information related to the price fluctuations of bitcoin can be observed. Twitter volumes have been shown to accurately reflect trading volumes and to be more accurate than trading volumes in reflecting price fluctuations. In this sense the volume of tweets can be seen as a proxy of sentiment.

(RQ3). Does sentiment merely reflect bitcoin price movements or cause them?

Finding 3. Twitter sentiment related to bitcoin leads the change in bitcoin exchange rate

It has been shown through multiple cross correlations that Twitter sentiment leads the bitcoin exchange rate. Bitcoin exchange rate lags sentiment by approximately 24 hours based on the sample size in this study. One of the main difficulties in studying bitcoin is the fact that the market is 24/7. That both the 8 hour time frame for aggregation and the 24 hour value had the same result for a lag time is interesting. A much larger analysis would be required to determine the optimum time frame for aggregation and lag. The fact that correlation between retweets and bitcoin price was strongest when aggregated over 24hrs could be a reflection of the fact that, for strong waves of sentiment, it takes that duration to filter through to the majority of users.

(RQ4). Are retweets a better gauge of sentiment and more closely linked to bitcoin price changes?

Finding 4. Retweets are a better measure of sentiment than regular tweets.

Retweets have been shown to have a better correlation to price changes than regular tweets in the sample size of this study. However this finding may not hold true for a larger sample size. Retweets are useful for propagating news events quickly. For a sentiment model, this could be less effective when there are no major news events related to bitcoin. An approach to capture the increased quality of information held in retweets while still capturing the important individual investor sentiment is outlined in the Opportunities for Future Research section.

These findings can also be viewed in terms of two wider research questions that have received much focus of research.

1. Does the content of tweets contain useful information?

While manually trawling through the Twitter data, it is true to say that there is much indecipherable, unprintable, irrelevant content in the Twitter stream. With machine learning techniques, large volumes of data can be processed that make the useless data statistically relevant. Even a simple measure like the number of tweets related to a specific topic has been shown to be a useful barometer of real life events. The ability to quickly capture and analyse the data makes Twitter and excellent source of sentiment and as a predictor for financial market movements.

2. Are investors prone to sentiment?

This study also supports the notion of the sentimental investor trading on irrational noise. If the change in price of bitcoin is in reaction to sentiment, it is clear that the investors are being affected by sentiment. The bad or good news stories are often spread on Twitter, as shown previously with the retweet:

RT @BitcoinAgile: Bitcoin Price Sharply Drops in Wake of US Government Report

The bad news spread across the network. Given that it is difficult to put a fundamental price on bitcoin it is not a surprise that investors are affected by such news stories.

5.3 Limitations

There are several ways the research could be extended. Running the data capture over a longer period would help to validate the results and give a higher confidence level in the correlations. One of the main issues encountered during this research was with collecting a continuous stream of tweets. Gaps in the data severely affect the analysis when trying to show a cause and effect relationship, i.e. if 1 day of data is lost it invalidates the data. The solution was eventually moved to a cloud based server to alleviate some of the pain points around connectivity that hampered live data collection.

One of the major limitations in studying bitcoin is the fact that the market is 24/7. Choosing a timeframe to aggregate data proved difficult, as there is a sliding window of time when the sentiment can take effect. This differs from the stock market, with a defined window of closure that can be used to aggregate sentiment around, as most of the studies with Twitter and stocks have done. One approach that was considered was to base the study on one exchange in a particular time zone. There is then an added difficulty in associating tweets from users from a particular time zone, which was deemed to be more difficult. Using the index value of multiple bitcoin exchange rates and the all the Twitter stream was deemed to be more complete. With a larger data set over several months an optimum sentiment aggregation time frame might emerge. For a trading model based on this approach the optimum time frame would be of great importance to maximise profit.

The training data set used to build the classification model is quite small, a bigger model should tend to be more accurate. The model is also very domain specific, and seems to capture the sentiment that appears currently in terms of bitcoin quite well. This is probably why the model based on the Twitter corpus performed worse than the custom model. As bitcoin becomes more mature and enters into the mainstream, terms such as boom, bust and bubble may no longer be used. Then a more generic model of sentiment may prove more effective. Another issue noted was that the model could become stale, as the terms that are associated with bitcoin now may not be in the future. Tweets about government regulation are normally of a negative connotation. Over time a classification model would need to be updated to reflect the latest trends and terms.

On building a model for classification, the approach used by the Stanford researchers, Go, Bhayani et al. (2009) in building a Twitter corpus on emoticons is certainly an interesting idea. As stated previously, such an approach was attempted at the beginning of this research but abandoned due to the low number of bitcoin related tweets that also had emoticons. As there are now more tweets related to bitcoin than there were at the beginning of the work (benched marked at 180,000 tweets per month in November 2013, now there is up to 1 million), it may be easier to collect the training automatically.

5.4 Opportunities for Future Research

Twitter and bitcoin seemly offer the perfect combination of publicly accessible data. All bitcoin transaction data is public (but with anonymous users). As shown in the literature review, Twitter has been proven to be an excellent source of user sentiment. Research in both areas will continue to grow. A number of papers are just now appearing related to bitcoin market prices and doubtless many more will follow.

