Summer 2019

Two Essays on Bitcoin Price and Volume

Mohammad Bayani Khaknejad
Old Dominion University, Mbaya001@odu.edu

TWO ESSAYS ON BITCOIN PRICE AND VOLUME

By

Mohammad Bayani Khaknejad

MBA, July 2014, Sharif University of Technology, Tehran, Iran

B.Sc., September 2004, Sharif University of Technology, Tehran, Iran

Dissertation submitted to the Faculty of

Old Dominion University

In partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

BUSINESS ADMINISTRATIONS – FINANCE

OLD DOMINION UNIVERSITY

AUGUST 2019

Approved by:

Dr. Mohammad Najand (Director)

Dr. Licheng Sun (Member)

Dr. David Selover (Member)

ABSTRACT

TWO ESSAYS ON BITCOIN PRICE AND VOLUME

Mohammad Bayani Khaknejad

Old Dominion University, 2019

Director: Dr. Mohammad Najand

Bitcoin is a decentralized peer to peer digital transactions system that was introduced in 2009 in the aftermath of the financial crisis. Since its introduction, it has had a volatile journey, being adopted by computer programmers, cyber punk enthusiasts, criminals, and financial investors. While the future of bitcoin is still not clear, it has been widely adopted by many, not necessarily as a new method of transactions, but rather as a new investment vehicle. Being a new asset class, there are many unknown financial characteristics to be investigated about bitcoin and in this dissertation, we try to explore two of these characteristics: Price and volume.

In the first essay, we investigate the price-volume relationship. The term "price-volume relationship" in the finance literature, usually implies either relationship between volume and the magnitude of return, or relationship between volume and return per se. It has been established by previous studies that volume is positively related to the magnitude of return. We document that this is the case for bitcoin as well, and that this is merely because of the resampling of observations. The relationship between volume and return per se, however, is more controversial. It has not been studied as heavily and it is mostly observed only in spot markets, which has led scholars to believe it is caused by the restrictions imposed on short-selling in spot markets. We examine this relationship in bitcoin spot and futures markets and argue that while it is only observed in the spot market, the absence of short-selling cannot be the reason for this relationship.

In the second essay, we use market sentiment measures derived from a lexical analysis of news platforms and social media networks to try and forecast returns. We find that our sentiment measures do indeed granger-cause returns in the spot market. However, they do not explain much variation in returns, and therefore are not useful in forecasting prices in the absence of a fundamental model. This relationship is weaker in the futures market which is due to the higher level of investor sophistication in that market. We also examine the effect of our sentiment measures on volatility of returns, and on trading volume and find that they do drive these variables as well.

DEDICATION

To my parents, Hassan and Maryam, for their endless love and sacrifices.

And to my brothers, Mohsen and Mojtaba, for their wholehearted support.

ACKNOWLEGEMENT

I have many individuals to thank for the knowledge and experience I gained in the doctoral program at Old Dominion University. First, I would like to thank the chair of my dissertation committee, Dr. Najand, who not only helped me with his advice in the completion of the present dissertation, but also constantly supported me throughout the doctoral program. He has been my instructor, mentor, and often times a true friend and I truly believe I have been blessed by his presence and his insightful advice.

Moreover, I would like to thank my dissertation committee members because without their help this dissertation could not have been accomplished. I would like to thank Dr. Sun for all of his fantastic ideas that kept me going during the times I faced obstacles in my research, and I would like to thank Dr. Selover for kindly taking the time to share his deep knowledge of statistics with me.

I would also like to thank Dr. Ford, the coordinator of the doctoral program at Old Dominion University, for all his efforts to maintain the rich and diverse program. I consider graduating from this program an honor and I hope to be able to promote this program with my future accomplishments.

Finally, I would like to thank all of the staff, faculty, and classmates who I had the pleasure of interacting with while at Old Dominion, without whose help and guidance my journey through the doctoral program would not have been as enjoyable.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE: A BRIEF INTRODUCTION TO BITCOIN

What is Bitcoin?

Bitcoin is a decentralized electronic cash system. The units of the currency, also called bitcoin, can potentially be used as a means to store and transmit value to other participants on the network. In essence, bitcoin can be used by users to buy and sell products or services, pay bills, etc., just like any other currency. Transmission of bitcoins is possible using the bitcoin protocol via the internet, or any other network.

Introduced in 2008, bitcoin is the first digital currency that has, so far, been successful in gaining popularity and value. It is a purely peer-to-peer currency, which means transactions occur directly between the payer and the receiver. It provides users with pseudonymity, while preventing double-spending. It addresses the problem of double-spending by timestamping transactions into a hash-based proof-of-work, called the ledger (Nakamoto 2008).

Bitcoin is a fiat currency, i.e. its value is entirely virtual. Its value is determined by its supply and demand. Its supply is predetermined and based on the processing power of miners. The demand for it, on the other hand, is rooted in the confidence of users in the bitcoin protocol.

Since bitcoin is a distributed, peer-to-peer system, it does not rely on a third party to verify transactions. However, there are “miners” that facilitate transactions and keep track of all of the transactions in a ledger. These miners use processing power to find solutions to a mathematical problem.

This process is introduced to achieve scarcity. In return, the miners are rewarded with a bitcoin almost every 10 minutes. This way the miners perform the currency issuance and clearing functions of a central bank, but in a decentralized manner.

Built-in algorithms in the bitcoin protocol, manage the rate at which new bitcoins are issued. The difficulty of the mathematical problem that should be solved so that a new “block” of transaction records can be added to the ledger, i.e. the “blockchain”, adjusts dynamically to the processing power of the miners so that the rate at which new bitcoins are issued is almost constant. Moreover, every four years this rate will be halved. Finally, the total number of bitcoins that will eventually be issued are limited and less than 21 million. The last bitcoin will be issued in year 2140.

Background

Bitcoin is not the first attempt to create a digital currency. There have been around one hundred different digital pay schemes prior to bitcoin. Although some of them like PayPal have survived, most of them have failed and are not functional anymore. In order to get insight into how these pay schemes work we should first go over the different systems of transaction, and the history of transaction itself.

- **Economics of Transactions**

Even before the invention of currency, humans used to trade. In fact, the history of barter is at least as old as the recorded history of man (Davies 2010). Barter, is the most basic and most ancient form of transaction which still exists today. In its basic form it is a direct exchange of services and resources for mutual advantage. As an example, if Alice has food and needs clothes, and Bob has clothes but needs food, then Alice and Bob can exchange the food and clothes to get what each one of them desires. Although it seems very simple and convenient, there are obviously problems associated with this system of transaction: The first problem one can think of is that in order for the barter to be successful, both parties should be present at the same place and at the same time, which makes the coordination of the transaction problematic. Using the internet can help mitigate this problem, but then it brings up the issue of security and trust. How can individuals trust each other over the internet in the absence of a third party supervisor?

A second problem with barter, is the possible mismatch between the needs of the parties, with what they actually have. If, in the above example, Bob who has clothes doesn't need food and instead needs weapons, then the simple two party barter will not work. This problem can be solved by finding a third party, called Charlie who happens to have weapons and need food. This way a triangular barter would satisfy everyone’s demands. The mindful reader notices that as we worked our way around the second problem, we exacerbated the first problem. Coordinating the triangular barter is going to be even more difficult than the case where there were only two parties involved. Moreover, in most real life circumstances we need more than just two parties to be able to address all of the needs of the parties involved using barter.

Now to solve this coordination problem, one can think of two solutions: Credit, and cash. Credit, in the basic form of an ‘IOU’[1] can help solve the problem of coordination. In our previous example, Alice will receive the clothes from Bob and in return sign an ‘IOU’ piece of paper and give it to Bob. The owner of this ‘IOU’ paper will be able to ask Alice a favor in the future. Now Bob can get the weapons he needs from Charlie in exchange for the ‘IOU’ paper, and finally Charlie will take the ‘IOU’ paper to Alice and ask her for the food. Of course, this system has its flaws such as default risk, i.e. the risk that Alice defaults on her obligation. Therefore, it is necessary that in such a system the issuer of the `IOU' paper and her credit history is known. This is basically how credit card systems work.

The cash system, on the other hand, will start by an initial allocation of wealth, i.e. cash. Now each party can freely trade with the other party in exchange for cash. Since cash is acceptable by all parties and is not dependent on any individual's credit, it does not involve default risk. Therefore, it might not be necessary to track individuals and their credit history in such a system. It is also more convenient since it allows one to quantify goods, services, and credit. This transaction system is the basis for many digital currencies, including bitcoin.

