Effect of Sentiment on Bitcoin Price Formation

Brian Perry-Carrera

Professor Grace Kim, Faculty Advisor

Brian graduated with Distinction in Economics with a concentration in Finance and a certificate in Innovation and Entrepreneurship. He can be contacted at bperrycarrera@gmail.com. Presently, Brian works as an analyst at Bourne Partners, an Investment Bank in Charlotte, NC.

Acknowledgements

I would like to express an enormous amount of gratitude to Professor Grace Kim. Her willingness to help students is unrivaled. Her guidance and feedback throughout the process of writing this paper was pivotal to the completion and success of this paper. She truly gives her all to her students, and for that, I am extremely thankful.

Abstract

With the recent growth in the investment of cryptocurrencies, such as bitcoin, it has become increasingly relevant to understand what drives price formation. Given that investment in bitcoin is greatly determined by speculation, this paper seeks to find the econometric relationship between public sentiment and the price of bitcoin. After scraping over 500,000 tweets related to bitcoin, sentiment analysis was performed for each tweet and then aggregated for each day between December 1st, 2017 and December 31st, 2017. This study found that both gold futures and market volatility are negatively related to the price of bitcoin, while sentiment demonstrates a positive relationship.

JEL classification: G12; G41; Z00

Keywords: Asset Pricing; Bitcoin; Sentiment

1 Introduction

The rise in the value of bitcoin has resulted in the growth of a new asset: cryptocurrencies. According to Coinmarketcap.com, there are 1569 total cryptocurrencies (Coinmarketcap, 2018). The price of various coins range from $0.000000223 to $578,324 per token with a current total market capitalization of $326 billion (Coinmarketcap, 2018). Although the market capitalization is still an order of magnitude less than that of the S&P 500, the cryptocurrency market is a relatively new space, particularly when thinking about the longevity of the financial markets. Given the amount of known information with regards to traditional financial markets, we have many different pricing models that have been proven over time such as CAPM, APT, and the Black-Scholes Option pricing model. However, when we begin to examine the price formation for cryptocurrencies, we find that there has not been a comparable model yet developed. Further, attempting to use traditional asset pricing models to determine the price formation of cryptocurrencies results in inaccuracies because of the many differences between cryptocurrencies as an asset class and traditional assets. One of the main differences is the importance of public/investor sentiment with regards to each asset respectively. Due to the unproven nature of cryptocurrencies, the value of each coin rests much more on sentiment in comparison to the impact of sentiment on traditional asset valuation. This paper seeks to bridge the gap between traditional asset price formation and behavioral asset price formation to begin to understand the intricacies of bitcoin price formation, with a particular focus on the impact of public sentiment.

1.1 Background

2017 was the year of Bitcoin. The price rose from $996.34 to nearly $13,500 between January 1st, 2017 and December 31st, 2017 (Coinbase, 2018). Bitcoin is a peer-to-peer payment system developed by an individual that used the pseudonym Satoshi Nakamoto (Bitcoin.com, 2017). Bitcoin uses blockchain as a ledger that is cryptographically secured meaning the last line of the previous block appears as the first line of the next creating a chain of blocks. This ledger is distributed to all users and is immutable so that one cannot alter the history of the blocks to change previous transactions. Bitcoin uses SHA-256 as a hashing algorithm in order to create a secure network of blocks (Coindesk, 2017). Hashing is notably different than encryption because encryption implies that there can be decryption of the message, whereas reversing the hashing process is essentially impossible given current computing power. Currently blocks are broken up by transactions occurring during 10 minute periods because that is the required time to verify transactions. Transactions are verified by bitcoin "miners" that use specialized computers to solve the hash and subsequently secure the network. In return for the effort, miners receive awards for verifying the transaction in the form of partial bitcoins. When Nakamoto created Bitcoin he set the finite supply of bitcoin to 21 million where miners are rewarded from the supply pool, thus increasing the total supply (Coindesk, 2017). However, as time goes on and transactions continue, there will be less bitcoins available as a reward for miners, thus lowering the incentives to mine. Meanwhile the computing power required will become increasingly difficult to confirm transactions on the blockchain. Currently there are around 17 million bitcoins in circulation as of April 10th, 2018 (Blockchain.info, 2017). One additional bit of information to note regarding the supply of bitcoin is that some bitcoins have been stolen and lost. The most notable was the crash of Mt. Gox, one of the first bitcoin exchanges that lost 850,000 Bitcoins, which would have a current value of over $5 billion dollars at today's price (McMillan, 2017). Although some of the coins were recouped, the majority of lost coins are unable to be recirculated. This idea demonstrates the negative side of the anonymity of bitcoin. Although investigators can see the public key of the holders of the stolen coin, as long as those coins do not move into other accounts, there is no way of putting an identity to the public key.

Although Bitcoin creates numerous advantages, the technology does have some issues that will need to be worked out before it can become mainstream. The main current issue with Bitcoin is the matter of scalability. Bitcoin and the supporting blockchain can support approximately 3 to 4 transactions per second, which may initially sound sufficient, but when compared to the nearly 200 transactions per second average of PayPal and the 1,500+ transactions per second average performed by Visa, the issue of scalability becomes illuminated (Altcoin Today, 2017). Despite the limitation on the speed of transactions, at its peak in December 2017, Bitcoin had over around 1 million unique addresses used on the Bitcoin blockchain, yet now in April 2018 there are around 400,000 (Blockchain.info, 2017). Given that the Bitcoin blockchain is open source, many developers are currently working on various solutions to the issue of scalability (Coindesk, 2017). One potential solution to the issue of scalability was the Segwit2x hard fork which would have increased the base block size to 2MB (from 1MB), but the hard fork would have decreased miners' revenue and therefore they abandoned the hard fork (Parker, 2017). In addition to the decrease in revenue for miners, Segwit 2x would have taken some transactions off the blockchain, which gives rise to concerns about fraud. It is evident that the problem of scalability, especially when considering security, is still an issue today.

It is also important to recognize the types of transactions that Bitcoin is used for. Although bitcoin's price and market capitalization are at an all-time high, the market is mostly speculative. A small fraction of the market cap rate uses bitcoin to pay for goods and services, but mostly there is not enough surrounding infrastructure to facilitate bitcoin as a form of payment. The influx of recent investment is likely due to investors beginning to understand the future potential as a form of payment, but the current investment represents more of a speculative investment opportunity.

Although Bitcoin was the first mainstream blockchain other developers have begun to develop other systems that utilize a blockchain in order to improve efficiency, remove central authority, and eliminate the double spending problem. The blockchain with the second largest market capitalization, Ethereum, allows for developers to build on top of the Ethereum network to create decentralized apps that involve smart contracts (Coin Market Cap, 2017). The smart contracts are executed using Ethereum token, Ether, in order to facilitate the exchange of goods or services between parties where one party receives the token as payment immediately as the good or service has been confirmed received by the blockchain network. Smart contracts eliminate counter-party risk and eliminate the need for a third party settlement agency like the Automated Clearing House. Although Bitcoin is the first application of blockchain technology, there has been enormous growth in the startup space dedicated to using blockchains in new ways, which has the potential to disrupt numerous industries where blockchain can improve efficiency by creating a peer-to-peer network.

1.2 Bitcoin as an Asset

Bitcoin adds benefits in several areas: verification of ownership, and removal of intermediaries which improves efficiency and limits fraud. Given the recent growth in the valuation, an important question to ask is, what factors are driving the growth? Although some literature has been written about correlations between equities and bitcoin, there is little research on other factors that affect the valuation of bitcoin. In this paper, I will discuss the role that social media plays in the price of bitcoin. This research will give an important insight on the relationship between bitcoin and a variable that has yet to be econometrically examined. If a significant result is proven, there would be important implications for future investment strategies regarding bitcoin and potentially other cryptocurrencies as well. Given the speculative nature of bitcoin, much of the investment in bitcoin is based on sentiment rather than valuation of the asset's tangible value. Because of this investment behavior, it is particularly important to examine public sentiment.

Although there are several ways to examine public sentiment, this paper utilizes Twitter data in combination with Google Trends for several reasons. First, up until the end of 2017 there was no simple way to trade bitcoin on an institutional level, meaning that the investment opportunity was flooded with average investors. Given the unique type of investors, public tweets represent a larger majority of sentiment than it would for a typical asset.