To further this research a model based on weighted retweets may prove more accurate. Discarding regular tweets would not seem like a good long term approach. As observed for the 8 hour run, the combined data performed better than the retweets. It was also observed that there was a period where no retweet values were recorded. A model based on retweets will suffer as a result. Retweets will perform well when major news events have a significant impact on sentiment. A model based on retweets will pick up the repeatedly retweeted value and predict the price accordingly. However, in times without a major news event, and of relative stability, the individual investor tweets will be lost with a retweet only approach. Thus a model based on weighted retweets, for example instead of a +1/-1 for a positive/negative, retweets will be marked as +2 for Positive, -2 for Negative, may be more effective. A long running analysis would need to be performed to find the optimum weighted value.

Another improvement for future research would be the use of a sentiment control in order to cross reference results against Twitter. This would be used to establish with more certainty that the sentiment of tweets is actually providing the useful information, and that Twitter is not merely acting as a proxy of bitcoin related news. This approach was considered for this project, but no suitable control could be found. No mainstream news outlet covering the markets currently cover the main Bitcoin related news. Occasionally a bitcoin story makes its way into the mainstream media but it could not be relied on. Also, as many of the bitcoin related sites seem to favour positive news stories, the control may be skewed. There are more objective sites appearing like Coindesk (2014) that could possibly be used for any future work.

In order to test the correlations present between Twitter sentiment and bitcoin price, a trading model could be built based on the findings and approach in this paper. Even with a weak correlation (weaker than found in this research), a trading model built on sentiment should be profitable should the predictive power be is as projected. Building a trading model is the only real way to prove the predictions. Such a model, if effective, would be of particular interest to those interested in trading in bitcoin.

References

ALI, A., HWANG, L.-S. & TROMBLEY, M. A. 2003. Arbitrage risk and the book-to-market anomaly. *Journal of Financial Economics,* 69**,** 355-373.

ANTWEILER, W. & FRANK, M. Z. 2004. Is all that talk just noise? The information content of internet stock message boards. *The Journal of Finance,* 59**,** 1259-1294.

ASUR, S. & HUBERMAN, B. A. Predicting the future with social media. Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 IEEE/WIC/ACM International Conference on, 2010. IEEE, 492-499.

BAKER, M. & STEIN, J. C. 2004. Market liquidity as a sentiment indicator. *Journal of Financial Markets,* 7**,** 271-299.

BAKER, M. & WURGLER, J. 2000. The equity share in new issues and aggregate stock returns. *The Journal of Finance,* 55**,** 2219-2257.

BAKER, M. & WURGLER, J. 2006. Investor sentiment and the cross-section of stock returns. *The Journal of Finance,* 61**,** 1645-1680.

BAKER, M. & WURGLER, J. 2007. Investor sentiment in the stock market.

BARBERIS, N., SHLEIFER, A. & VISHNY, R. 1998. A model of investor sentiment. *Journal of financial economics,* 49**,** 307-343.

BARRATT, M. J. 2012. Silk road: eBay for drugs. *Addiction,* 107**,** 683-683.

BITCOIN.ORG. 2014. *Bitcoin FAQ* [Online]. Available: https://bitcoin.org/en/faq.

BITCOINATMMAP. 2014. Available: http://bitcoinatmmap.com/.

BLACK, F. 1986. Noise. *The journal of finance,* 41**,** 529-543.

BLOOMBERG. 2014a. *Bitcoins Can't Shake Bubble Image in Poll After 45% Drop* [Online]. Available: http://www.bloomberg.com/news/2014-07-17/bitcoins-can-t-shake-bubble-image-in-poll-after-45-drop.html.

BLOOMBERG. 2014b. *Can the Bloomberg terminal be toppled?* [Online]. Available: http://fortune.com/2014/03/20/can-the-bloomberg-terminal-be-toppled/.

BOLLEN, J., MAO, H. & ZENG, X. 2011. Twitter mood predicts the stock market. *Journal of Computational Science,* 2**,** 1-8.

BROWN, G. W. & CLIFF, M. T. 2004. Investor sentiment and the near-term stock market. *Journal of Empirical Finance,* 11**,** 1-27.

CHOI, H. & VARIAN, H. 2012. Predicting the present with google trends. *Economic Record,* 88**,** 2-9.

COINDESK. 2014. *CoinDesk* [Online]. Available: http://www.coindesk.com/price/.

CONOVER, M., RATKIEWICZ, J., FRANCISCO, M., GONÇALVES, B., MENCZER, F. & FLAMMINI, A. Political polarization on twitter. ICWSM, 2011.

DAS, S., MARTÍNEZ-JEREZ, A. & TUFANO, P. 2005. eInformation: A clinical study of investor discussion and sentiment. *Financial Management,* 34**,** 103-137.

DAS, S. R. & CHEN, M. Y. 2007. Yahoo! for Amazon: Sentiment extraction from small talk on the web. *Management Science,* 53**,** 1375-1388.

DAVIDOV, D., TSUR, O. & RAPPOPORT, A. Enhanced sentiment learning using twitter hashtags and smileys. Proceedings of the 23rd International Conference on Computational Linguistics: Posters, 2010. Association for Computational Linguistics, 241-249.