Before the introduction of bitcoin, there were many attempts to create a digital transaction system. In order to have a good view of the benefits provided by bitcoin, we are providing an

[1] Read ‘I owe you’

overview of the systems that preceded bitcoin. Understanding how these systems function, is vital to having an appreciation of bitcoin and the blockchain technology.

- **Credit Based Systems**

Here, we would like to introduce three types of credit-based systems that will give the reader insight on the processes that enable a transaction to be completed on the internet:

1. In the most basic type of credit transaction system, a user will interact directly with a merchant, like a website, that asks the user for their payment information (PI), e.g. credit card information. The merchant also has the order information (OI), i.e. what is being sold to the user. The merchant then sends both information to a credit card processor which in turn communicates with the credit card issuer to complete the transaction. In order for this type of transaction to be functional, all information should be shared with both the merchant and the credit card issuer. This system offers the lowest anonymity, as a user's PI and OI is shared by all parties in the system.

2. The second type of credit transaction systems, involves the presence of an intermediary that will interact directly with the user and the merchant. This will cause a disconnect between the user and the merchant. The intermediary receives the PI from the user, and the OI from the merchant, and then shares both, or only the PI with the credit card issuer. This denies the merchant from access to the user's PI, and can also deny the issuer access to the user's OI. A popular example of this type of systems is PayPal. This system provides some level of anonymity, but still both the user and the merchant should have a profile with the intermediary, and the intermediary has access to all of the information.

3. The third type is not really a transaction system, but rather a security protocol called Secure Electronic Transaction or SET. It was introduced in 1996 and established by VISA and MasterCard in cooperation with some IT giants. This system is similar to the first type, but uses an innovation called the dual signature (Stalling 2003). Using the dual signature the user

encrypts his/her information in two parts and then sends them to two different users. The OI is only shared with the merchant and the PI is only shared with the processor/issuer.

- **Cash Based Systems**

In general, cash-based systems, have two major advantages over credit based systems: It is not possible to double-spend cash, and it offers higher anonymity. Double spending is using a certain amount of money/credit for multiple transactions. An example is when someone with a balance of $100 writes two checks to two different entities each for $100. This cannot happen with traditional cash, because it cannot be copied, and it is spent at the same time that the transaction is taking place. The problem with digital cash, however, is that you cannot have both of these features at the same time. On the one hand, since it is digital it is copiable like any other digital data and on the other hand, appointing a third party to supervise and confirm all transactions, will take away the anonymity. Chaum (1983) was the first to address both of these seemingly conflicting concerns. He proposed that using a traceable serial number and a blind signature a user can receive an “unconditionally untraceable electronic money” from a bank and spend it with full anonymity. Nevertheless, the requirement that the bank should always be online limited its practicality.

Later on, Chaum, Fiat et al. (1988) argued that just like paper checks, we do not need to prevent, but rather, detect and trace double spending. They proposed a way to be able to identify users who double-spend electronic money. This system would not have been able to prevent a user from double spending, but as soon as they did, their identity would be revealed to the bank with almost certainty.

At this time, as the cyberpunk movement was gaining more and more attention, many scholars were trying to propose a way to widely use digital cash for transactions. Chaum and Pedersen (1992), Brands (1997), Ferguson (1993), Camenisch, Hohenberger et al. (2005), and Okamoto and Ohta (1991) among others helped pave the way to inventing a feasible electronic cash system.

In 1989, David Chaum, also founded DigiCash that offered anonymous payments and at the same time used tamper-resistant hardware to prevent, and not just detect, double-spending. Ultimately, DigiCash went into bankruptcy in 1998. Its failure was mainly rooted in two flaws: First, it was difficult to persuade merchants and banks to adapt this technology so there was no demand from these parties. Second, only the senders were anonymous and therefore it did not gain strong support from user to user transfers.

In 1998Wei Dai proposed the idea of b-money. It relied on two protocols: One was an unjammable broadcast channel which was not feasible, and the other was the separation of users and servers. Although it never became practical, the idea of b-money was incorporated in bitcoin and referenced by Satoshi Nakamoto (Nakamoto 2008).

At the same time Nick Szabo designed a decentralized digital currency that he named “bit gold”. Bit gold too was never implemented, but it had a lot in common with bitcoin. It offered pseudonymity and also prevented double-spending.

How does Bitcoin Work?

In order to get a deep understanding of how bitcoin functions, it is necessary to know some terms, functions, and protocols. Therefore, we will try to elaborate on some of these components hereunder:

- **Public and Private Keys**

Each wallet, or account, in bitcoin has at least one pair of public and private keys. The pair are mathematically related and are large 256-bit numbers that are difficult to guess or reproduce. The combination of these two will allow a user to spend a bitcoin. The public key is just like an email address that others can use to send a bitcoin to the user. The private key is like the password that allows the rightful owner of the bitcoin to sign the transaction. Upon signing the transaction, every node that knows the public key will be able to validate the transaction. Basically, any user who has access to the private key of a bitcoin is the true owner of it.

- **Digital Signature**

A digital signature, like a traditional signature on a check, validates a transaction. In other words, it proves that a user is the rightful owner of that coin and therefore has the right to authorize the transaction. It is different from a traditional signature in that upon being revealed, the signature cannot be used or forged in the future to create other transactions on behalf of the user.

- **Node**

A node is a computer of significant processing power that runs the bitcoin software. It relays information to several other nodes in the network which in turn relay information to other nodes, and in this fashion help information to ow pretty quickly. Some nodes are “miners”. Miners help bitcoin operate by solving a mathematical puzzle, also known as proof-of-work, and keeping track of all of the transactions in the network.

- **Proof of Work (PoW)**

Introduced by Dwork and Naor (1992) and formalized by Jakobsson and Juels (1999), it was originally a protocol to deter denial of service attacks, such as spams, on a network by requiring some processing time by a computer. It can be thought of as a puzzle that is feasible for a computer to solve, but at the same time difficult enough so that it takes some time. Spamming in its basic form is posting one same message repeatedly on a website in large numbers to make its service unavailable to other users. The idea was that if each posting requires some significant processing time from a computer, then the computer will not be able to efficiently spam the website. Using PoW, one can make a process as time consuming as they want.

Hashcash is a PoW introduced by Back (2002) that is used in electronic cash systems. While some electronic cash systems use PoW to control the rate of minting cash and to make it scarce, bitcoin uses PoW to update the blockchain, i.e. the ledger containing all the accounting entries.

- **Hash, Hashing, and Hash Functions**

Hashing is transforming strings of characters with variable lengths into values of fixed, and usually shorter, length called hash. This transformation is achieved by using hash functions. Bitcoin uses a hash function called SHA-256 to transform information from each transaction, to a hash which then becomes an entry in the blockchain, i.e. the ledger. Simply put, miners are trying to _nd a number that when combined with the transactions data in the block and passed through the hash function produces an output in a certain range. The hash function makes it impossible to solve for the number, and therefore, the most efficient way of finding that number is guessing it using random numbers.

The miner that manages to find the desired number is rewarded with some bitcoins, currently 12.5 bitcoins. The reward for mining halves every four years and the total number of bitcoins in circulation are expected to be capped at 21 million sometime in the year 2140.

- **Difficulty**

To ensure a stable rate of minting new bitcoins, the difficulty of the PoW is variable. This is a feature that is built into the bitcoin algorithm so that technological advancement and improved processing power will not accelerate the supply of bitcoins. The higher the processing power of the miners, the higher the difficulty of the puzzle they need to solve, so that on average a bitcoin is generated almost every 10 minutes.