Several papers either examine bitcoin price formation with factors such as sentiment or typically correlated financial markets, but few to none incorporate both factors**.** Bitcoin is similar to equities in the sense that both assets can be bought and sold through exchanges, and the price is determined by many of the same factors including the sentiment of market participants. However, a crucial difference is that the value of bitcoin is derived based on the value that we believe it has. This notion is in contrast to equities because the value of equities is related to the performance of a specific company relative to past performance as well as their competitors. Gold also has some attributes that mirror bitcoin, but has aspects that are different. The two are similar in the sense that, although gold has some intrinsic value because it is a precious metal and has value in the use of computer electronics, the majority of investment in gold is due to the fact that investors view gold as a safe-haven asset. Similarly, bitcoin has no intrinsic value besides the value we believe it to have. Additionally, for those invested in gold or gold futures, the investor likely never takes hold of tangible gold, but rather, like bitcoin, an account says that an investor has a certain amount of gold and we trust this information to be true. However, the largest difference between gold and bitcoin is the level of risk. Investors believe that gold will likely have value into perpetuity, but the future value of bitcoin is extremely uncertain. Additionally, given that the value of gold into perpetuity would remain constant, the commodity is also protected against inflation. These similarities and differences of bitcoin and equities and gold give rise to the question: why invest in bitcoin and not another, less risky asset? In order to answer this question, we must think about the types of people investing in bitcoin.

Broadly speaking, there are two types of people who invest in bitcoin: those who believe the technology underlying bitcoin has value and those who like the riskiness of the asset. The former group understand the power of blockchain technology and see the value of a system that is immutable and decreases risk during the exchange of information between a party and a counter-party. The latter group of people are those who have seen the enormous potential return on investment in bitcoin as an asset. In the context of this study, it is also important to recognize whose sentiment is represented through the examination of public tweets relating to bitcoin. It seems likely that the sentiment determined through tweet analysis represents the opinions of people who are evaluating the value of the technology, and that the sentiment variable created in this study does not reflect the opinions of those who simply see bitcoin as an alternative financial asset. Given that the sentiment analyzed in this study does not reflect the opinions of those investing in bitcoin as an alternative asset, it should be noted as a limitation that the sentiment variable used does not reflect 100% of the sentiment towards bitcoin.

Section 2 offers literature review that comments on prior relevant studies, section 3 follows with the theoretical framework that provides an understanding of the theoretical approach to this study, while section 4 gives a clear picture of what and how the data was captured for this study. Section 5 describes the empirical specification used for the econometric approach to understanding bitcoin price formation with section 6 following, expressing the findings of the econometric tests and regressions. The final section, section 7, summarizes the findings and gives a broad conclusion of the effect of sentiment on bitcoin price formation and potential future implications.

2 Literature Review

2.1 Papers on Price Formation

There are several different ways to approach modeling price formation of assets. In Saban Celik's paper, *Theoretical and Empirical Review of Asset Pricing Models: A Structural Synthesis*, he gives a detailed picture of the various methods of asset pricing (Celik, 2012). He describes the two main perspectives with regard to asset pricing: neoclassical based asset pricing and behavioral based asset pricing. Although most assets are priced by either of these methods, price formation of bitcoin does not fit into solely one of the aforementioned methods of asset pricing. Given the importance of sentiment towards bitcoin with regard to investment, bitcoin pricing should be modelled through the inclusion of both neoclassical and behavioral based asset pricing methods.

One study examined bitcoin price formation following the neoclassical pricing model. The researchers examined the relationship between bitcoin and supply and demand fundamentals, macroeconomic trends, as well as the attractiveness to investors (Ciaian et al., 2014). They found that both supply and demand fundamentals have a significant impact on the price of bitcoin as well as the attractiveness to investors; however, they did not find a significant result between macroeconomic indicators, such as the Dow Jones Index and the price of oil, and the price of bitcoin (Ciaian et al., 2014). Additionally they found that "the size of the BitCoin economy and the velocity of BitCoin circulation, have the strongest impact on price" (Ciaian et al., 2014). However, this study simply used Wikipedia views as a measure of public sentiment. Given their results, it seems reasonable to suggest that adding an additional variable capturing public sentiment through a more nuanced approach would give additional insights into the power of sentiment.

In another paper examining the determinants of bitcoin price, the authors describe how the level of demand for bitcoin takes into account the same supply and demand fundamentals as other currencies, commodities, and equities, but additionally recognize that the demand for bitcoin is related to the expectation of future price movements (Georgoula et al., 2015). They indicate the unique nature of bitcoin price formation by describing how, in additional to traditional pricing fundamentals, investment in bitcoin is also based on the public feeling of future price expectations (Georgoula et al., 2015). This distinct price formation further indicates the need to examine bitcoin price formation in a way that reflects the notion that price is determined by some classical pricing fundamentals, and some behavioral pricing theory.

2.2 Papers on Sentiment and Financial Markets

Previous studies conclusively show that social media sentiment analysis can predict future returns (Sul et al., 2016; Nofer & Hinz, 2015). However, each study has certain limitations. Nofer and Hinz's study did not find a significant relationship between "Social Mood" and the stock market, but they did find a significant relationship between "Follower-Weighted Social Mood" and the stock market. They found that an increase in the follower weighted social mood levels by 1 implies an increase of 3.3 basis points in the DAX (Nofer & Hinz, 2015). However, it is notable that this study focused on German tweets and German equities, so it is unclear if a similar relationship occurs in the U.S.. One interesting analysis provided in their paper was how they categorized sentiment. Most of the other studies in the field simply categorize the text into positive, neutral, or negative, but their study created five mood categories: grief, hopelessness, tiredness, anger, and positive mood (Nofer & Hinz, 2015). In one study attempting to use social media sentiment to predict stock returns, the study found a unique result that provides an interesting insight into the effect of a user's sentiment influence on future returns (Sul et al., 2016). They found that the speed of sentiment diffusion effects future returns, meaning that tweets sent by people with fewer than the median number of followers and that are not retweeted take longer to have an impact on future stock prices (Sul et al., 2016). Further, their result indicated that using sentiment from tweets from users with less than the median number of followers had an impact on future returns 10 and 20 days later yielding an annualized return of 15.65% and 11.41% respectively, while sentiment analysis of tweets by users with many followers had no impact on future returns (Sul et al., 2016). Their paper analyzed sentiment with a daily constraint rather than on an individual Tweet basis (Sul et al., 2016). A daily outlook on sentiment may fit with the equity market because the financial market has daily hours of trading, while this outlook may create a limitation when examining the bitcoin market because the bitcoin market is open 24/7.

One of the main differences between the various studies is how sentiment is calculated. Most use machine learning algorithms to determine the sentiment, but various programs are used and in some cases created. This difference is important because the very basis of determining the relationship is reliant on successful categorization of sentiment. One study used a sentiment analyzing program to analyze each word in the Tweet and subsequently gave an aggregate sentiment score to each Tweet (Mai et al., 2016). An interesting finding from their study was that they found that sentiment from bitcoin forums had a stronger impact than sentiment derived from tweets (Mai et al., 2016). The paper described a limitation of their findings based on their sentiment analysis given that they used a financial sentiment dictionary when categorizing their Twitter data (Mai et al., 2016).

2.3 Papers on Sentiment and Bitcoin

There has been a good deal of research performed on the relationship between Twitter data and stock market returns; however, the work regarding Twitter data and cryptocurrency returns is much less developed. However, among the current literature, there does seem to be a common theme that there is some level of a significant relationship between Twitter sentiment and bitcoin price. This theme validates the further examination at this relationship. One peer paper found no significant relationship between sentiment and bitcoin price below the 10% level (Xie et al., 2017); however, it is also notable that this sentiment analyzed in this paper was from particular Bitcoin forum data. This paper utilizes Twitter data to represent public sentiment, which is notably different than the sentiment found on Bitcoin forums. It seems reasonable to assume that there is a divide in the type of user that posts on Twitter versus a Bitcoin forum. Intuitively, Twitter data may encompass more of the lay-man's sentiment regarding Bitcoin, while more active members of the Bitcoin community are likely to post discussions on Bitcoin forums.

An additional paper examined Twitter sentiment in order to create a predictive trading algorithm for bitcoin (Colianni et al., 2017). The model they created achieved a day to day accuracy of over 90% (Colianni et al., 2017). There are several notable remarks based on their research. First, their data collection is extremely large, relative to the data collected in other papers. They collected over 1 million tweets that took over 21 days to collect. It seems likely that their high predictive accuracy is in part due to the size of their data. In addition, they noted that over half of the tweets collected were posted by users who only posted once about Bitcoin (Colianni et al., 2017). This finding demonstrates the importance of individual tweet collection as opposed to focusing on individual users with a larger influence. Given their finding, this study will collect tweets based on keywords, rather than tweets from particular identified users that frequently post about bitcoin.

One notable finding from examining previous papers is time period selection. Some of the relevant studies of Twitter sentiment and bitcoin have seemingly arbitrary time periods. These studies chose their time periods based on collecting enough data, rather than specifying a time period based on significant events or growth in the bitcoin market. One study collected data from 2013 to 2016 based on their claim that this period was when Bitcoin gained a more mainstream popularity and media attention (Kim et al., 2016). An additional paper examining the relationship between sentiment and options returns collected data between July 2009 and September 2012 (Houlihan & Creamer, 2017); however, this time period seems relevant for financial markets given this period followed the mortgage crisis, but it does not suggest that this same period is most relevant in the examination of bitcoin pricing and valuation. Based on the previous studies, careful consideration as to the time period selection for this paper was taken (described in detail in the data section).