DE LONG, J. B., SHLEIFER, A., SUMMERS, L. H. & WALDMANN, R. J. 1990. Noise trader risk in financial markets. *Journal of political Economy***,** 703-738.

EDELEN, R. M., MARCUS, A. J. & TEHRANIAN, H. 2010. Relative sentiment and stock returns. *Financial Analysts Journal***,** 20-32.

EDELMAN, B. 2012. Using internet data for economic research. *The Journal of Economic Perspectives***,** 189-206.

EDMANS, A., GARCIA, D. & NORLI, Ø. 2007. Sports sentiment and stock returns. *The Journal of Finance,* 62**,** 1967-1998.

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

FORBES. 2014. *Bitcoin's Mt. Gox Goes Offline, Loses $409M* [Online]. Available: http://www.forbes.com/sites/cameronkeng/2014/02/25/bitcoins-mt-gox-shuts-down-loses-409200000-dollars-recovery-steps-and-taking-your-tax-losses/.

FOX, J. 2011. *The myth of the rational market: a history of risk, reward, and delusion on Wall Street*, Harriman House Limited.

GALACTIC, V. 2013. Available: http://www.virgin.com/richard-branson/bitcoins-in-space.

GANDAL, N. & HALABURDA, H. 2014. Competition in the Cryptocurrency Market.

GINSBERG, J., MOHEBBI, M. H., PATEL, R. S., BRAMMER, L., SMOLINSKI, M. S. & BRILLIANT, L. 2008. Detecting influenza epidemics using search engine query data. *Nature,* 457**,** 1012-1014.

GLASER, F., HAFERKORN, M., WEBER, M. C. & ZIMMERMANN, K. 2014. How to Price a Digital Currency? Empirical Insights on the Influence of Media Coverage on the Bitcoin Bubble. *Empirical Insights on the Influence of Media Coverage on the Bitcoin Bubble (April 29, 2014). MKWI*.

GO, A., BHAYANI, R. & HUANG, L. 2009. Twitter sentiment classification using distant supervision. *CS224N Project Report, Stanford***,** 1-12.

GOMEZ, GONZALEZ, J. E., PARRA & POLANIA, J. A. 2014. Bitcoin: something seems to be jfundamentallykwrong.

GRAHAM, C. C. A. P. 2014. *Swiss, UK watchdogs step up scrutiny on forex traders* [Online]. reuters. Available: http://www.reuters.com/article/2014/03/31/us-swiss-forex-investigation-idUSBREA2U0EN20140331.

GRANELLO, D. H. & WHEATON, J. E. 2004. Online data collection: Strategies for research. *Journal of Counseling & Development,* 82**,** 387-393.

GRETZEL, U. & YOO, K. H. 2008. Use and impact of online travel reviews. *Information and communication technologies in tourism 2008***,** 35-46.

GU, B., KONANA, P., LIU, A., RAJAGOPALAN, B. & GHOSH, J. 2006. Predictive value of stock message board sentiments. *McCombs Research Paper No. IROM-11-06*.

HENNIG-THURAU, T., WIERTZ, C. & FELDHAUS, F. 2012. Exploring the "Twitter Effect:" An Investigation of the Impact of Microblogging Word of Mouth on Consumers' Early Adoption of New Products. *Available at SSRN*.

HILL, K. 2014. *SECRET MONEY: LIVING ON BITCOIN IN THE REAL WORLD*, Forbes Media.

HONG, Y. & SKIENA, S. The Wisdom of Bookies? Sentiment Analysis Versus. the NFL Point Spread. ICWSM, 2010.

HOWARD, P. N., DUFFY, A., FREELON, D., HUSSAIN, M., MARI, W. & MAZAID, M. 2011. Opening closed regimes: what was the role of social media during the Arab Spring?

HUANG, C. 2011. Facebook and Twitter key to Arab Spring uprisings: report. *The National. Abu Dhabi Media,* 6.

JANSEN, B. J., ZHANG, M., SOBEL, K. & CHOWDURY, A. 2009. Twitter power: Tweets as electronic word of mouth. *Journal of the American society for information science and technology,* 60**,** 2169-2188.

JONES, C. M., KAUL, G. & LIPSON, M. L. 1994. Transactions, volume, and volatility. *Review of Financial Studies,* 7**,** 631-651.

KEYNES, J. M. 1936. *General theory of employment, interest and money*, Atlantic Publishers & Dist.

KHONDKER, H. H. 2011. Role of the new media in the Arab Spring. *Globalizations,* 8**,** 675-679.

KRISTOUFEK, L. 2013. BitCoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era. *Scientific reports,* 3.

KWAK, H., LEE, C., PARK, H. & MOON, S. What is Twitter, a social network or a news media? Proceedings of the 19th international conference on World wide web, 2010. ACM, 591-600.

KYLE, A. S. 1985. Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society***,** 1315-1335.

LEE, C., SHLEIFER, A. & THALER, R. H. 1991. Investor sentiment and the closed-end fund puzzle. *The Journal of Finance,* 46**,** 75-109.

LIU, B. 2012. Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies,* 5**,** 1-167.