ESSAY ONE: PRICE-VOLUME RELATIONSHIP IN BITCOIN MARKETS

Introduction

The relationship between trading volume and return (sometimes referred to as volume and price relationship) has been studied extensively. In particular, there are two relationships that have been the focus of studies in this stream of research: The relationship between volume and the magnitude of return, and the relationship between volume and return *per se*. The former relationship, i.e. the relationship between volume and the magnitude of return, has been found to be positive for most asset classes at calendar time frequencies (i.e. not transactions or tick frequency) and in both the spot and futures markets. Mixture of distributions (MDH) and the sequential arrival of information are the two prominent theories that try to explain the nature of this relationship. In both theories, the existence of a period of time during which multiple transactions occur are key assumptions which in turn explains why tests conducted at transactions frequencies may not result in a positive relationship between volume and the magnitude of return (Harris and Gurel 1986). Therefore it is safe to say that this relationship arises from the resampling of observations in calendar time intervals. The literature on the second relationship, i.e. the relationship between volume and return per se, is not as developed. The relationship between volume and return per se is usually observed in spot markets and not in futures markets. This has led scholars to suggest that this is due to the absence, or limitation, of short selling in the spot market. There is an explanation provided by Karpoff (1987) that attributes this relationship to the costliness, or absence of short selling in spot markets. In this study we use the unique opportunity that bitcoin markets provide, to examine this relationship. Bitcoin, introduced in 2009, is now widely traded with trading volumes exceeding 30,000 bitcoins in an hour (Nakamoto 2008). Due to the absence of debt instruments denominated in bitcoins, there is no reliable way of short selling bitcoin. Therefore, we can assume no significant short selling occurs in the bitcoin spot market. There also exists a futures market in which there is no extra cost to short selling. This provides a unique opportunity for us to test the relationship between volume and return per se in bitcoin markets and its possible ties to short selling or its costliness.

Our findings on the relationship between volume and the magnitude of return are consistent with those of the majority of studies in this field, i.e. there exists a positive and statistically significant relationship between trading volume and the magnitude of returns for both five minutes and hourly frequencies. At the transactions frequency, however, we find a negative relationship. As for the relationship between trading volume and return *per se*, while we document a negative and statistically significant relationship in the spot market, we fail to observe any significant relationship in the futures market. The significance and signs hold at all three frequencies. This is result is surprising in that although it is the opposite of what is expected by referring to Karpoff's (1987) explanation, and at the same time, it is only observed in the spot market, which suggests this relationship might be yet caused by the absence of short selling in the spot market. To investigate this further, we design a new test that separates positive and negative returns and we study them separately. The results show a higher trading volume for periods with negative returns which is in direct contradiction with Karpoff's (1987) explanation. A possible explanation for this anomaly is behavioral patterns, such as herding behavior in down market periods, which are more pronounced in the spot market due to lower investor sophistication in this market compared to the futures market (Poyser 2018). Finally, this anomaly might be unique to bitcoin markets because of their extremely speculative nature at this time.

Future studies on other asset classes can help clarify this contradiction. We contribute to the literature by finding an anomaly that cannot be explained by present theories. Understanding the nature of volume return relationship is important for at least three reasons. First, it will provide information on how financial markets function. Theories suggested to explain these relationships discuss the flow of information among traders, incorporation of the information in making trading decisions, and short sale considerations (Copeland 1976, Epps and Epps 1976, Karpoff 1987). It could also offer insight into behavioral reactions to price movements (Krafft, Della Penna et al. 2018). This study tries to examine the soundness of these theories by applying them to bitcoin spot and futures markets.

Second, these relationships can have huge impacts for assets traded in both the spot and the futures markets. Given the sharp increase in trading volume around delivery dates of futures contracts is associated with a significant change in futures prices, arbitrageurs will make sure that there is a spillover effect to the spot market price. This in turn can help address the issue of destabilizing effect of speculative assets on future prices (Rutledge 1979).

And third, this relationship can have important implications in event studies where return and trading volume are used jointly to determine the informational content of an event. Some event studies use volume as a measure for disagreement among investors, and return as the informational content of the event. If volume and return are coupled and determined together, event studies that use tests on volume and return jointly can utilize these relationships to provide more robust results (Beaver 1968).

The article is structured as follows: Part II will review select articles from the literature on price and volume relationships discussing the findings and theoretical explanations. Parts III and IV will explain the data and the methods used. Part V will discuss the findings and possible explanations, and part VI will conclude the paper.

Review of the literature

The literature on the relationship between volume and return is one of considerable size. Although this relationship has been studied over and over, there is little consensus on the relationship itself, let alone the theories explaining it. We can broadly classify the findings in two categories: The relationship between volume and the magnitude of price change, and the relationship between volume and price change *per se*. Karpoff (1987) summarizes most of the previous findings on these relationships.

1. Relationship between volume and the magnitude of price change

It is a common belief among practitioners that it takes a large volume to move prices significantly (Najand and Yung 1991). Even if this statement does not imply causality, at least it indicates a positive relationship between volume of trade and the magnitude of price change, i.e. the absolute value of price change. This relationship has been documented by many scholars for different asset classes, over different time intervals, and using different frequencies. Stock market indices have been studied and have been shown to maintain this relationship using different frequencies (Crouch 1970), from hourly (Jain and Joh 1988) to daily (Ying 1966). The relationship has been documented for common stocks using different frequencies, from trades (Epps and Epps 1976) to daily (Westerfield 1977), to monthly (Morgan 1976), to annual (Comiskey, Walkling et al. 1987). Moreover, the same relationship has been found to hold for futures contracts on commodities (Cornell 1981), currencies (Grammatikos and Saunders 1986), and T-bills (Tauchen and Pitts 1983).

On the other hand, Chen, Firth et al. (2001) report mixed results for price volume relationship in multiple national indices and Huang and Yang (2001), and Abba Abdullahi, Kouhy et al. (2014) find no evidence for a relationship between price and volume.

As for the causal[2] relationship between volume and price movements, Luu and Martens (2003), and Rashid (2007) find a one way relationship from volatility to volume. However, Gervais, Kaniel et al. (2001) find that lagged values of volume drive returns, and Hiemstra and Jones (1994) as well as Malliaris and Urrutia (1998) find a bidirectional relationship between return and volume.

[2] causality implies Granger causality throughout the article

Theoretical Explanations

The two prominent theories that try to explain this anomaly are the Epps (1975) "Mixture of Distribution Hypothesis" and the "Sequential Arrival of Information" of Copeland (1976).

- "*Mixture of Distributions Hypothesis"* indicates that volatility of price and trade volume are correlated for speculative assets. According to Clark (1973), Epps (1975), and Epps and Epps (1976) both of these variables have distributions based on one common underlying variable. This variable is interpreted as the flow of information to the market. This model suggests a contemporaneous relationship between volume and volatility since both of these variables are driven by introduction of new information.

- *"Sequential Arrival of Information"* model assumes there are two groups of investors, 'optimists' and 'pessimists' in the market and upon arrival of news they interpret it as 'good' and 'bad' news respectively. However, according to this model information is disseminated gradually and from one investor to another. Investors are assumed not to be able to infer the information from the change in the price. The model predicts that volume of trade depends on two factors: strength of news, and the mix of optimists and pessimists (or degree of disagreement in the market). Trading volume is greater for news with greater strength and in markets with higher agreement. In other words, if all of the investors are optimists or pessimists, then the trading volume would be maximum. This seems a little counterintuitive, because one would expect zero trades in a market with all investors sharing the same opinion. On the other hand, the model predicts that if the new information is accessible to everyone at the same time, then volume of trade would have a negative relationship with the magnitude of price change.

2. Relationship between volume and price change per se

Yet another belief among practitioners is that bull markets experience heavier trading volume while bear markets are accompanied by lower trade volume (Karpoff 1987). This would imply a positive relationship between volume and return per se (as opposed to magnitude of return). This relationship, too, has been heavily studied by scholars but is not found to be as significant. Some of the studies that have documented evidence of this relationship in common stocks are Morgan (1976), Epps (1977), Rogalski (1978), and Comiskey, Walkling et al. (1987). These studies use frequencies ranging from transactions to annual. Ying (1966) has found this relationship in stock market indexes and Epps (1975) reports the same relationship for NYSE bonds using transactions frequency. Nevertheless, most of these results show rather weak statistical significance (Karpoff 1987).

Theoretical Explanations

- Epps (1975) categorizes investors as either 'bulls' or 'bears'. Market demand depends entirely on bulls' demand while market supply depends solely on bears' supply. The model is based on investors' anchoring (also known as confirmation) bias, implying that interpretation of news will only lead to reinforcement of prior beliefs. Thus, each group of investors will only react to one type of news, bulls to 'good' news and bears to 'bad' news. The model indicates that in light of good news, bulls will react by changing their demand curve (and in turn market demand curve). This change includes an upward shift in the intercept and an increase in the absolute value of the slope of their demand curve. On the other hand, upon the arrival of bad news only bears react and will only change the intercept of their supply curve (and in turn market supply curve). This asymmetrical change in the slope of demand curve, will cause the volume from a positive price change to be greater than the volume from an equal but negative price change. In other words:

$$\frac{V^+}{\Delta P^+} > \frac{V^-}{\Delta P^-}$$

- Karpoff (1987) builds a model based on the asymmetry in the relationship between volume and price changes. The model assumes that volume of trade for negative returns is smaller than the volume of trade for equal but positive returns due to restriction (or costliness) of short positions. The fact that most studies on futures contracts fail to document this relationship borrows support to this model.