Prior studies evaluate sentiment using several different tools, but they all essentially create sentiment variables based on the same approach. Many studies used sentiment tools with a financial lens such as “Aktuelle Stimmungsskala” and “StockTwits” (Hinz & Benlian, 2014; Houlihan & Creamer, 2017). These tools use a word classification specifically designed for financial context in order to determine sentiment. Although these tools may be appropriate when analyzing sentiment with regard to financial markets, they are not the most appropriate given that Bitcoin is far from a traditional financial asset. Additionally, a finance dictionary found in an article in *The Journal of Finance* titled: “Measuring Readability in Financial Decisions” was used by two studies focusing on the effect of social media sentiment on bitcoin pricing (Loughran & McDonald, 2015). The method of using word classifying dictionaries all function in the same way. Each word has a sentiment value and then the whole text receives a sentiment value based on the aggregate sentiment score of each word.

The most common method among the peer research of determining sentiment is VADER sentiment analysis (Valence Aware Dictionary and sEntiment Reasoner). In addition, the studies that used the VADER sentiment analysis found that VADER outperformed other sentiment tools in terms of polarity classification and consistency compared to eleven other sentiment tools (Stenqvist & Lonno, 2017). Further, VADER utilizes three different lexicons, one of which is the

Linguistic Inquiry and Word Count (LIWC), which was used as the basis of sentiment in another study (Mai et al., 2016). Lastly, while some studies used other sentiment tools as described above, VADER is widely regarded as one of the best sentiment analysis tools when it comes to analyzing social media and is therefore a better fit for this paper than other tools might be (Stenqvist & Lonno, 2017).

The common variables collected among the peer papers are: bitcoin price, tweet text, and tweet sentiment. Although some studies performed additional analysis by capturing variables such as bitcoin transactions, and influence of the twitter/forum post, my paper seeks to examine a more comprehensive group of factors that may influence price formation through the examination of a larger, more recent, dataset than prior papers that spans a particularly interesting period of price movement of bitcoin.

3 Theoretical Framework

In Saban Celik's paper *Theoretical and Empirical Review of Asset Pricing Models: A Structural Synthesis*, he presents a clear outline of the two types of asset pricing methods: the neo-classical based asset pricing and the behavioral based asset pricing (Celik, 2012). The divide between the two methods of asset pricing can be seen below in Figure 1**,** drawn from Celik's paper.

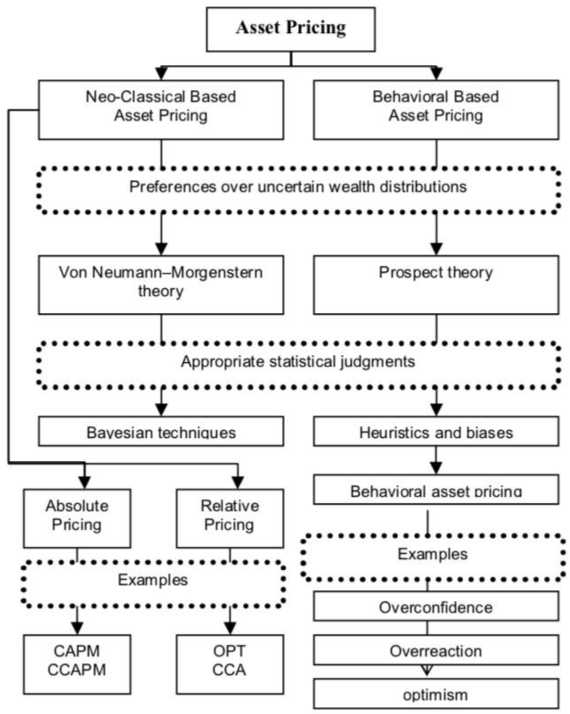

Figure 1

He further describes that the initial distinction between the two pricing methods is based on how individuals asses the relevant factors when considering investment in a risky asset (Celik, 2012). The neo-classical framework gives way to a more statistical approach based on Von Neumann-Morgenstern theory and Bayesian techniques, while individuals who utilize a behavioral method focus on how people actually behave (Celik, 2012). Some financial economists adamantly believe that absolute and relative pricing, under the neoclassical asset pricing framework, are more accurate because they take into account more information such as macroeconomic risk and the pricing of other assets; however, Celik defends the behavioral based approach by describing that psychologists working in behavioral finance have found that individuals do not actually make decisions as if they had Von Neumann-Morgenstein preferences, but rather they follow prospect theory (Celik, 2012). Prospect theory describes how individuals are more likely to choose a decision with known risk over an option where the outcome is unknown (Celik, 2012).

When thinking about the price formation of other assets such as equities and commodities, there are several similarities to bitcoin price formation. Although the capital asset pricing model (CAPM) would not accurately forecast the expected return of bitcoin, some of the variables used in calculating the return of a typical asset also contribute to bitcoin's price formation. If we look at the CAPM equation where: $r_a = r_f + \beta_a(r_m - r_f)$, the calculated return of a typical asset is based on rf, the risk-free rate, the beta of the asset (the volatility of the asset relative to the market as a whole), and the excess market return. It seems reasonable to assume that although bitcoin differs from traditional assets in many ways, the risk-free rate would still have an impact on bitcoin's price formation. Calculating the beta of bitcoin would not make sense if calculated relative to financial market riskiness; however, if the beta could be calculated based on the riskiness of bitcoin relative to the riskiness of the entire cryptocurrency asset class, it could be incorporated into bitcoin's price formation. However, based on the scope of this paper, a beta calculation of bitcoin will not be performed and instead, other measures of relative risk will be used.

Further, there are problems with using other relative pricing models like the Black-Scholes Model to model bitcoin price formation. Several fundamental assumptions underlying the Black-Scholes model are violated when thinking about bitcoin. The Black-Scholes model assumes that there is no transaction cost in buying an option; however, processing bitcoin transactions requires far more effort from miners and therefore is more costly to exchanges. In turn, this results in exchanges charging a transaction cost to the purchaser. In addition, the Black-Scholes model assumes that the volatility of the underlying asset is known and constant, but this is certainly not the case with bitcoin.

One of the key attributes of price formation derived from prospect theory is that utility is derived from both consumption levels *and* changes in the consumer's financial wealth (Barberis et al., 1999). Prospect theory also suggests that utility is a function of previous gains and losses (Barberis et al., 1999). This idea is important in the context of bitcoin because of the riskiness of the asset. Investor habits are altered based on the performance of their previous investments in cryptocurrencies. In a sense, this idea reflects the importance of sentiment. Sentiment is a result of the utility of the investor, and utility is a reflection on prior investment, and therefore suggests prior sentiment determines future demand.

After examining the various methods of typical asset price formation, it becomes clear that there is no current method to accurately model bitcoin price formation. Bitcoin's price formation is unique compared to typical financial assets. The price of bitcoin takes into account many of the attributes that the neoclassical asset pricing model uses, but is also affected by behavioral aspects that would be found in a behavioral asset pricing model. Therefore it would be inaccurate to examine bitcoin price formation through the lens of either the neoclassical pricing models or behavioral pricing models. Instead, a hybrid between them must be applied. This paper seeks to incorporate a hybrid empirical approach to develop an accurate pricing method for understanding bitcoin price formation.

Given the short lifespan of bitcoin relative to other financial assets, investor behaviors are heavily reliant on their own sentiment as well as the sentiment of the cryptocurrency market. Although there is some intrinsic value behind bitcoin's technology, much of the recent investment has been speculative and therefore demonstrates the power of sentiment. Other financial assets have years of historical data that give a sense of stability to the market. Although there are periods with higher volatility than others, overall the value of these assets fluctuate far less than bitcoin. Bitcoin is just over ten years old and only in the last two years has it gained shown a significant increase in value. There are two main factors that cause investment to be based more on speculation than a typical financial asset. First is the age of bitcoin. Investors are unsure of the true value and therefore implied price of bitcoin because there are few to no comparable assets. As bitcoin ages and the cryptocurrency market becomes more mature, there will be an increase in stability and therefore decrease in the importance of sentiment. Additionally, bitcoin is still a novelty. Although there is intrinsic value in the technology supporting bitcoin, it is unclear of the future need and value of bitcoin as a currency itself. Further, the technology itself is new and still underdeveloped. There is potential for bitcoin to drastically alter the foundation of exchange across the globe, but at the same time, many problems still remain unsolved that will likely seal the fate of bitcoin. Given this high level of uncertainty, people naturally feel comfort in like thought and therefore rely more on the opinions of others i.e. public sentiment.