LIU, J., CAO, Y., LIN, C.-Y., HUANG, Y. & ZHOU, M. Low-Quality Product Review Detection in Opinion Summarization. EMNLP-CoNLL, 2007. 334-342.

LOTAN, G., GRAEFF, E., ANANNY, M., GAFFNEY, D. & PEARCE, I. 2011. The Arab Spring| the revolutions were tweeted: Information flows during the 2011 Tunisian and Egyptian revolutions. *International Journal of Communication,* 5**,** 31.

MEHRA, R. & PRESCOTT, E. C. 1985. The equity premium: A puzzle. *Journal of monetary Economics,* 15**,** 145-161.

MISHNE, G. & GLANCE, N. S. Predicting Movie Sales from Blogger Sentiment. AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs, 2006. 155-158.

MOORE, T. & CHRISTIN, N. 2013. Beware the middleman: Empirical analysis of Bitcoin-exchange risk. *Financial Cryptography and Data Security.* Springer.

NAKAMOTO, S. 2008. Bitcoin: A peer-to-peer electronic cash system. *Consulted,* 1**,** 2012.

NEWS, Y. 2013. *Twitter, by the numbers* [Online]. Available: https://news.yahoo.com/twitter-statistics-by-the-numbers-153151584.html.

NEWSWEEK. 2014. *Ex-J.P. Morgan Trader Joins Bitcoin Bulls Launching Hedge Funds* [Online]. Available: http://www.newsweek.com/ex-jp-morgan-trader-joins-bitcoin-bulls-launching-hedge-funds-258494.

O'CONNOR, B., BALASUBRAMANYAN, R., ROUTLEDGE, B. R. & SMITH, N. A. 2010. From tweets to polls: Linking text sentiment to public opinion time series. *ICWSM,* 11**,** 122-129.

OH, C. & SHENG, O. Investigating Predictive Power of Stock Micro Blog Sentiment in Forecasting Future Stock Price Directional Movement. ICIS, 2011.

OXFORDENGLISHDICTIONARY. 2014. *Oxford Dictionary - Sentiment* [Online]. Available: http://www.oxforddictionaries.com/definition/english/sentiment 2014].

PAK, A. & PAROUBEK, P. Twitter as a Corpus for Sentiment Analysis and Opinion Mining. LREC, 2010.

PANG, B. & LEE, L. 2008. Opinion mining and sentiment analysis. *Foundations and trends in information retrieval,* 2**,** 1-135.

PANG, B., LEE, L. & VAITHYANATHAN, S. Thumbs up?: sentiment classification using machine learning techniques. Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10, 2002. Association for Computational Linguistics, 79-86.

PELAT, C., TURBELIN, C., BAR-HEN, A., FLAHAULT, A. & VALLERON, A.-J. 2009. More diseases tracked by using Google Trends. *Emerging infectious diseases,* 15**,** 1327.

QUIGGIN, J. 2013. *The Bitcoin Bubble and a Bad Hypothesis* [Online]. Available: http://nationalinterest.org/commentary/the-bitcoin-bubble-bad-hypothesis-8353.

RESERVE, B. 2014. Available: http://bitcoinsreserve.com/about.

SABHERWAL, S., SARKAR, S. K. & ZHANG, Y. 2011. Do internet stock message boards influence trading? Evidence from heavily discussed stocks with no fundamental news. *Journal of Business Finance & Accounting,* 38**,** 1209-1237.

SADIKOV, E., PARAMESWARAN, A. G. & VENETIS, P. Blogs as Predictors of Movie Success. ICWSM, 2009.

SAUNDERS, M., LEWIS, P AND THORNHILL, A 2012. *Research Methods for Business Students, 6th edition*, Pearson.

SEBASTIANI, F. 2002. Machine learning in automated text categorization. *ACM computing surveys (CSUR),* 34**,** 1-47.

SEIFTER, A., SCHWARZWALDER, A., GEIS, K. & AUCOTT, J. 2010. The utility of "Google Trends" for epidemiological research: Lyme disease as an example. *Geospatial Health,* 4**,** 135-137.

SENTIMENT140. 2014. Available: http://help.sentiment140.com/for-students.

SHLEIFER, A. & VISHNY, R. W. 1997. The limits of arbitrage. *The Journal of Finance,* 52**,** 35-55.

SINHA, S., DYER, C., GIMPEL, K. & SMITH, N. A. 2013. Predicting the NFL using Twitter. *arXiv preprint arXiv:1310.6998*.

SPRENGER, T. O., TUMASJAN, A., SANDNER, P. G. & WELPE, I. M. 2013. Tweets and trades: The information content of stock microblogs. *European Financial Management*.

SUL, H., DENNIS, A. R. & YUAN, L. I. Trading on Twitter: The Financial Information Content of Emotion in Social Media. System Sciences (HICSS), 2014 47th Hawaii International Conference on, 2014. IEEE, 806-815.

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

TRENDS, G. Available: http://www.google.com/trends/.

TUMASJAN, A., SPRENGER, T. O., SANDNER, P. G. & WELPE, I. M. 2010. Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. *ICWSM,* 10**,** 178-185.

TWITTER. 2014a. Available: https://about.twitter.com/company.