The study of relationship between volume and price in currency markets has not gained much attention (Kumar 2017). Bauwens, Rime et al. (2006), and Bjonnes, Rime et al. (2005) find support for the mixture of distributions hypothesis, while Hagiwara and Herce (1999), and Mougoué and Aggarwal (2011) find no evidence for this hypothesis. Moreover, there are very few studies focusing on the relationship between trading volume and return per se.

The bitcoin market is a rather new, and more importantly, unregulated market. While all the transactions are safely recorded in the blockchain, there are no legal (or widely accepted) debt instruments offered on bitcoin. This means that the volume of short selling in the bitcoin spot market is negligible, if not absolutely zero. At the same time, short selling in bitcoin futures market is costless (The costs associated with entering long and short contracts are symmetrical). This provides a unique opportunity to test whether the relationship between trading volume and return per se is observed because of limitations on short selling. If this is the case, according to Karpoff (1987), we would expect to see a positive significant relationship between volume and return in the spot market but not in the futures market.

There are fewer studies on this relationship in the context of currency markets and even then the results are not consistent. In the futures market, while Kumar (2017) finds support for the mixture of distribution hypothesis, Mougoué and Aggarwal (2011) fail to do so. As for the spot market, Hagiwara and Herce (1999) find no support for the mixture of distributions hypothesis but Bauwens, Rime et al. (2006), and Bjonnes, Rime et al. (2005) find evidence supporting the hypothesis.

Data

The first Bitcoin block (also known as the genesis block) was created on Jan. 3, 2009 by someone using the alias “Satoshi Nakamoto” and the first transaction occurred on Jan. 12, 2009. The price of one Bitcoin has increased from less than a hundredth of a cent in 2009 to near $20,000 in 2018 and has decreased to $3,820 at the time of writing this article. Figures 1 and 2 show the price and return, and Figure 3 shows the trading volume of Bitcoin over time. Several articles have associated Bitcoins performance to fundamental, macroeconomic, and speculative factors (Adjei 2019). While it is being adopted as a new type of currency by investors, it would be interesting to look at the price volume relationship in Bitcoin spot and futures markets.

[Insert Figure 1 About Here]

[Insert Figure 2 About Here]

The data for this study includes Bitcoin price and transaction volume for both spot and futures markets. Spot market data have been collected from bitcoincharts.com and span over the period from Dec 2014 to Sep 2018. In this study only data from Coinbase pro, formerly known as GDAX (Global Digital Asset Exchange), have been used. Coinbase pro is among the largest cryptocurrency exchanges in the world and operates in 32 countries. Bitcoin prices across different exchanges vary slightly, therefore making the use of data from multiple exchanges problematic. This study uses observations with transactions, 5 minutes, and hourly frequencies. Futures market data is obtained from tickdata.com and covers the period from Dec 2017 to June 2018. This data is gathered from CBOE futures exchange. Frequencies used in this study are transactions, 5 minutes, and Hourly. Each contract is on one Bitcoin and is cash settled.

The listed price at any time is for the contract with the nearest expiration date. Summary statistics for both spot and futures data are shown in tables 1 and 2.

[Insert Figure 3 About Here]

Summary statistics for Bitcoin spot market at the 5 minutes frequency are provided in panel B of table 1. Return for every 5 minute interval is calculated as a growth rate, i.e. by subtracting the log of open price from the log of close price. Mean return is zero with a standard deviation of 0.005 and return over each 5 minute period is positive in 44 percent of observations and negative in 49 percent of observations. Volume shows the total number of Bitcoins transacted over the 5 minute period and has a mean of 37.56 with a standard deviation of 69.17. As a test for robustness, natural logarithm of volume has also been used and has been found to have the same relationship structure with return.

Summary statistics at hourly frequency, provided in panel C of table 1 show the mean return to be pretty close to zero and the standard deviation of return to be around 0.01. The return is positive in 53 percent of observations and negative in 47 percent of observations. As for volume, the mean is 448.41 and the standard deviation is 649.70.

[Insert Table 1 About Here]

Panel B of table 2 shows the summary statistics for Bitcoin futures at the 5 minute frequency. Mean return is zero with a standard deviation of 0.0047. Return over each 5 minute interval is equally positive and negative in 34 percent of observations and is zero in almost 32 percent of observations. Volume is the number of contracts where each contract is for one bitcoin. Mean of number of contracts transacted is 10.18 with a standard deviation of 17.02.

Finally, Panel C of table 2 shows the summary statistics for Bitcoin futures at hourly frequency. The mean return is close to zero and standard deviation is 0.014. The price is up and down equally in half of the observations. Mean volume is 97.81 and the standard deviation of volume is 117.54.

[Insert Table 2 About Here]

Methodology

In the first step, vector autoregression (VAR) has been used to study the relationship between volume and price in terms of precedence. This will help us examine the causal relationship between volume and return. Based on AIC and BIC, we use a VAR (2) with two lags. Next, OLS and GARCH (1, 1) have been used to investigate the relationship between volume and price return. We use simple OLS, OLS with heteroskedasticity robust errors to correct for heteroskedasticity bias, as well as GARCH (1, 1) to correct for GARCH effects in the error terms.

We start by using a simple OLS regression to investigate the relationship. We regress volume of trade on both return per se and the absolute value of return. The OLS regressions are as follows:

$$Volume = \alpha + \beta . Return + \varepsilon \tag{I}$$

And

$$Volume = \alpha + \beta . |Return| + \varepsilon \tag{II}$$

Then, in order to further test the model suggested by Karpoff (1987) we use the following specification:

$$Volume = \alpha + \beta.down + \gamma.up \times Return + \delta.down \times |Return| + \varepsilon \qquad \text{(III)}$$

[Insert Figure 4 About Here]

Where down is a dummy which is assigned a value of 1 when the return is negative and a value of 0 otherwise. The coefficient of the dummy (down), shows the difference between the intercepts. Figure 4 helps explain the interpretation of the coefficients. The dashed line on the right (line I) shows the relationship between volume and positive returns. The dashed line on the left (line II) shows the hypothetical relationship between the volume and negative returns only (i.e. when the market is down and the dummy 'down' is equal to 1). The dashed line on the right (line II') is the mirror image of line II which is the regression line achieved by regressing volume on the absolute value of all negative returns. If Karpoff (1987) hypothesis is correct, one would expect to find the slope of line I to be significantly larger than that of line II'. That means the coefficient of the interaction of down and return (i.e. δ) should be smaller than that of the interaction of up and return (i.e. γ).

Furthermore, to take care of the serial correlation in the error terms, we also include lagged values of volume, down, return, and the interaction of down and return. Moreover, we run regressions (I, II, and III) using MLE estimation with a GARCH (1, 1) specification.

Referring back to figure 4, it is apparent that the line that represents regressing volume on magnitude of returns will lie somewhere between lines I and II'. Moreover, if we assume data points lie around lines I and II', then we would expect that magnitudes of error terms from the aforementioned regression should increase with the magnitude of return, hence we should expect heteroskedastic residuals. Karpoff (1987) mentions this as one of the testable implications of the explanation provided. However, it is important to observe that heteroskedastic error terms simply mean a difference in the slopes of lines I and II' and not the slope of line I being greater than that of line II'. In other words, while it helps to test whether the volume-return relationship is symmetrical for up and down returns, it does not necessarily imply that this asymmetry is caused by short selling constraints. Nevertheless, we conduct this test as well.

Finally, to further test whether the limitation on short selling is the cause of the volume - return relationship, we conduct one more test. For this test we want to use a proxy that can capture the effect of short selling constraints in the spot market and see whether this is the sole driver of the volume-return relationship. We know that the absence of short selling in the bitcoin spot market implies that in periods when transactions are initiated by buyers, we would expect higher participation in the market than in periods when transactions are initiated by sellers. Given the abundance of liquidity traders in the market makes we can assume that buy and sell orders have a symmetrical probability of being filled and therefore, the initiation of the order at any period will determine whether it is a seller market or a buyer market.

[Insert Table 3 About Here]

Results

We start by running a vector autoregression of two lags to help us determine the specification. The results of the VAR for bitcoin spot market are provided in table 3. For the spot market, we find a bidirectional (feedback) relationship at the 5 minute frequency, i.e. lagged values of either variable seem to help explain some of the variation in the other variable. However, at the hourly frequency the relationship seems to be one-directional. Lagged values of return help explain variation in volume, but not the other way around. This same one-way relationship can be seen in the VAR results of the futures market.