4 Data

For the purpose of this paper, several data sources have been used to create a unique dataset capturing over 500,000 tweets and their respective sentiment, as well as five additional variables relating to the price formation of bitcoin. For the purpose of this study, data regarding Bitcoin, tweets from Twitter, and applicable financial information were required. In order to capture meaningful data, the data selected spans December 1st, 2017 to January 1st, 2018 on a daily basis. This time frame was selected because it was a particularly interesting month with regard to bitcoin volatility and sentiment. The time period selected was also in part due to the sheer number of tweets recorded in a month. Scraping tweets for a period longer than one month would require millions of tweets to be collected, which was infeasible given the timeframe for this study. Prior to December 1st Bitcoin had been on an upward trajectory. On November 28th, 2017 bitcoin broke $10,000 for the first time in its history. (99 Bitcoins, 2017). As the price of bitcoin began to rise, on December 11th, CBOE begins to offer bitcoin futures for the first time. This change was monumental for bitcoin because it was the first time any cryptocurrency could be exchanged on an institutional level. Additionally, the creation of bitcoin futures meant that investors could short sell the coin without ever holding the asset for the first time. Bitcoin received an enormous amount of buzz through media outlets and the price continued to rise until it hit its all-time peak just below $20,000 on December 18th. Following the all-time high, there were a wide range of opinions regarding the future direction of bitcoin. One managing director of a cryptocurrency trading firm, Dave Chapman, said he believed bitcoin could hit $100,000 by the end of 2018, while one Business Insider article wrote on December 22nd, 2017 that bitcoin could trigger the next financial crisis (Murphy, 2017; Hodgson, 2017). The price of bitcoin continued to fall with only minor rebounds. On December 28th, South Korea, one of the largest locations of investment, threatened to shut down all cryptocurrency trading in the country (99 Bitcoins, 2017). By January 1st, 2018, bitcoin had fallen to around $13,000 with many unsure the future direction of bitcoin in 2018. The time period selected encompasses both a large price increase to its all-time high, as well as a huge sell off that resulted in a 28% loss over a span of two weeks (Coinbase, 2018).

4.1 Bitcoin Data

For the data regarding bitcoin, data was collected from Blockchain.info, one of the most respected Bitcoin data sources commonly used in similar research. The variables captured are: price, market capitalization, trade volume, difficulty of mining, transaction fees paid to the miners, cost per transaction, miners revenue, output volume, the number of transactions, the number of unique addresses used on the bitcoin blockchain, the number of bitcoin blockchain wallets created, and the total number of bitcoins that have been mined, all collected on a daily basis for the time period selected (December 2017). Additionally, a new variable was created from recoding bitcoin price to capture the natural logarithm of bitcoin price in order to limit the power of outliers. Blockchain.info allowed for data capture without many limitations. The Bitcoin related variables were downloaded from Blockchain.info by exporting their chart data for the selected time period on a daily basis to a .csv file. Then the various variables were compiled into one file in order to capture all of the variables of interest and allow for ease of use in the regression process. It should be noted that it would be too difficult to capture a beta for bitcoin because it is unsuitable to compare bitcoin's risk relative to a typical financial market such as the S&P 500, and it would be extremely difficult to capture the risk of the overall cryptocurrency market.

4.2 Twitter Data

Twitter data was collected through a Python based web scraping process to collect tweets with the keywords "Bitcoin" and "BTC." The Twitter scraping process was extremely challenging. Twitter does not like to make public tweets easy to download, which meant finding roundabout ways to collect the tweets. The tweets downloaded in this study were captured by using Taspinar's Twitterscraper program, which uses python to request the tweet data iterate through the tweets related to the keyword, and then uses Beautifulsoup4 to parse through the HTML code to return the tweet data. The program allows the user to scrape tweets based on keyword parameters, the language of the tweet, and the time period requested. The tweets are then returned as a JSON file. In order to export the tweets to an excel file, as well as add a sentiment variable, an additional python program was created. This python program iterates through each tweet and appends the relevant data to a list. The tweet data selected was the full text of the tweet, the number of favorites, number of replies, number of retweets, the username of the tweet's post, and the date and time the user posted the tweet.

Sentiment variables were created using the VADER sentiment analysis tool. A more detailed description of how sentiment scores are created can be found below in the Empirical Specifications section. The sentiment variables used are a positive sentiment score, a neutral sentiment score, a negative sentiment score, and a compound sentiment score. The specific sentiment scores are useful to understand the breakdown of sentiment for the text, and allows for a logic test to check if the compound score is accurate based on the breakdown of sentiment scores. The compound score is more useful for determining overall sentiment of the text. Based on the summation of the valence scores for each word, the text receives a normalized score between -1 and 1, with negative one the most extreme negative and positive one the most extreme positive.

An additional variable was created to give a weight to the sentiment. It seems logical to assume that there is a difference in the sentiment of a tweet between a post with zero favorites, replies and retweets, and a tweet that has a much larger response activity. In order to account for this difference, a weight variable was created. Further, it is reasonable to assume that a user's agreement to the sentiment of the tweet would differ based on their action of favoriting, replying, and/or retweeting the original post. To represent this difference, the weight for each tweet sentiment was calculated by the summation of: 1.5 times the number of favorites, the number of replies, and 2 times the number of retweets. This method of weighting based on different actions represents the differences between how a user views a favorite, reply, and retweet. If a user retweets a post they demonstrate agreement with the sentiment and actively spread this sentiment with their own followers. Similarly, a favorite demonstrates some level of agreement with the sentiment, but should not be weighted as strongly as a retweet given that the user did not spread the sentiment to others. It is difficult to properly weigh the action of a reply because it is unknown whether a user's reply demonstrates agreement or disagreement with the tweet sentiment. It should be noted that there is a possibility that this method could cause an exaggeration in the weighted sentiment because it does not account for the possibility that the same user performed multiple response actions to the tweet. Although a favorite and retweet on the same tweet does reflect a stronger agreement of sentiment than solely performing one of the actions, there is some exaggeration using this method.

In order to perform time-series analysis for the price of bitcoin with regards to sentiment, the sentiment about bitcoin must be aggregated for each day. To do so, an arithmetic mean was taken for each day. Note that this calculation was an average of the sentiment of the tweet after multiplying the compound sentiment score created by the VADER analysis and the weight assigned as explained above. However, one problem occurred when calculating the weighted sentiment of each tweet. Given the procedure for weighing each tweet's sentiment, if a post had zero retweets, replies, and favorites, the weight would have been zero and therefore would result in a weighted sentiment score of zero. However, this calculation is inaccurate because it implies that the sentiment of such a tweet is 0, meaning neutral, yet the VADER sentiment score was not necessarily zero. To take this issue into account, a conditional statement was used so that if a tweet had zero retweets, replies and favorites, the original VADER compound sentiment score was used instead the tweet receiving a sentiment score of zero.

4.3 Data Collection Issues

Over the course of the data collection, the scraping process ran into many problems. In the original scraping program, a different collection process was used that acted like a bot that automatically scrolled through tweets, extracting the data through the HTML code. However, this process was unsuccessful because it collected data too slowly to allow for enough data to be collected. Several attempts to speed up the scraping process were made but with no avail, which resulted in a significant setback in data collection. A new approach had to be made which required starting from a blank code file. After further research, the Twitter API scraping method seemed to be better alternative. However, this new method came with its own problems that had to be debugged.

The scraping process itself took several days to run. Initially several problems occurred because of unique instances that shut the scraping down because it tried to collect data that was not available on a certain tweet. After creating multiple if statements to allow the scraper to continue if there were tweets with missing information another attempt to scrape was performed, but again unsuccessful. Twitter's API records tweet data in different ways depending on if the tweet was retweeted by a user or not. More if statements were then incorporated to account for more conditions of tweet data and the scraping process was resumed once again; however, the next problem arose when the network used for internet connection crashed and several days of tweet scraping was lost. After inquiring to the network provider, they resolved the issue and the program began scraping again.

However, another issue occurred after 250,000 tweets were collected. Upon examination of the data collected, the dates of the tweets collected seemed too random to be accurate based on the program created. After a closer look, a new problem was revealed because of how Twitter keeps track of the tweets. As mentioned earlier, Twitter records data differently depending on if the tweet has been retweeted by a user. This difference in the data caused an error with the date and time recorded. Instead of recording the date and time that the post was created, the scraper collected the date and time the user created their account. This meant that the 250,000 tweets collected were worthless because they did not have accurate dates, which would result in meaningless regression results. After fixing this error, the python program could accurately capture all the necessary data. Yet another problem arose after collecting another 100,000 tweets. The tweets collected only spanned the previous two days from the time the scraping process began. Clearly this time interval is not a significant enough time period to allow for substantial or significant results, so an additional if statement was created. The new code allowed the program to check the date of the tweet and only if the date was within the selected timeframe would the tweet be captured. Although this fixed the problem with the time period selection, this method resulted in an additional problem.