TWITTER. 2014b. *Twitter Data Grant* [Online]. Available: https://blog.twitter.com/2014/twitter-datagrants-selections.

VINCENT, A. & ARMSTRONG, M. 2010. Predicting break-points in trading strategies with Twitter. *Social Science Research Network, Rochester, NY, SSRN Scholarly Paper ID,* 1685150.

VOCKLER, P. 2011. *Financial Reform: Unfinished Business* [Online]. Available: http://www.nybooks.com/articles/archives/2011/nov/24/financial-reform-unfinished-business/.

WANG, H., CAN, D., KAZEMZADEH, A., BAR, F. & NARAYANAN, S. A system for real-time twitter sentiment analysis of 2012 us presidential election cycle. Proceedings of the ACL 2012 System Demonstrations, 2012. Association for Computational Linguistics, 115-120.

WU, L. & BRYNJOLFSSON, E. 2013. The future of prediction: How Google searches foreshadow housing prices and sales. *Economics of Digitization.* University of Chicago Press.

ZHU, F. & ZHANG, X. 2010. Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of Marketing,* 74**,** 133-148.

Appendix

Appendix A – Introduction

Google trends and exchange rate for bitcoin

Figure A.1 Bitcoin search term as displayed in google trends service

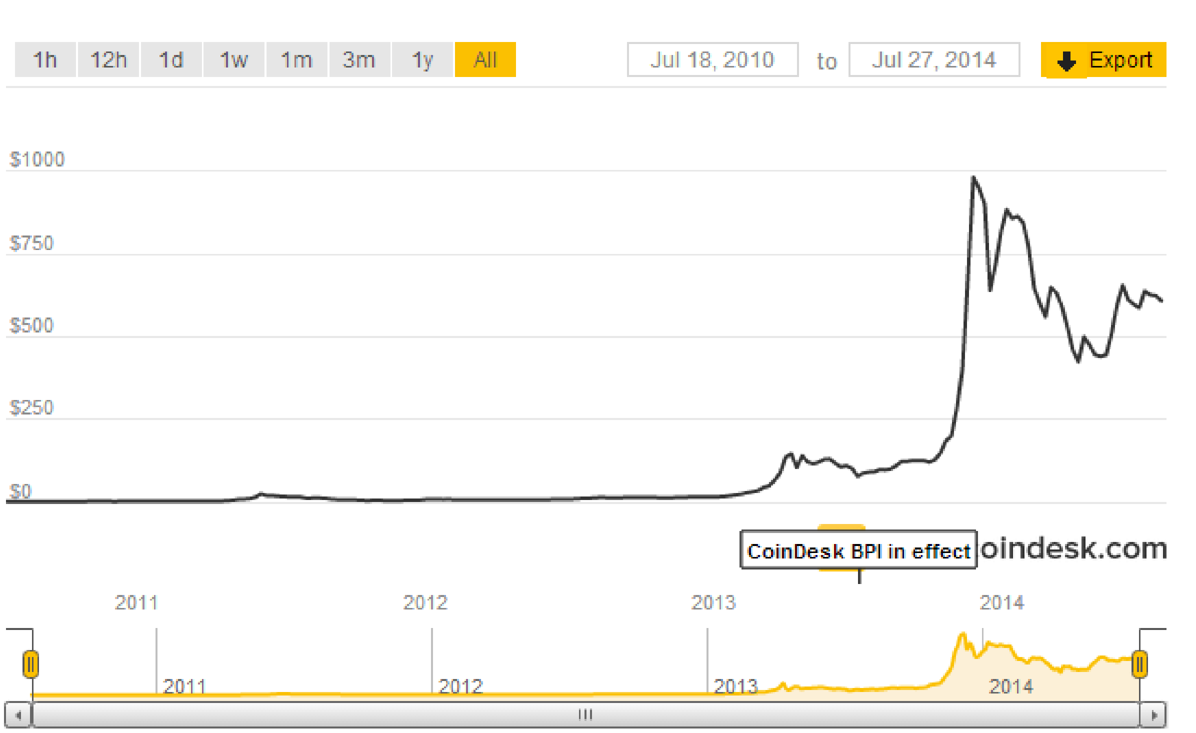


Figure A.2 Bitcoin Exchange rate in dollars

Appendix B – Methodology and Fieldwork

Confusion Matrix

===

Summary

Correctly Classified Instances : 114 80.2817%

Incorrectly Classified Instances : 28 19.7183%

Total Classified Instances : 142

===

Confusion Matrix

----

a	b	c	<--Classified as		
49	2	6	| 57	a	= Negative
6	24	5	| 35	b	= Neutral
6	3	41	| 50	c	= Positive

===

Statistics

Kappa 0.6443

Accuracy 80.2817%

Reliability 59.1341%

Reliability (standard deviation) 0.4012

Twitter Corpus confusion matrix

```
=========================================================
Summary
-------------------------------------------------------
Correctly Classified Instances      :     43     52.439%
Incorrectly Classified Instances    :     39     47.561%
Total Classified Instances          :     82
=========================================================
Confusion Matrix
-------------------------------------------------------
a     b     <--Classified as
24    16    |  40     a   = Negative
23    19    |  42     b   = Positive
=========================================================
Statistics
-------------------------------------------------------
Kappa                               0.086
Accuracy                            52.439%
Reliability                         35.0794%
Reliability (standard deviation)    0.3126
```

Bitcoin Price Index calculation

Information Taken from CoinDesk: How exactly is the BPI calculated?