[Insert Table 4 About Here]

The futures market results are reported in table 4. At the 5 minute frequency, lagged values of return are statistically significant in explaining the variation in volume, but lagged volume has no explanatory power over variations in return. At the hourly frequency, while lagged values of return help explain variation in volume, lagged values of volume lack the explanatory power. These results suggest that it is best to pick volume as the dependent variable and return as the explanatory variable. To further verify this choice, we conduct a granger causality test of two lags and we report the results in table 5. These results indicate that at least one of the two first lags of return Granger causes volume. This confirms our specification and we move forward with the choice of volume as the dependent variable and return as the independent variable. The results for OLS regressions are reported in table 6.

[Insert Table 5 About Here]

In panel A of table 6 we see the results for regressions (I) and (II), i.e. regressing volume on the absolute value of returns and volume on return per se, in the spot market. Looking at specification (I), at both 5 minutes and hourly frequencies the beta is positive, and statistically and economically significant. The coefficient is statistically significant at the 0.01 level of significance. A one standard deviation change in the absolute value of returns, is associated with almost a 3 standard deviation change in volume, hence the economic significance. The R-squared of the regression is very low meaning there is not much explanatory power from the absolute value of returns alone. The positive coefficient provides support for the ‘Mixture of Distributions Hypothesis’ and also favors ‘Sequential Arrival of Information’ model over the ‘Simultaneous Arrival of Information’ model.

[Insert Table 6 About Here]

Turning our attention to specification (II), again, at both frequencies we document statistical and economical significance for beta. The coefficient is significant at the 0.01 level of significance and a one standard deviation change in return, is associated with a .42 standard deviation change in volume. The sign of the coefficient however, is negative. This is contrary to the evidence provided in the literature and suggests that negative returns are followed by higher trading activity in the Bitcoin market. This result does not support Karpoff (1987) model. This anomaly could be unique to the cryptocurrency, or even Bitcoin, markets. Again, the R-squared is very low which implies that return *per se* cannot explain much of the variation in trading volume.

Panel B of table 6 shows the results of regressions (I) and (II) in the spot market, correcting for GARCH (1, 1) effects. We see that including GARCH (1, 1) corrections do not change the significance or the sign of coefficients.

Panel C of table 6 reports the results of regressions (I) and (II) in the futures market. Results from regression (I), at both 5 minutes and hourly frequencies, still show positive and significant coefficients for the magnitude of return. This too, is evidence for the 'Mixture of Distributions Hypothesis' and 'Sequential Arrival of information'. The results of regression (II) in the futures market show insignificant coefficients for return per se in the futures market. This is generally consistent with the previous literature on volume and price relationship.

Finally, Panel D of table 6 reports the results for regressions (I) and (II) in the futures market, after correcting for GARCH (1, 1) effects. Again, the findings are robust to inclusion of GARCH (1, 1) corrections. Next we try to investigate the contradictory results we find from regression (II) in the spot market.

According to Karpoff (1987) explanation, and in line with previous literature one would expect to find a positive relationship between volume and return per se. In order to further investigate this, we first conduct the test proposed by Karpoff (1987). Referring back to figure III, the difference in the slopes of lines (I) and (II') means that the error terms from regression (II) should be heteroskedastic. Considering the fact that the V ~ |R| line should fall between these two lines, and assuming data points are scattered uniformly around these two lines, we expect the distance between the V ~ |R| line and data points, and hence the magnitude of error terms from regression (II), to increase as |R| increases. That would mean a correlation between the error terms from regression (II) and |R| which is conditional heteroskedasticity. Therefore, we can test the asymmetry in volume-return relationship by simply running a conditional heteroskedasticity test. The results of Breusch-Pagan heteroskedasticity test are presented in table 7. At both frequencies, the null hypothesis of "No conditional heteroskedasticity" is rejected at the 0.01 confidence level. These results can be interpreted in two ways: one is that there exists a relationship between volume and returns per se that is asymmetric for positive and negative returns, in both the spot and futures market. Another explanation for these results is that this is simply showing a flaw in our specification which is causing the heteroskedasticity. In any case it should be noted, again, that at best this test merely confirms that the nature of volume-return is asymmetric for positive and negative returns and does not necessarily show which one is larger and for what reason. Another helpful test would be to separate positive and negative returns and regressing volume on them separately.

[Insert Table 7 About Here]

Panel A of table 8 shows the results of regression (III) in the spot market. The significance of β implies a difference in the intercepts of lines (I) and (II), i.e. $\alpha_1 - \alpha_2$. More importantly, coefficients of both interaction terms are statistically significant. Moreover, the difference in the slopes of lines (I) and (II) are implied by the coefficients γ and δ. The results show that for the spot market, and at the 5 minute frequency, δ is greater than γ and that implies that the slope of line (II) is larger than that of line (I), i.e. the line representing negative returns is steeper than the one representing positive returns. While this finding is consistent with the negative β we find from running regression (I), it is not consistent with the findings of previous studies and it does not support the model proposed by Karpoff (1987).

[Insert Table 8 About Here]

In a similar model lagged values of all variables, including the dependent variable, have been added to the specification to take care of the serial correlation between the error terms from regression (III). Still the coefficients of interest, γ and δ, are both statistically and economically significant. Moreover, δ is still greater than γ consistent with the negative β found in regression (I), and contradicting Karpoff (1987) explanation. Generally, the findings of this test are robust to the choice of either of the specifications we have used.

In panel B of table 8 we report the results of regression (III) after correcting for GARCH (1, 1) effects. The significance and signs of the coefficients do not change and the firings are again robust to inclusion of GARCH (1, 1) effects.

Looking at panel C of table 8 and for the futures market, we find γ and δ from regression (III) to be also significant and somewhat the same. In all instances δ is greater than γ which is what we would expect in the presence of short selling constraints. However, the difference between δ and γ diminishes as the frequency of observations decreases from 5 minutes to hourly.

In panel D of table 8 GARCH (1, 1) effects have been included in the model. The difference between δ and γ is even larger for this specification which is, again, what we would expect in the presence of short selling limitations. These results show that the volume returns relationship is asymmetrical for positive and negative returns even in the futures market where short selling is subject to no restriction.

According to Karpoff (1987) explanation, in the absence of short-selling (or in the case of costliness of short-selling) one would expect to find δ to be smaller than γ, and that in turn would be the reason why volume and return may have a positive and significant relationship. In Bitcoin spot market, where short selling is not allowed, we find the opposite of what is predicted by this model. In the futures market we do not find a significant relationship between volume and return, but we find δ to be smaller than γ. It can be inferred from these results that the relationship we find in specification (II) is not necessarily due to a restriction in short selling. The tests conducted in this study cannot explain the reason behind this relationship. However, these results imply that current theories cannot explain this relationship. Finally, the results observed in this study could be specific to bitcoin or cryptocurrency markets and conducting the same tests in other markets would help clarify this matter.

Conclusion

We revisit the volume-return relationship for Bitcoin spot and futures markets. Many scholars have studied this relationship in various contexts and as a result there is a sizable literature available on this subject. However, not all of the evidence are consistent (Kumar 2017). Moreover, this relationship has not been the focus of many studies in currency markets. Finally, the Bitcoin market is a new market and very few studies have been conducted on its characteristics. Our results suggest that there is a positive and significant (both economically and statistically) relationship between volume and the absolute value of return. This gives support to Copeland (1976) sequential arrival of information theory. Moreover, we find evidence that volume also has a positive and significant relationship with volatility (measured by the standard deviation of returns), which lends support to Epps (1975) mixture of distributions hypothesis. Our results concerning the relationship between volume and return *per se* however, are not in line with previously documented results in the literature.

The negative and significant relationship between volume and return per se is not documented by previous studies. Most scholars working on the subject have documented positive relationships and have attributed this result to the restrictions, or costliness of short selling in the spot market (Karpoff 1987). According to Karpoff (1987), expensive short selling causes the volume of trade for negative returns to be lower than that for positive returns. This should be reflected in a smaller coefficient for negative returns compared to the coefficient for positive returns. Thus aggregating all observations, positive and negative returns, should result in a positive correlation between volume and return *per se*. The negative results from this study, however, do not support this explanation. We further test this model by using dummy variables to separate positive and negative returns. Although we find different coefficients for positive and negative returns, the larger coefficients for negative returns imply the exact opposite of Karpoff (1987) model. This implies that volume of trade was higher for negative returns.