Based on the way that Twitter's API returns tweet data, data is returned in a stream, meaning it is returned from most recent posts and then begins to go back in time. Given the sheer volume of tweets related to bitcoin, the scraper was left running for four days yet only returned data from the day the process was started and the previous day. This meant that it would be infeasible to capture the Twitter data for the time period selected (four months prior to when the scraping process began). Due to this result, a new method of capturing tweets was required. Using Taspinar's python Twitterscraper as the base of the program, a new python program was created to allow for time period selection. After months of debugging code and resolving problems, the tweet data ultimately spanned from December 1st 2017 to December 31st 2017 with a total of 513,660 tweets.

These tweets were then cleaned to remove tweets that were collected, but do not actually relate to the public sentiment of bitcoin. Many tweets were captured that can be considered spam, meaning a user promoted their own cryptocurrency strategies, but simply do not reflect a sentiment towards bitcoin. After the removal of the spam tweets, the dataset resulted in 511,087 tweets related to the keywords "bitcoin" and "btc." It should be noted as a limitation of this study that more tweets unrelated to the public sentiment of bitcoin were included and not removed due to the sheer size of the dataset. Although the initial plans for this research included capturing Bitcoin forum data and the forum sentiment, the Twitter data scraping process required much more work and time than initially anticipated. It should be recognized as a limitation of this study that Twitter data reflects the sentiment of the lay investor, whereas the forum data would have given an insight to the sentiment of those more in tune with the technology itself based on the difference between the types of people that post on Twitter versus a technical bitcoin forum.

4.4 Variable Description and Specification

Based on prior studies and economic rationale, several other traditional explanatory variables are used. One study found that the most significant explanatory variable in bitcoin price creation was the price of the VIX index ETF (Estrada, 2017). They concluded that the relationship between the volatility index and bitcoin's price was because investors believe bitcoin to be a risky asset and therefore implies a relationship between the riskiness of bitcoin and the ETF measuring the volatility of the S&P 500 (Estrada, 2017). Additionally, other prior studies included other financial related variables that are typically correlated with other assets such as the price of gold and the price of the S&P 500 ETF (Mai et al., 2016; Kim et al., 2017). The VIX price data and the S&P 500 ETF price data was collected from Yahoo Finance for the same time period as the bitcoin data, with the daily adjusted closing price used. The price of gold futures was captured using a Bloomberg terminal to download the daily GC1 COMB price (gold futures price) for the selected time period. The table on the next page in Table 1 depicts a table with the variables used and a description of what the variables represent.

Variable	Description	Predicted Effect
Ability to supply	Difficulty of finding a new block	Negative
Demand	# of daily confirmed bitcoin transactions	Positive
Supply	# of Bitcoins unmined	Negative
Market Volatility	price of volatility index ETF	Negative
Price of Gold	the price of GC1 COMB gold futures	Negative
Price of Equities	the price of the S&P 500 ETF	Negative
Public Sentiment	sentiment based on public tweets	Positive
Investor Sentiment	sentiment based on investor beliefs	Positive
Riskless Rate	rf based on US 10Y Treasury Note	Negative
Riskiness of Asset	volatility of BTC price	?
Liquidity of Asset	the level of difficulty of buy/sell of BTC	Positive

Table 1

The ability to supply new bitcoin transactions can be measured by the relative difficulty that the miners face to find a new block. The more difficult it is for the miners, the more it will cost the miners due to increased electrical power costs and costly upgrades to increase computing power. This increased cost therefore would suggest a negative relationship to price. In the case of supply and demand, the predicted effects are straightforward. If demand increases, the price will increase, indicating a positive relationship. Given that supply and demand have an inverse relationship, supply will have a negative relationship with price.

Since bitcoin is a risky asset and the VIX measures market volatility, one might suspect that there is a positive relationship between the VIX and the price of bitcoin. However, a prior study found that there was a statistically significant negative relationship between the VIX and the price of bitcoin (Estrada, 2017). This relationship can be rationalized by thinking about investor behaviors. Given that the VIX gives investors an opportunity to hedge against market risk, bitcoin and VIX related products become substitute assets, meaning an increased demand of one results in the decreased demand of the other. The price of gold futures would likely have a negative relationship to the price of bitcoin for the same reason that gold futures are negatively correlated with equities. Bitcoin is a risky asset and gold is seen as a safe-haven asset, which implies an inverse relationship. The price of equities, using the S&P 500 ETF as a proxy, would have a negative relationship to the price of bitcoin because the two are seen as alternative investments. Given the differing levels of risk between the assets, it seems to reason that investment in one of the assets would cause a decrease in the other. The riskless rate using the yield on a 10 year U.S. Treasury bond has a negative relationship for the same reason that it has a negative relationship to equities. If bond yields are higher, there is less incentives to invest in risky assets.

Note that riskiness and the liquidity variables have been re-coded. In order for the volatility of bitcoin price to accurately represent the riskiness of the asset, information regarding the overall riskiness of the cryptocurrency market would be required. This data would be extremely difficult to capture and would require creating an index of the entire cryptocurrency market. Alternatively, the volatility of bitcoin could be proxied by the volatility of the bitcoin futures market; however, this proxy would be inaccurate for this study given the time period of interest. The bitcoin data examined in this study occurred prior to opening of bitcoin futures exchanges and therefore do not reflect the true volatility of bitcoin for the selected time period. Additionally, the liquidity of bitcoin had to be re-coded since no existing variable captures the liquidity of bitcoin. As a proxy of liquidity, the number of transactions in the mempool was used as a proxy, where the mempool is the number of unconfirmed transactions on the blockchain on each day. The variable was re-coded such that the average of the three mempool values is taken for each day, and then every day's mempool was divided by the total number of transactions for each day in order to reflect the relative relationship between the number of unconfirmed transactions and the total number of confirmed transactions. However, this value would have a negative relationship with liquidity so the variable was then inverted by taking 1 divided by the variable created to reflect a positive relationship with liquidity. It should be noted that performing lognormal transformations are often done to deal with extreme outliers and can normalize the variance across observations; however, these transformations are best suited when there are large changes in unit or if the data is log-normally distributed, so for the purpose of the data used in this study, it would necessarily improve the data. Basic information about each variable can be seen on the next page in Table 2.

Variable	Interaction	If Y, which variables	Min Value	Max Value	Average Value	Median Value	Standard Deviation
Ability to supply	Y	Demand, Supply	1.347E+12	1.873E+12	1.668E+12	1.59E+12	2.00E+11
Demand	Y	Supply	228,926	490,644	360,898	370,277	56,226
Supply	Y	Demand	16,713,925	1,6774,500	16,745,135	16,746,025	18,362
Market Volatility	Y	SPDR, Gold	$9.34	$11.68	$10.26	$10.17	$0.73
Price of Gold	Y	VIX, SPDR	$1,238.50	$1,309.30	$1,265.88	$1,262.50	$19.98
Price of Equities	Y	VIX, Gold	$260.81	$267.13	$264.43	$264.91	$1.98
Public Sentiment	Y	Price	-	-	-	-	-
Investor Sentiment	Y	Price	0.2514	1.7544	0.8399	0.7560	0.3897
Riskless Rate	Y	SPDR, Gold, VIX	2.33%	2.50%	2.40%	2.39%	5.13%
Riskiness of Asset	Y	Price	11083.67	18894.50	15222.15	15276.52	2308.90
Liquidity of Asset	Y	Supply	223.61%	1120.13%	431.16%	329.90%	233.69%

Table 2

The predicted sign of the relationship between the price of bitcoin and the riskiness of the asset puts forth an interesting question. Do investors recognize more opportunity for higher expected returns with bitcoin, or are they fearful of it? Typically investors view risk as a negative, which would imply a negative relationship; however, bitcoin is not like other assets, and it is possible that investors like the risk associated with bitcoin, especially those shorting it. Lastly, the liquidity of bitcoin would have a positive relationship with the price of bitcoin because an increase in the ability to trade the asset results in an increase in demand, which therefore increases price.

5 Empirical Specifications

5.1 Sentiment Specification

After collecting the scraped Twitter/forum data as well as the Bitcoin data, previous studies have performed several different sentiment analysis on the social media data. All studies used a program to analyze the sentiment. Several studies used an algorithm called VADER for sentiment analysis, while other studies used more simple analytical tools to determine sentiment. Studies classified sentiment in several different ways. Some classified social media posts between positive, neutral, or negative, while other studies broke the sentiment down further to very positive, positive, neutral, negative, and very negative. Many of the studies in the field performed the Granger Causality test in order to determine which variable was affecting which; in other words, to determine if social media affected price, or if price affected sentiment. These studies come to the consensus that OLS regression can determine short run affects, while vector modelling is needed to determine the more long term results.

In this study, the aforementioned VADER (Valence Aware Dictionary and eEntiment Reasoner), sentiment analysis tool was applied. This method of sentiment analysis was used because several prior research studies had used this method and found the results to be the most accurate they tested (site some sources). Additionally, the VADER sentiment analysis is a lexicon and rule-based sentiment tool designed specifically for social media sentiment analysis. The VADER analysis takes into account typical negations such as “not,” punctuation such as exclamation points, slang words, emoticons, word shape such as capital letters, and degree modifiers such as “very” (‘Cjhutto”, 2018).