The main features and criteria are as follows:

1. The CoinDesk BPI is a simple average of leading XBT/USD and XBT/CNY exchange prices.
2. The BPI is expressed as the midpoint of bid/ask spread.
3. The BPI is updated every 60 seconds.
4. If an exchange does not update its price for more than 30 minutes, it is omitted from the live BPI calculation until it is updated again.
5. New index historical data commences on 1 July 2013.
6. Prior index historical data is obtained via Mt. Gox.
7. End-of-day high, low, and closing BPI is based on Coordinated Universal Time (UTC).
8. Non-USD and non-CNY BPI prices are implied based on rates obtained via openexchangerates.org.
9. Any updates to the BPI criteria and formula shall occur as necessary.

Why is the BPI not volume-weighted?

The decision to apply a simple average, as opposed to a volume-weighted average, for the CoinDesk BPI was made because the bitcoin market currently lacks sufficient depth and regional liquidity.

Since trading volume now favours particular regions, a volume-weighted approach would not act as a proper global indicator, because each international bitcoin exchange is not equally available to all national trading participants.

A simple average does not favour a regional exchange with high volume and ensures that the BPI is meaningful for the largest number of market participants. Also, a simple average approach minimizes the impact of volume irregularities and accidentally excluding an exchange.

As overall liquidity improves and the number of global exchange choices increases, the impact of regional variances should diminish and a volume-weighted approach may become more appropriate.

CoinDesk Bitcoin Price Index API

CoinDesk provides a free API to access the current price of their BPI (Powered by CoinDesk). The service provides a simple JSON response that is easy to query

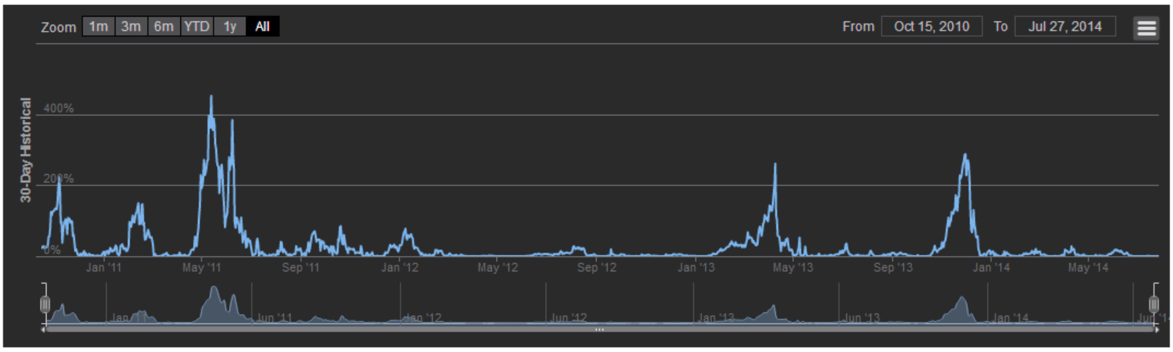
For example the url below can be queried for the

http://api.coindesk.com/v1/bpi/currentprice/USD.json

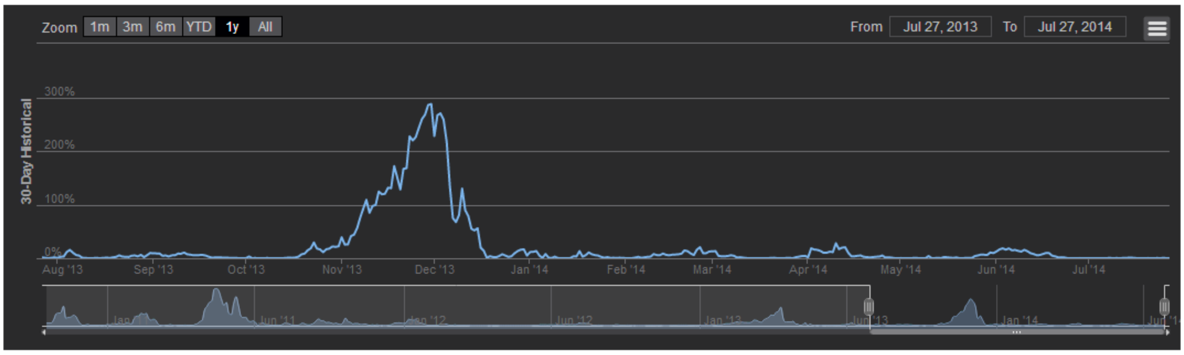
Example JSON response is:


```
{
   "time": {
      "updated": "Jul 28, 2014 08:41:00 UTC",
      "updatedISO": "2014-07-28T08:41:00+00:00",
      "updateduk": "Jul 28, 2014 at 09:41 BST"
   },
   "disclaimer": "This data was produced from the CoinDesk Bitcoin Price Index (USD). Non-USD currency data converted using hourly conversion rate from openexchangerates.org",
   "bpi": {
      "USD": {
         "code": "USD",
         "rate": "578.2025",
         "description": "United States Dollar",
         "rate_float": 578.2025
      }
   }
}
```