We also study the above-mentioned relationships in Bitcoin futures market. The relationship between volume and magnitude of return shows the same results supporting both Epps (1975) mixture of distributions hypothesis and Copeland (1976) sequential arrival of information. As for the relationship between volume and return *per se,* the relationship is not significant in Bitcoin futures market. This implies that the relationship between volume and return per se is unique to the spot market and most likely happens due to differences between the spot and futures markets. However, our results rule out the restriction of short selling as the reason for this relationship. This relationship could be rooted in behavioral biases and exclusive to the spot market because of the lower level of investor sophistication. Alternatively, it could be due to algorithm trading and speculation. This calls for more investigation into Bitcoin markets as well as studies in other asset classes and time periods.

We contribute to the literature by studying the relationship between volume and return in Bitcoin markets. This will add to the young but rapidly growing finance literature on Bitcoin markets. Furthermore, we find support for the mixture of distributions hypothesis and sequential arrival of information theory. We also confirm Karpoff (1987) assertion that the relationship is asymmetrical between positive and negative returns, but find no support for short selling being the reason for the relationship between volume of trade and return per se.

ESSAY TWO: INVESTOR SENTIMENT AND ITS EFFECT ON BITCOIN PRICES

Introduction

Since its introduction in 2009, bitcoin has gained a lot of attention, not only from individual investors with speculation intent, but also from businesses trying to hedge their bitcoin denominated income in the futures market, and sophisticated institutional investors such as hedge funds. In the aftermath of the financial crisis of 2008, the idea of bitcoin was presented by someone under the alias of Satoshi Nakamoto as a solution to an old problem: Double spending in digital currencies. Using a distributed peer to peer network system, bitcoin not only addresses the issue of double spending, but also offers users some anonymity, also referred to as pseudonymity, in transactions (Nakamoto 2009). Moreover, a distributed network system implies no need for a central governing body. All of this helped make bitcoin an appealing transactions system to a certain type of investors at the time of its inception. However, later on, and after its price surge in 2014 and again at the end of 2018, more and more investors are showing interest in bitcoin, not as an alternative transactions system, but rather as a new asset class that provides a new and unique investment opportunities (Glaser, Zimmermann et al. 2014). Also, the accelerating rise of bitcoin price in 2014 and again in 2018, gained the attention of so many unsophisticated individuals with speculative investment attitudes. But investors are not the only group of people who are interested in bitcoin. As a new asset class, with its unique and unknown features, it has also made finance and economic scholars interested. There have been a few studies on bitcoin trying to explain different features and characteristics of bitcoin. Ciaian, Rajcaniova et al. (2016) show that bitcoin price is not driven by macro-financial developments. Cheah and Fry (2015) find the fundamental price of bitcoin to be zero and attribute its current price to a speculative bubble in the market. In spite of all the research that has been done on the subject, there is no agreement on financial and economic fundamentals of bitcoin price. This also means there is no established way of forecasting bitcoin prices. Nevertheless, most studies on bitcoin mention speculative prices, and bubbles (Yermack 2015) and this hints to behavioral anomalies in the market. Most notably, Garcia, Tessone et al. (2014)relate bitcoin return patterns to social interactions, especially those occurring on social media networks. This has led us to examine the relationship between sentiment and emotions among bitcoin investors, or enthusiasts, and bitcoin returns.

In order to do so, we use five sentiment measures provided by Thomson Reuters MarketPsych Indices (TRMI) and test for their relations, both contemporaneous and causal, with bitcoin returns. These sentiment indices have been created using a lexical analysis over texts relating to bitcoin, which are downloaded from news and social media platforms. We use simple ordinary least squares and control for GARCH (1, 1) effects. We also use vector autoregression to check for causal relationships[3]. We find that while our sentiment measures do indeed affect bitcoin returns, they do not explain much of the variation in the return, and therefore, cannot be used to forecast bitcoin prices. Our findings confirm most of the previous findings both on bitcoin price patterns and on behavioral finance and market sentiment studies.

Following the behavioral finance literature, we also look at the relationship between sentiments with trading volume and return volatility and find that generally positive sentiment leads to periods of lower return volatility and lower trading volume, and that negative sentiment usually leads to periods with high trading volume.

We contribute to the literature by trying a behavioral approach to modelling bitcoin returns. We establish that investor sentiments, as gathered through news and social media networks does in fact affect bitcoin prices, but that it does not account for much of the return patterns, and therefore, cannot be effectively used to forecast the price of bitcoin.

[3] Causal relationship does not imply causality, but rather chronological precedence, i.e. Granger causality.

Review of the Literature

Even though it is a new phenomenon, bitcoin has been studied by many finance and economics scholars. The majority of these studies try to examine the nature of bitcoin as a new asset class. Dyhrberg (2016) looks at diversification benefits of bitcoin and classifies bitcoin as something in between gold and the US dollar. Yermack (2015) takes a more fundamental approach too classifying bitcoin. He asserts that bitcoin is not a useful medium of exchange, does not show any diversification benefits, and is not used to denominate any financial debt instruments, and claims it behaves more like a speculative asset than a currency. Glaser, Zimmermann et al. (2014) look at the intentions behind bitcoin investments and argue that bitcoin is mostly desired not as a currency, but rather an alternative investment asset. At the same time some scholars try to look at price formation and return patterns in the bitcoin market. Ciaian, Rajcaniova et al. (2016) use traditional determinants of currency price to model bitcoin price and conclude that macro-financial developments do not in fact drive bitcoin prices. Many studies document the lack of a fundamental valuation for bitcoin and point to the existence of speculative bubbles in the bitcoin market. Cheah and Fry (2015) argue that the fundamental value of bitcoin is zero and attribute its price completely to speculative bubbles. Cheung, Roca et al. (2015) use generalized sup augmented Dickey-Fuller test and find two major bubbles in bitcoin price patterns until 2015. Some studies take a different approach and use information from bitcoin investors to try and gain insight into the return patterns of bitcoin. Yelowitz and Wilson (2015) analyze Google Trends data to find out the intent behind bitcoin investments. They claim that computer programming enthusiasts and criminals comprise the majority of participants in bitcoin markets. Garcia, Tessone et al. (2014) find that word-of-mouth communication on social media networks, as well as growth of new adopters can explain bitcoin’s rapid returns. This highlights an opportunity to use the collective behavioral foot print of bitcoin investors to understand its return patterns. Market sentiment derived from social media can help make this a possibility.

There are a large number of studies on investor sentiments, since it gained attention in the past two decades. The sentiment indicators employed in these studies can be broadly categorized in three groups: economic-based sentiment, survey-based sentiment, and media-based sentiment. Survey-based sentiment measures can be grouped as: 1) sentiment surveys from American Association of Individual Investors (AAII) and 2) sentiment surveys development by using algorithmic content analysis. Of the latter group media-based sentiment indicators from textual analysis of news stories and social media blogs have been extensively used for sentiment research.

Two common methods often used in textual analysis are the dictionary-based approach and machine learning (Kearney and Liu 2014). General Inquirer and Diction are two most widely used software that use built-in dictionaries in English language. Moreover, public media platforms provide opportunities for financial analysts and researchers to measure market sentiment through analyzing their contents using these powerful software. Borovkova (2011) and Borovkova and Mahakena (2015) have used news sentiment on crude oil and natural gas provided by Thomson Reuters News Analytics to study their futures price movements. Additionally, Smales (2015) used sentiment data provided by Thomson Reuters News Analytics to study the futures price of gold. They find significant relationship between news sentiments and returns on these commodity futures contracts. Today, Thomson Reuters MarketPsych Indices (TRMI) are available on many assets classes including cryptocurrencies.

The effect of investors' emotional and cognitive biases on their decision making processes has been documented by many scholars in different markets. Hirshleifer and Shumway (2003) have reported a statistically significant relationship between morning sunshine and daily market index returns in 26 countries. Goetzmann, Kim et al. (2014) introduced weather-based indicators of mood, and found that cloudy weather affects stock purchase, stock returns (overpricing), and stock return co-movements. Blasco, Corredor et al. (2012) find that in periods with high stress, investors rush towards the market portfolio. Moreover, Edmans, Garcia et al. (2007) have documented the effect of a negative general mood, arising from a soccer match win/loss, on a significant market decline.