5.2 Econometric Specification

It should also be recognized that there is likely interaction between bitcoin price and sentiment. If we consider sentiment to be an important factor in bitcoin price formation, it makes sense that the resulting price would then have an effect on sentiment creating interaction between the variables. Given this interaction, it becomes necessary to perform Granger Causality tests. First it was necessary to time set the data to a delta of 1 day because each variable was captured and/or aggregated for each day from December 1st, 2017 to December 31st, 2017. The next step was to perform a unit root test on the price and sentiment variables. In this study, Augmented

Dickey-Fuller (ADF) tests were used in order to determine stationarity. Given the initial hypothesis that sentiment may granger cause price and/or vice-versa, it is also possible that the lag of sentiment may have an effect on the lag of price or on itself. The same relationship may also occur for the lag of price on the lag of sentiment or a lag of itself. This possible relationship requires a test for optimal lag selection, which was performed through using Stata.

The optimal number of lags for both the sentiment and price variables were determined using the Akaike Information Criterion (AIC) because it has been found to yield the most accurate results for datasets with under 60 observations (Khim & Liew, 2004). The ADF was then performed for both the sentiment and price variable with their respective optimal lag number. For the ADF test, the null hypothesis is: $H_0: \Theta = 0$, meaning the data is non-stationary and therefore may have random walk or drift, and the alternative hypothesis is: $H_A: \Theta < 0$, meaning the data is stationary. If the null hypothesis cannot be rejected, a first difference must be performed on that variable in order to make the time-series stationary.

The next step is to find the optimal number of lags for the Vector Autoregression (VAR) model. In order to do so, the Stata was used to test the optimal number of lags, but with respect to both sentiment and price in the same model. It should be noted that a maximum number of lags had to be inputted as a parameter when testing for the optimal number of lags. There is no commonly accepted calculation for determining the maximum number of lags to test. If too many lags are considered, the power of the test is diminished because of erroneous lags, yet not enough lags could result in the optimal lag selection to be affected if the true optimal lag is greater than the maximum number tested. Given this information it is standard practice to use economic rationales for the maximum number of lags tested. Since the data represents daily values, the maximum number of lags should be relatively small. It seems reasonable to assume that lag effect of greater than a week, 7 days, would have a significant impact, and therefore for the purpose of this study, the max lag of 7 is used. Based on the VAR model with a max lag of 7, the optimal number of lags for the model is then selected once again by the AIC.

Once the VAR model including the optimal number of lags is created, the Granger Causality test can be performed using Stata. Given that we are testing Granger causality for sentiment and price, it is the Granger causality bivariate linear autoregressive model that is used. In mathematical form the tests conducted can be represented in a two model equation:

$$x_t = a_1 + \sum_{i=1}^{p} a_i x_{t-i} + \sum_{i=1}^{p} \beta_i y_{t-i} + \epsilon_{1t}$$

and

$$y_t = a_2 + \sum_{i=1}^{p} \gamma_i x_{t-i} + \sum_{i=1}^{p} \delta_i y_{t-i} + \epsilon_{2t}$$

where x_t and y_t represent the stationary lagged variables (sentiment and price), with p as the optimal number of lags, and ϵ_{1t} and ϵ_{2t} as the residual error terms (Bai et at., 2010). Given these equations, two null hypotheses are tested:

$$H_0^1 \colon \beta_1 = \cdots = \beta_p = 0$$

and

$$H_0^2 \colon \gamma_1 = \cdots = \gamma_p = 0$$

Based on the two null hypotheses, there are four possible outcomes. If neither null hypotheses can be rejected, then there is not a statistically significant linear causal relationship, meaning *x* does not granger cause *y*, and *y* does not granger cause *x*. If H_0^1 is rejected, but H_0^2 cannot be rejected, then it suggests that *y* granger causes *x*, but *x* does not granger cause *y*. Likewise, if H_0^1 cannot be rejected, but H_0^2 can be rejected, it implies that a linear causality exists such that *x* granger causes *y*, but *y* does not granger cause *x*. The last possible result is that both H_0^1 and H_0^2 can be rejected, which therefore implies a two-way causal relationship such that *x* causes *y* and *y* causes *x* through feedback.

Depending on which outcome occurs from the Granger causality test determines the next appropriate econometric procedure. Although initially an ordinary least squares (OLS) linear regression was going to be used to determine the impact of the independent variables, if significant, on the dependent variable, but after initial testing and further thinking, this technique creates several issues. Although sentiment and price were tested for stationarity through performing an ADF test, the remaining independent variables were not tested for a unit root. If an OLS regression is run with non-stationary variables in a time-series, the regression can give spurious results. If this misleading result is in fact a result of using non-stationary time-series variables, more advanced econometric tools must be performed such as cointegration tests. Although a simple linear OLS regression can yield a regression output with a high R^2 value, indicating a high level of the goodness of fit of the model, if stationary and non-stationary variables are used but are not found to be cointegrated then it implies that they have different trends and therefore do not provide any meaningful insights in their relationship in the long run.

An additional problem may occur, particularly in time-series analysis, if the error term observations are correlated, known as autocorrelation. Autocorrelation occurs when a variable is correlated to a lag of itself. If there is a significant relationship between an independent variable and its lag, but it is not included in the regression, then the linear regression will not yield accurate results. It is also possible that an OLS regression would give inaccurate results because the independent variables are correlated with each other. If this correlation occurs, it can result in a high R^2 value because the variables do in fact fit the model well, but the t-values would be inaccurate and therefore the true significance of the independent variables would be meaningless. It should be noted that an OLS regression will underestimate errors in the presence of serial correlation in errors. If correlation between independent variables occurs, it would be necessary to create interaction terms to account for this correlation and improve the linear regression.

Given the independent variables used in this study, it is likely that some of these potential problems would occur while using an OLS regression, and therefore a VAR model has been selected to run the econometric regression. A mathematical depiction of the VAR model can be represented by:

$$y_t = c + \prod_1 y_{t-1} + \cdots + \prod_p y_{t-p} + \in_t$$

where y_t is an $(n\ x\ 1)$ vector of the time series variables containing $y_t = (y_{1t}, \ldots, y_{nt})'$, $\prod_t$ are $(n\ x\ n)$ coefficient matrices, and $\in_t$ is an $(n\ x\ 1)$ vector of the error terms (Richards, n.d.). The VAR is an autoregressive model that forecasts the all variables to determine the impact of each variable on all of the other variables, and itself if lags are included. However, it should be noted that, as previously discussed, the VAR model also requires the variables to be stationary. Additionally, the VAR model assumes that each variable effect all of the variables causing the coefficients to be difficult to interpret (Hydnman & Athansopoulous, n.d.). Due to time constraints of this study, a basic VAR model is performed, but more time would be required to develop a more detailed VAR model.

6 Findings

6.1 Augmented Dickey-Fuller Tests

Based on the procedure outlined in the empirical specification section, both sentiment and price were first tested for stationarity given their respective optimal number of lags by an ADF test. In the case of sentiment, the optimal lag given by the AIC was lag 0 and the resulting ADF test for sentiment given lag 0 was a p-value of 0.00, which indicates that the null hypothesis can be rejected at both the 5% and the 1% level, meaning the variable is stationary and no first differencing is required. For the price of bitcoin, the optimal lag selection given by AIC was lag 1 and the resulting p-value from the ADF test of price with lag 1 was 0.1804, which is greater than 0.05, and therefore we cannot reject the null hypothesis, meaning the variable is non-stationary and the first difference is needed.

6.2 Vector Autoregressive Model and Granger Causality Test

The optimal number of lags for the VAR model with both price and sentiment as endogenous variables was found to be lag 7, once again based on the AIC selection. The VAR model was then created with lag 7 and then the Granger causality test was performed on this VAR model. The resulting Stata output can be seen below in Figure 3**:**

vargranger

Granger causality Wald tests

Equation	Excluded	chi2	df	Prob > chi2
price	sentiment	**.77385**	**1**	**0.379**
price	ALL	**.77385**	**1**	**0.379**
sentiment	price	**1.3447**	**1**	**0.246**
sentiment	ALL	**1.3447**	**1**	**0.246**

Figure 3

This output indicates that neither null hypothesis can be rejected, meaning that there is not a linear causal relationship from price to sentiment nor from sentiment to price.

6.3 Vector Autoregressive Model

In order to run the VAR model, the optimal number of lags is first required. When including the price of bitcoin, the VIX, the S&P 500 ETF, gold futures, the yield of the U.S. 10 Year Treasury Note, twitter sentiment, and the proxied volatility of bitcoin and liquidity, there were two different optimal lag selections based on AIC. Depending on the maximum of lags to be tested through while using the "varsoc" function, gave two different optimal lag numbers. When the maximum was set to 7, the optimal lag found was lag 6, while when the maximum was set to 1, the optimal lag number was 7. Given these results, two VAR models were run based on the two optimal lags found. Although the VAR model outputs the tests statistics for every variable on every other variable, for the purpose of this study, the most relevant output is the effect of each variable on bitcoin price. The resulting output for the VAR model with lag 6 and lag 7 can be found in Appendix A and Appendix B**,** respectively.