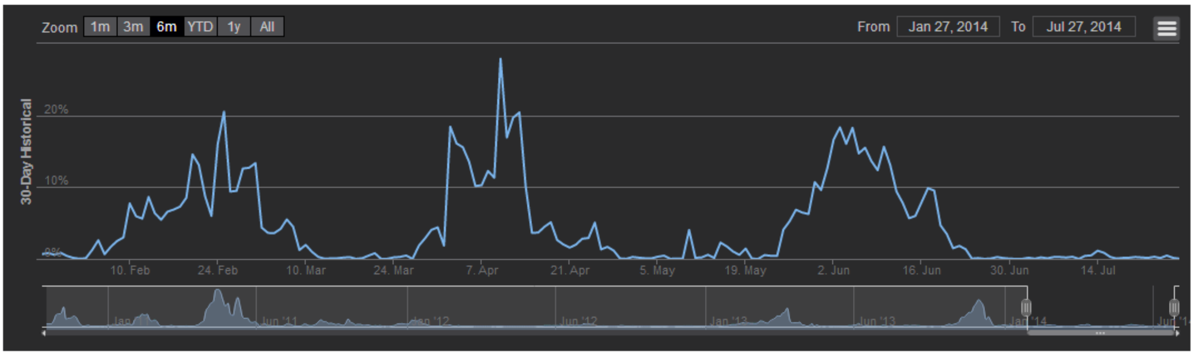
Price volatility of Bitcoin



All time - since late 2010



In the last year



In the last 6 months. Source http://www.coinometrics.com/bitcoin/vix

Appendix C – Findings and Analysis

Twitter Daily Volumes, Transaction Volume of Bitcoin and Daily Price Fluctuations

Date	No. of Tweets	Transaction Volume	Bitcoin Daily Price Fluctuations
28/07/2014	39116	64744	1.059311278
29/07/2014	34580	61607	0.425866699
30/07/2014	38383	66153	3.06252147
31/07/2014	30264	69761	3.008664529
01/08/2014	33852	67915	2.361744216
02/08/2014	29667	54989	1.309067688
03/08/2014	27787	53621	0.303087061
04/08/2014	35725	67812	0.213489095
05/08/2014	30283	72823	0.622060127
06/08/2014	37461	80402	0.012004596
07/08/2014	39054	69913	0.74780461
08/08/2014	39060	68297	0.532856656
09/08/2014	31180	60599	0.413188153
10/08/2014	29952	57913	0.231257121
11/08/2014	40038	75575	2.738145729
12/08/2014	38491	76982	0.889571087
13/08/2014	39453	75738	4.160433642
14/08/2014	40205	79082	6.614393007
15/08/2014	40645	73193	2.34588536
16/08/2014	37174	67974	4.673593492
17/08/2014	34020	60223	5.170921263

Weekday Correlations

		Weekday Tweets	Weekday Volumes	Price Fluctuation
Weekday Tweets	Pearson Correlation	1		
	Sig. (2-tailed)			
	N	15		
Weekday Volumes	Pearson Correlation	.294	1	
	Sig. (2-tailed)	.288		
	N	15	15	
Price Fluctuation	Pearson Correlation	.258	.348	1
	Sig. (2-tailed)	.353	.204	
	N	15	15	15

Weekend Price Correlations

Shifted data with lag applied.

8 hour

Bitcoin Price Change	Bullishness
-2.78	-0.05
-3.9	0.17
3.78	0.01
-2.16	-0.16
-5.62	0.06
-7.98	0.11
-0.73	-0.05
18.83	0.37
0.63	0.31
-0.37	0.22
16.5	-0.03
-6.79	-0.16
-5.31	0.02
-4.78	-0.26
2.28	0.15
-8.68	-0.11
8.96	0.04
-2.56	-0.12
-0.44	-0.26
2.41	-0.15
-2.39	-0.22
-1.4	0.12
-0.41	-0.03
0.04	0
-0.7	-0.07
0.99	-0.12
-0.08	-0.06
0.01	0.16
3.64	0.27
0.6	0.11
5.82	0.21

-1.8	0.1
-0.85	-0.26
-3.39	-0.37
0.46	-0.43
0.45	0.31
-1.59	0.14
3.41	0.17
-0.37	0.04
-1.79	0.05
-2	0.17
-12.54	0.21
-2.83	-0.15
-3.74	-0.01
1.43	0.18
-14.3	-0.01
-16.92	-0.17
7.53	0.01
-31.53	-0.15
-9.97	0.14
4.94	0.43
6.48	0.11
-12.57	0.34
-4.8	0.3
-6.83	0.29
13.49	0.23
16.07	0.06
-12.17	-0.1
-14.03	-0.2
-0.26	-0.18
-5.65	-0.21
-22.82	-0.22
-4.34	-0.27