Of the different sentiment measures used, optimism, joy, fear, and gloom have been most widely used in previous studies. Perhaps the most commonly documented important sentiment measure in the finance literature is fear. Financial crises often cause a spike in fear levels in the future market which in turn can lead to a slowdown in financial recovery, or even more turmoil in the market (Rubbaniy, Asmerom et al. 2014). It has been proxied by different implied volatility indices and it has been shown to be useful in forecasting stock returns (Esqueda, Luo et al. 2015). Da, Engelberg et al. (2014) create a daily "Financial and Economic Attitudes Revealed by Search" (FEARS) index based on households' internet searches and found that the FEARS index can predict asset prices reversals, volatility, and flow of funds. Often times, after going through a long-term recession investors become pessimistic in the stock market. The gloomy stage of a market downturn may take a long time to reverse because investors are more sensitive to the more fragile market (Lauricella 2011). Azzi and Bird (2005) find a relation between periods of boom and gloom, and market analysts' recommendations with more recommendations encouraging high momentum growth investment styles in periods of boom.

Karabulut (2013) uses Facebook's Gross National Happiness (GNU) index to proxy for happiness and found a statistically and economically significant relationship between the index and market returns in the following day.

In the context of finance literature, optimism can be interpreted as overestimation of future financial outcomes, which immediately leads to overpricing (Balasuriya, Muradoglu et al. 2013).

The relationship between sentiment measures and asset prices often tend to be bidirectional. (Brown and Cliff 2004) find that stock returns granger cause sentiments, and that sentiments cannot be used to forecast stock returns, while Verma, Baklaci et al. (2008) use sentiments to successfully forecast stock returns. Some studies have used sentiment measures to forecast index returns. Kurov (2010) examines the role of noise traders in stock index futures and its returns. Simon and Wiggins III (2001) use market-based sentiment measures to predict S&P 500 index returns and find statistically and economically significant forecasting power. The use of sentiment measures in finance is not limited to asset returns. Gao and Süss (2015) find a strong relationship between sentiment measures and the co-movement among commodity futures. They also find that periods with high optimism subsequently lead to an increase in liquidity and trading volume. Some researchers have looked into the relationship between sentiments and price volatility. Verma and Verma (2007) find a relation between sentiments and stock return volatility. Wang, Keswani et al. (2006) document that in predicting price volatility, adding lagged values of return renders the forecasting power of sentiments insignificant.

In this study we use selected Thomson Reuters MarketPsych Indices (TRMI) on cryptocurrencies to study the relationship between bitcoin sentiments and bitcoin prices. Recently, Sun, Najand et al. (2016) have used sentiment data on stocks provided by TRMI to predict the S&P 500 index returns. In this study we use five sentiment measures that reflect the general public attitude toward bitcoin. The focus of study will be on the relationship between these sentiment measures and bitcoin returns. However, in the later section we also investigate the relationship of sentiment measures with trading volume and volatility.

Data and Methodology

In this study we use data from MarketPsych of Thomson Reuters. The Thomson Reuters MarketPsych Indices (TRMI) are updated every minute and are derived from a collection of premium news outlets, global internet news coverage, and a broad group of social media (Peterson 2013). TRMI utilizes a lexical analysis paired with a complex grammatical framework on contents derived both from news and social media platforms to reflect sentiments from both professional and individual investors. For the first category, MarketPsych sources of input text include The New York Times, The Wall Street Journal, Financial Times, Seeking Alpha and dozens more sources available to professional investors. Less formal news sources are obtained from Yahoo! and Google News. For the second category, TRMI utilizes over two million social media websites including StockTwits, Yahoo! Finance, Blogger, chat rooms and other sources.

MarketPsych employs lexical analysis to extract sentiment indices by scrapping all sources every minute, and this can add up to over two million news articles and posts every day. Each sentiment index measure is the weight of words and phrases related to that sentiment in the overall text from the combined news and social medial content and each minute value is a simple average of the past 24 hours, or 1440 minutes, of information(Peterson 2013). Thus, the TRMI represent an unmatched collection of premium news and a broad range of social media. There are 49 sentiment indices defined related to bitcoin, out of which we are going to use only five. These are overall sentiment, optimism, joy, fear, and gloom. Overall sentiment shows the overall attitude of market participants, measured as positive references to bitcoin net of negative references, in the abovementioned sources. Optimism reflects positive attitude of market participants and is measured as mentions of or references to optimism, net of those related to pessimism. In the same fashion, joy, fear, and gloom are also calculated as mentions or references to these emotions net of negative references.

[Insert Table 9 About Here]

Price and volume data for the spot market have been collected from bitcoincharts.com and span over the period from Mar 2018 to Dec 2018. In this study only data from Coinbase pro, formerly known as GDAX (Global Digital Asset Exchange), have been used. Coinbase pro is among the largest cryptocurrency exchanges in the world and operates in 32 countries. Bitcoin prices across different exchanges vary slightly, therefore making the use of data from multiple exchanges problematic. Given the fluctuations in TRMI values in high frequencies, and assuming sentiments of market participants are somewhat more stable than what can be measured using textual analysis of news and social media platforms, this study uses observations with hourly frequency only. Futures market data is obtained from tickdata.com and covers the period from Mar 2018 to June 2018. This data is gathered from CBOE futures exchange. Frequency of observations is hourly. Each contract is on one Bitcoin and is cash settled. The listed price at any time is for the contract with the nearest expiration date. Table 9 shows a summary statistics of the data used in this study. Panel A shows descriptive statistics of the spot market, and panel B shows those of the futures market. The five sentiment measures have a range of -1 to 1. Optimism and joy are regarded as positive sentiments since they reflect positive attitudes toward the asset, while fear and gloom are negative sentiments. Overall sentiment shows the overall feeling toward the asset in the public. Since these measures have been developed using a lexical analysis, and since there exists a considerable overlap between the keywords associated with different sentiment measures, it is important to look at the correlations between these measures. Table 10 shows the correlation matrix between these sentiment measures, return, trading volume, and return volatility. It can be seen that there is a high correlation between overall sentiment and other measures of sentiment in both the spot market (panel A) and the futures market (panel B). Overall sentiment measure is highly correlated with optimism (0.69 in the spot market and 0.64 in the futures market). It is also negatively correlated with gloom (-0.40 in the spot market and -0.37 in the futures market)[4]. These high correlations between independent variables could be problematic in a regression. Therefore, in order to ensure our regression results are robust to the inclusion/exclusion of sentiment variables, we use three different specifications. Model 1 uses all five sentiment measures as independent variables, model 2 only uses the overall sentiment measure, and model 3 uses the other four as independent variables.

[Insert Table 10 About Here]

We are interested in investigating the contemporaneous, as well as the causal relationship between sentiments and bitcoin prices. To achieve that we run ordinary least squares regressions, vector autoregressions, and granger causality tests. We also correct for GARCH (1, 1) effects to make sure our results are robust given heteroskedastic error terms.

Results

We begin by looking at the contemporaneous relation between sentiment and returns. Table 11 shows the results of regressing returns on sentiment measures using our three different models in both the spot and the futures market. Panel A shows the OLS results in the spot market. Overall sentiment and joy are the two measures that are significant across different specifications. Overall sentiment has a positive coefficient and is significant at the 10% level. This is consistent with the findings of Verma and Soydemir (2009). The effect of joy is also positive and significant at the 10% in model 1 and 5% in model 2. Correcting for GARCH (1, 1) effects, however, changes these relationships. Looking at panel B, after correcting for

[4] All of the pairwise correlations between sentiment measures are significant at the 1% level.

GARCH (1, 1), the only sentiment measure with a significant coefficient is fear. The negative and statistically significant effect of fear on returns is consistent with the findings of Griffith, Najand et al. (2019). Panel C shows the regression results in the futures market. Here optimism and gloom both have positive and significant effects on returns. The coefficient of optimism is significant at the 10% in model 1 and 1% in model 3. The coefficient of gloom is positive and significant at the 5% in model 1 and 10% in model 3. Again, however, introducing GARCH (1, 1) effects changes the significance of these effects. After controlling for GARCH (1, 1) effects (Panel D of table XI), fear is the only sentiment measure with a significant and this time positive effect on returns. Fear seems to have a positive contemporaneous effect on spot prices and at the same time a negative effect on futures prices. Its effect is robust to GARCH (1, 1) corrections and inclusion/exclusion of the highly correlated overall sentiment measure as an independent variable. Next, we look at the causal relationship between sentiment measures and return.