Interpreting the coefficients of the VAR model is not a useful exercise because they cannot be interpreted in the same manner as coefficients of an OLS regression. In the case of the VAR model, it is more relevant to examine the z-statistic to determine if a significant correlation exists, but it does not give an understanding as to the direct impact of the relationship. There are several notable findings based on the results of both VAR models. In the case of the VAR model with the optimal lag selected to be 6, there are three significant relationships between the variables and the price of bitcoin. The result indicates that a statistically significant relationship exists between lagged gold price and the price of bitcoin. This relationship was also found when using an optimal lag of 7 for the VAR model. With a z-value of 0.00 for both lags, we can reject the null hypothesis at both the 5% and 1% level, indicating a statistically significant relationship. This result is interesting and should be examined further with more in-depth research because gold is an extremely stable asset relative to bitcoin. Given that the sign of the coefficient is negative, this result indicates that the relationship between gold futures and the price of bitcoin is negative. This negative relationship suggests that investors purchase bitcoin in periods of decreased volatility, given that we know investors purchase gold in more turbulent time periods.

The next finding based on the VAR model output is the relationship of the VIX to price. Although the z-value for the lag 6 of the VIX on price is too large to reject the null hypothesis, the resulting z-value is 0.00 when the optimal lag is changed to 7, indicating we can reject the null hypothesis. For a better understanding of the true relationship between the VIX and the price of bitcoin, an Impulse Response Function (IRF) graph was created to visually represent the relationship throughout the lags. The IRF graph was created in Stata and can be seen below in Figure 4**:**

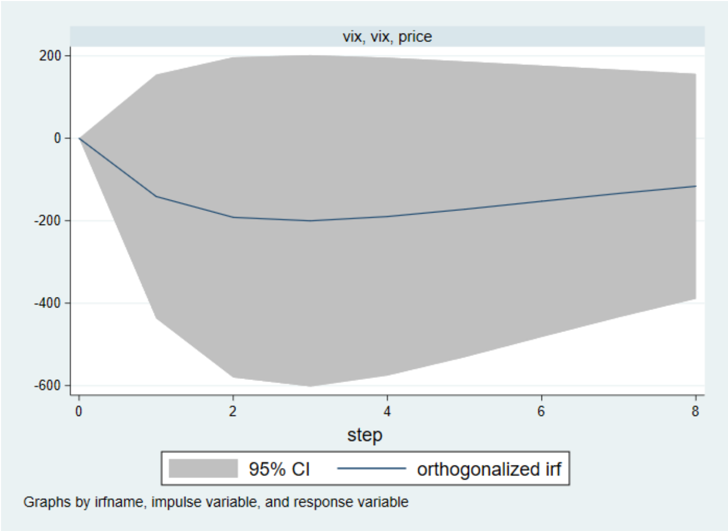

Figure 4

An impulse response function indicates the evolution of a specific variable reacting to a specific time period shock that impacts another variable, which provides a particularly useful insight when understanding the interaction between variables in a VAR model. From this graph, we can see that a negative relationship occurs between the VIX and the price of bitcoin. This relationship agrees with the initial predicted sign. Returning back to the original hypothesis, a negative relationship occurs due to how investors view the risk associated with bitcoin. When market volatility is high, investors tend to focus their investments on shorting the VIX and therefore have less to invest in bitcoin, while during periods of low volatility, investors like to find alternative sources of risk to maintain their target portfolio risk. This relationship agrees with the relationship found between gold and bitcoin.

A statistically significant result was also found between sentiment and price, when using the optimal lag of 6. With a z-value of 0.049, we can reject the null hypothesis at the 5% confidence level, but not the 1% level. Once again, to get a clearer picture of this relationship, an IRF graph was created and can be seen on the next page in Figure 5:

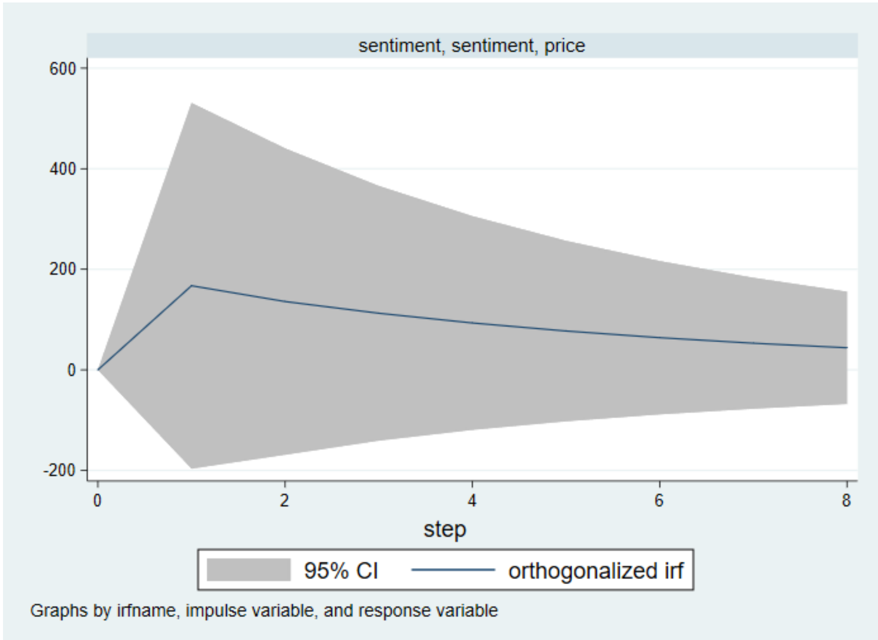

Figure 5

This graph is particularly interesting because it shows the diffusion of sentiment over time (through the various lags), as well as a large confidence interval at lag 1. However, given that the significance of sentiment was determined by the number of optimal lags, it is difficult to definitively describe the true relationship. Given that one prior found a positive relationship between twitter sentiment and the price of bitcoin, yet a prior study found a negative relationship, the result found in this study is somewhat understandable (Mai et al., 2015; Georgoula et al., 2015). Further, the relationship found in this study, in combination with the other two mentioned studies, may indicate that the statistical significance may be caused by time period selection. Additionally, the number of tweets as well as the accuracy of the sentiment analysis is crucial for an accurate econometric approach to understanding the true relationship between sentiment and price.

7 Conclusion

In this paper examining the effect of sentiment on the price formation of bitcoin, several conclusions can be made. Depending on what optimal lag is used, a significant, negative, relationship between the price of gold futures and the price of bitcoin exists. Additionally and even more conclusively, this study finds a statistically significant negative relationship between market volatility, as proxied by the price of the VIX ETF. These relationships give an important insight into how investors view bitcoin as an asset. Both relationships suggest that investors view bitcoin as a substitute, meaning they only invest in bitcoin or gold/VIX. This conclusion can be explained by thinking about how investors manage risk in their portfolios. Given that investors manage risk by investing in various assets based on a target amount of risk, they can simply choose which asset to invest in based on market volatility at any given time. If the expected return and associated risk of investing in the VIX ETF is sufficient for hitting their target level of risk, then they will forego investments in alternative risky assets like bitcoin, while during periods of low volatility, investors may turn to bitcoin for an additional source of risk to balance their portfolios.

With regards to the relationship between public sentiment and the price of bitcoin, it is difficult to conclusively determine if a significant relationship exists. This study found a positive, significant relationship at the 5% confidence level, but it was dependent on the maximum lag criteria and subsequent optimal lag selection. This result can be explained by several factors. One possibility is that sentiment effects price both positively and negatively. It is possible that the influence of the user behind the sentiment effects the relationship to a greater extent than anticipated. Although some level of sentiment weighing was performed based on the number of retweets, favorites, and replies that a post received, it is possible that only the sentiment of key influences effect the price of bitcoin, while users whose posts did not receive much activity have little effect on price, and therefore could affect the regression. Additionally, it is possible that the time period selected and therefore the implied number of tweets collected was insufficient to determine a statistically significant result. Lastly, it is also possible that due to the short time frame of bitcoin as an asset, there may not be a clear model for price formation, and therefore the current prices are essentially random. In this case, bitcoin price may become normalized in the future after the validity of bitcoin's value has been tested, and then a more accurate pricing model can be created. This paper should provide valuable insights into what are several key factors that impact the price of bitcoin, which can be expanded with future research. Other studies did not perform IRF tests, but they were found to be particularly insightful in this study, indicating that additional IRF tests of VAR models should be used in further research when examining the impact of variables on the price formation of bitcoin. Given the possible sources of error, future studies should take careful consideration when determining the weight of the sentiment for each tweet, as well as collecting more tweets spanning a period of longer than one month.