24 hour time shifted data

Bitcoin Price Change	Aggregate Sentiment Retweets
-2.9	281
-15.76	-343
18.72	1503
9.34	-92
-7.82	22
-2.27	-1
-0.42	-553
-1.78	66
0.22	-118
4.25	1231
3.17	-1028
-2.49	-306
1.45	404
-16.33	980
-5.14	57
-23.69	-451
-36.57	-110
-10.9	346
22.74	647
-26.46	-407
-32.81	-582

Next day predictions based on correlations

Bitcoin Price Change	Aggregate Sentiment Retweets	Movement	RESULT
-2.9	281	UP	Incorrect

-15.76	-343	Down	Correct
18.72	1503	UP	Correct
9.34	-92	DOWN	Incorrect
-7.82	22	UP	Incorrect
-2.27	-1	DOWN	Correct
-0.42	-553	DOWN	Correct
-1.78	66	UP	Incorrect
0.22	-118	DOWN	Incorrect
4.25	1231	UP	Correct
3.17	-1028	DOWN	Incorrect
-2.49	-306	DOWN	Correct
1.45	404	UP	Correct
-16.33	980	UP	Incorrect
-5.14	57	UP	Incorrect
-23.69	-451	DOWN	Correct
-36.57	-110	DOWN	Correct
-10.9	346	UP	Incorrect
22.74	647	UP	Correct
-26.46	-407	DOWN	Correct
-32.81	-582	DOWN	Correct

Analysis of 1 hour aggregation

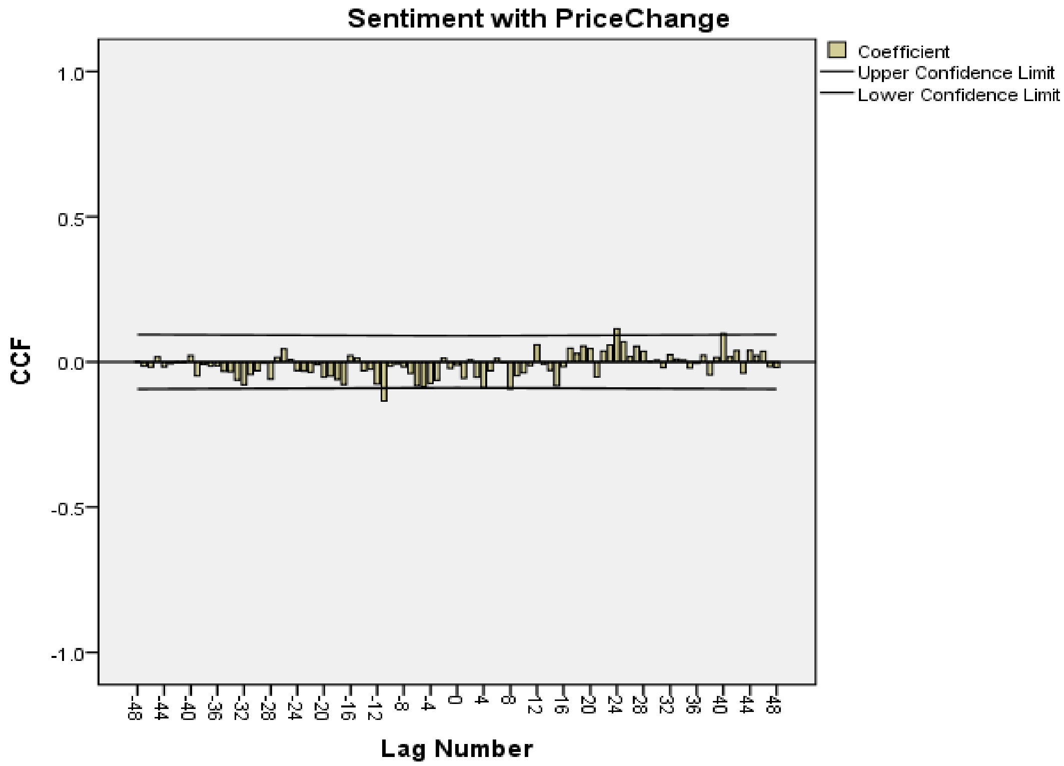
Cross correlation lags applied of 48 hrs

Cross Correlations

Series Pair: Sentiment with PriceChange

Lag	Cross Correlation	Std. Error[a]
-11	-.135	.045
-10	-.015	.045
-9	-.006	.045
-8	-.017	.045
-7	-.039	.045
-6	-.083	.045
-5	-.086	.045
-4	-.073	.045
-3	-.063	.045
-2	.013	.045
-1	-.023	.045
0	-.011	.045
12	.058	.045
19	.055	.046
23	.057	.046
24	.114	.046
25	.069	.046
26	.019	.046
27	.053	.046
40	.099	.047

Only significant values shown.



Of note the most signification cross correlation is at 24 hour point. When lag applied as below it is not significant but is positive.

Correlation with lag of 24 applied.

Correlations

		Sentiment	Bitcoin Price Lag Applied
Sentiment	Pearson Correlation	1	
	Sig. (2-tailed)		
	N	501	
Bitcoin Price Lag Applied	Pearson Correlation	.060	1
	Sig. (2-tailed)	.179	
	N	501	502

Table C.2. Correlation for 1 hour aggregation with prices moved 24 hour.