8 References

Acheson, N. (2018, April 16). Bitcoin Guides Archive. Retrieved from http://www.coindesk.com/information/.

Ali, J. I. (2017, March 09). The Blockchain will do to the financial system what the internet did to media. Retrieved from https://hbr.org/2017/03/the-blockchain-will-do-to-banks-and-law-firms-what-the-internet-did-to-media

Bai, Z., Zhang, B., & Wong, W. (2009). Multivariate linear and non-linear causality tests. *SSRN Electronic Journal*. doi:10.2139/ssrn.1408542

Barberis, N., Huang, M., & Santos, T. (1999). Prospect theory and asset prices. doi:10.3386/w7220

Bitcoin and Ethereum vs Visa and PayPal – transactions per second. (2017, April 22). Retrieved from https://altcointoday.com/bitcoin-ethereum-vs-visa-paypal-transactions-per-second/

Blockchain Wallet Users. (2018). Retrieved from https://blockchain.info/charts/

Bohme, R., Christin, N., Edelman, B., & Moore, T. (2015). Bitcoin: economics, technology, and governance. *Journal of Economic Perspectives, 29*(2), 213-238. doi:10.1257

Celik, S. (2012). Theoretical and empirical review of asset pricing models: a structural synthesis. *International Journal of Economics and Financial Issues, 2*(2), 141-178.

Ciaian, P., Rajcaniova, M., & Kancs, D. (2015). The economics of bitcoin price formation. *Applied Economics, 48*(19), 1799-1815. doi:10.1080/00036846.2015.1109038

Colianni, S., Rosales, S., & Signorotti, M. (2016). Algorithmic trading of cryptocurrency based on Twitter sentiment analysis. *SSRN Electronic Journal*.

Estrada, J. C. (2017). *Analyzing Bitcoin price volatility* (Unpublished master's thesis). University of California, Berkely.

Georgoula, I., Pournarakis, D., Bilanakos, C., Sotiropoulos, D. N., & Giaglis, G. M. (2015). Using time-series and sentiment analysis to detect the determinants of Bitcoin prices. *SSRN Electronic Journal*. doi:10.2139/ssrn.2607167

Global Charts. (2018). Retrieved from https://coinmarketcap.com/charts/

Houlihan, P., & Creamer, G. G. (2017). Can sentiment analysis and options volume anticipate future returns? *Computational Economics, 50*(4), 669-685. doi:10.1007/s10614-017-9694-4

Hung, M., & Wang, J. (2005). Asset prices under prospect theory and habit formation. *Review of Pacific Basin Financial Markets and Policies, 08*(01), 1-29. doi:10.1142/s0219091505000324

Khim, V., & Liew, S. (2004). Which lag length selection criteria should we employ? *Economics Bulletin, 3*(33), 1-9.

Kim, Y. B., Kim, J. G., Kim, W., Im, J. H., Kim, T. H., Kang, S. J., & Kim, C. H. (2016). Predicting fluctuations in cryptocurrency transactions based on user comments and replies. *Plos One, 11*(8). doi:10.1371/journal.pone.0161197

Kim, Y. B., Lee, J., Park, N., Choo, J., Kim, J., & Kim, C. H. (2017). When Bitcoin encounters information in an online forum: using text mining to analyze user opinions and predict value fluctuation. *Plos One, 12*(5). doi:10.1371/journal.pone.0177630

Kristoufek, L. (2015). What are the main drivers of the Bitcoin price? Evidence from wavelet coherence analysis. *Plos One, 10*(4). doi:10.1371/journal.pone.0123923

Loughran, T., & Mcdonald, B. (2014). Measuring readability in financial disclosures. *The Journal of Finance, 69*(4), 1643-1671. doi:10.1111/jofi.12162

MacDonell, A. (2014). *Popping the Bitcoin bubble: an application of log-periodic power law modeling to digital currency* (Unpublished master's thesis). University of Notre Dame.

Mai, F., Bai, Q., Shan, Z., Wang, X. (Chiang, R. H. (2015). From Bitcoin to big coin: the impacts of social media on Bitcoin performance. *SSRN Electronic Journal*. doi:10.2139/ssrn.2545957

Matta, M., Lunesu, I., & Marchesi, M. (2015). The predictor impact of web search media on Bitcoin trading volumes. *Proceedings of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*. doi:10.5220/0005618606200626

Matta, M., Lunesu, L., & Marchesi, M. (2015). Bitcoin spread prediction using social and web search media. *Research Gate*.

McMillan, R. (2018, February 09). The inside story of Mt. Gox, Bitcoin's $460 million disaster. Retrieved from https://www.wired.com/2014/03/bitcoin-exchange/

Nielsen, H. B. (n.d.). *Non-stationary time series and unit root tests*. Lecture.

Nofer, M., & Benlian, A. (2015). *The value of social media for predicting stock returns preconditions, instruments and performance analysis* (Unpublished doctoral dissertation). Zugl.: Darmstadt, Techn. Univ., Diss.

Nofer, M., & Hinz, O. (2015). Using Twitter to predict the stock market: where is the mood effect? *The Value of Social Media for Predicting Stock Returns,* 63-88. doi:10.1007/978-3-658-09508-6_4

Patton, A. J. (2006). Volatility forecast comparison using imperfect volatility proxies. *SSRN Electronic Journal*. doi:10.2139/ssrn.932890

The relationship between cryptocurrency and social media. (2018). Retrieved from https://99bitcoins.com/cryptocurrency-and-social-media/

Stenqvist, E., & Lonno, J. (2017). *Predicting Bitcoin price fluctuation with Twitter sentiment analysis* (Unpublished master's thesis). KTH Royal Institute of Technology.

Sul, H. K., Dennis, A. R., & Yuan, L. I. (2016). Trading on Twitter: using social media sentiment to predict stock returns. *Decision Sciences, 48*(3), 454-488. doi:10.1111/deci.12229

Things you need to know – Bitcoin.com. (2018). Retrieved from http://www.bitcoin.com/you-need-to-know

Torres-Reyna, O. (n.d.). *Time series*. Lecture presented in Princeton University.

Vector Autoregressive Models for multivariate time series. (n.d.). *Modeling Financial Time Series with S-PLUSÂ®,* 385-429. doi:10.1007/0-387-32348-1_11

What is social media's role in the cryptocurrency world? (2016, December 21). Retrieved from https://news.bitcoin.com/social-media-role-cryptocurrency-world/

Xie, P., Chen, H., & Hu, Y. J. (2017). Network structure and predictive power of social media for the Bitcoin market. *SSRN Electronic Journal*. doi:10.2139/ssrn.2894089

9 Appendices

A. VAR Model with Dependent Variables on Price, Lag 6

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
price						
price						
L6.	-.8454684	.3308036	-2.56	0.011	-1.493831	-.1971052
vix						
L6.	-498.3773	601.521	-0.83	0.407	-1677.337	680.5823
spy						
L6.	69.38802	337.3249	0.21	0.837	-591.7567	730.5328
us10yr						
L6.	-1665.808	6210.969	-0.27	0.789	-13839.08	10507.47
gold						
L6.	-137.8056	25.30037	-5.45	0.000	-187.3935	-88.21783
sentiment						
L6.	774.1959	393.9798	1.97	0.049	2.009754	1546.382
btcvol						
L6.	.2413895	.4845856	0.50	0.618	-.7083809	1.19116
btcliquidity						
L6.	-108.5746	129.3241	-0.84	0.401	-362.0452	144.8959
ntransactions						
L6.	-.0042908	.0064196	-0.67	0.504	-.016873	.0082915
transactionfees						
L6.	-1.194711	1.159179	-1.03	0.303	-3.46666	1.077237
_cons	192191.6	72952.46	2.63	0.008	49207.41	335175.8

B. VAR Model with Dependent Variables on Price, Lag 7

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
price						
price						
L7.	-.7345962	.2555261	-2.87	0.004	-1.235418	-.2337742
vix						
L7.	-1646.145	463.791	-3.55	0.000	-2555.159	-737.1311
spy						
L7.	-330.2895	259.4978	-1.27	0.203	-838.8958	178.3169
us10yr						
L7.	19598.8	4776.963	4.10	0.000	10236.12	28961.47
gold						
L7.	-106.0753	19.46088	-5.45	0.000	-144.218	-67.9327
sentiment						
L7.	725.5426	308.5704	2.35	0.019	120.7558	1330.329
btcvol						
L7.	.2145492	.3735527	0.57	0.566	-.5176007	.946699
btcliquidity						
L7.	44.27232	99.68304	0.44	0.657	-151.1029	239.6475
ntransactions						
L7.	.0083985	.0050336	1.67	0.095	-.0014672	.0182643
transactionfees						
L7.	-2.925334	.8921008	-3.28	0.001	-4.67382	-1.176849
_cons	212564.8	56222.31	3.78	0.000	102371.1	322758.5