[Insert Table 11 About Here]

Table 12 shows the results of regressing return on lagged values of sentiment measures. Looking at panel A, overall sentiment is again positively and significantly correlated with subsequent returns. However, after correcting for GARCH (1, 1) (Panel B), the statistical significance of this relationship disappears. However, fear and gloom seem to be statistically significant after correcting for GARCH (1, 1). Fear has a strong positive effect on subsequent returns that is significant at the 1%. The positive sign shows a reversal of sign from the contemporaneous results, where we found a negative relationship. This short term reversal is significant with the findings of Tetlock (2007). The effect of gloom is negative and significant at the 5% level which is consistent with most of the literature. Panel C shows the results of regressing return on lagged values of sentiment measures in the futures market and Panel D shows the same results corrected for GARCH (1, 1) effects. There is no statistically significant relationship found between the sentiment measures and subsequent returns in the futures market. In order to further investigate the causal relationship between sentiment measures and returns, we run a vector autoregression.

[Insert Table 12 About Here]

Table 13 shows the results of the vector autoregression for both the spot and futures markets. Looking at panel A, we see that lags of overall sentiment measure and fear both drive returns. The first lag of overall sentiment has a positive effect that is significant at the 5% in model 1 and 10% in model 2. The first lag of fear also has a strong positive effect that is significant at the 1% level in both models 1 and 3. From panel B, it can be seen that the sentiment measures do not have any robust significant causal effect on futures prices. The positive significant effect of overall sentiment on returns is in line with findings of previous studies documented in the literature, as well as our expectations. However, our results indicate that fear, as a negative sentiment, also has a positive effect on returns. Although the sign changes for the second lag, this is still puzzling and warrants further attention. Next, we are going to look at the effects of our sentiment measures on trading volume and return volatility.

[Insert Table 13 About Here]

Table 14 summarizes the vector autoregression results for volume and the five sentiment measures. From panel A, results for the spot market, it can be seen that both lags of overall sentiment have a strong negative effect on volume. The coefficient for the first lag is significant at the 1% level under both models 1 and 2. The negative and statistically significant second lag in both models implies this effect is persistent. Although these results are in contrast to the findings of Gao and Süss (2015), They are consistent with our previous findings and the results reported in our first essay. Previously, we found a positive causal relationship between overall sentiment and returns. Moreover, in the first essay we observed that periods of price increase are associated with lower trading volume. Therefore, the negative causal effect of overall sentiment on trading volume is expected. The first lag of gloom appears to have a strong positive effect on trading volume. Again, considering that gloom is a negative sentiment and is negatively, although statistically insignificantly (See table XIII), related to subsequent returns, these results are consistent with our previous findings. The second lag of fear has a positive and significant effect on trading volume. Although this is consistent with our previous findings on the causal relation between fear and returns, we do not observe the same sign changing behavior in the short term.

[Insert Table 14 About Here]

In panel B of table 14, results for the futures market, fear is the only sentiment measure with a statistically significant coefficient. The effect of fear on futures returns are negative and significant at the 10% level. This implies that fear will subsequently lead to less transaction volume. It is interesting that while fear does not affect futures returns, it does affect futures transaction volumes. This might have to do with the uncertainty that exists in the market concerning the outlook of bitcoin.

[Insert Table 15 About Here]

We now turn our attention to volatility. Table 15 shows the results of vector autoregression for return volatility and sentiment measures. In panel A, for the spot market, optimism has a negative and significant effect on volatility. The coefficient for the first lag is significant at the 5% level only in model 3, and that of the second lag is significant at the 1% level in both models 1 and 3. This means optimism is negatively associated with our measure observed market risk, i.e. return volatility. Moreover, first lag of joy also seems to have a strong positive and statistically significant causal effect on volatility. While joy, as a positive sentiment, has mostly positive, although statistically insignificant, effects on return, it does cause observed risk to increase. Finally panel B of table 15 shows vector autoregression results in the futures market. Here, too, the second lag of optimism is negatively associated with volatility. In other words, high optimism in both the spot and the futures market leads to periods with lower return volatility. Overall sentiment also has negative coefficients. The coefficient for the second lag of overall sentiment, in model 2 is statistically significant at the 5% level. This result is consistent with the effects we observed in the spot market. A general observation from comparing results from the spot market and the futures market is that return, trading volume, and return volatility, all have stronger associations with sentiments in the spot market, which is probably due to the high presence of noise traders who not only reflect these sentiments in social media platforms, but also are more affected by such general sentiments due to lower sophistication and higher behavioral biases.

Conclusion

We use five sentiment measures provided by Thomson Reuters MarketPsych Indices to study the relationship between bitcoin return, trading volume, and return volatility with market sentiment. By doing so, we aim to achieve two goals. First, to try and test whether market sentiment does in fact affect bitcoin prices. Given that most of the investors in bitcoin markets are unsophisticated investors with speculative attitudes toward investing, it is more likely to find a meaningful relationship between sentiment measures and bitcoin prices, rather than other asset classes with more robust financial and economic fundamentals.

Second, we examine forecasting capability of the sentiment measures we are using in bitcoin markets. Bitcoin is a rather new currency, and while the number of studies exploring its financial characteristics is growing rapidly, there is still no solid ground upon which one can fundamentally analyze bitcoin price movements. Since we lack the economic insight to forecast bitcoin prices based on fundamentals, and considering the existence, and abundance, of unsophisticated investors in bitcoin markets, using market sentiments could prove to be promising in forecasting bitcoin prices.

We document that sentiment measures do influence subsequent bitcoin prices. More specifically, fear has a powerful effect on subsequent returns in the spot market. This is consistent with the findings reported by Griffith, Najand et al. (2019). Moreover, overall sentiment has a positive relationship with returns in the spot market. However, the relationship disappears after controlling for GARCH (1, 1) effects. This is consistent with the results documented by Wang et al. (2006). Our sentiment measures do not significantly affect results in the futures market, which is probably due to higher level of investor sophistication. In our VAR tests we also noted that return has a weak causal effect on all five of our sentiment measures. This implies a bidirectional relationship between sentiment and return and is therefore consistent with the findings of Brown and Cliff (2005).

Our sentiment measures do not explain much variation in bitcoin returns. $R^2$of all regressions is extremely low ($R^2 < 0.01$) meaning although the effect of sentiments is statistically significant, it cannot be used for forecasting the return of bitcoin.

We also look at the relationship between sentiment and trading volume and find that in the spot market while overall sentiment has a negative effect on trading volume, gloom and fear positively affect volume. Of the three, gloom has the strongest effect. At the same time, in the futures market, fear has a negative effect on trading volume. These results do not confirm the findings of Gao and Süss (2015).

Finally, we look at the relation between sentiments and return volatility. In both spot and futures markets overall sentiment and optimism tend to decrease return volatility. Moreover, in the spot market joy seems to cause a subsequent increase in return volatility. These findings are consistent with the results of Verma and Verma (2007).

We contribute to the literature in two ways. First, we look at the relation between sentiments and returns in bitcoin markets and while we confirm most of the previous findings in the realm of behavioral finance, we also add to it by exploring this relationship in the context of a new asset class, i.e. bitcoin.

Second, we try to forecast bitcoin prices using only previous returns and sentiments and find that our sentiment measures lack the capability to forecast bitcoin prices. While these results show the ineffectiveness of our sentiment measures in predicting bitcoin prices, using sentiments and behavioral approaches might still be a feasible approach to modelling and forecasting bitcoin prices. This can be tested by using other sets of sentiment measures or other technics of creating the sentiment indices.

References

Abba Abdullahi, S., et al. (2014). "Trading volume and return relationship in the crude oil futures markets." Studies in Economics and Finance **31**(4): 426-438.

Adjei, F. (2019). "Determinants of Bitcoin Expected Returns." Journal of Finance and Economics **7**(1): 42-47.

Azzi, S. and R. Bird (2005). "Prophets during boom and gloom downunder." Global Finance Journal **15**(3): 337-367.

Back, A. (2002). "Hashcash-a denial of service counter-measure."

Balasuriya, J., et al. (2013). "Is It Better to Be Optimistic?--Financial Optimism and Well-Being." Financial Optimism and Well-Being (July 1, 2013).

Bauwens, L., et al. (2006). "Exchange rate volatility and the mixture of distribution hypothesis." Empirical Economics **30**(4): 889-911.

Beaver, W. H. (1968). "The information content of annual earnings announcements." Journal of accounting research: 67-92.

Bjonnes, G., et al. (2005). "Volume and volatility in the foreign exchange market: does it matter who you are." Exchange Rate Economics: Where Do We Stand: 39-62.

Blasco, N., et al. (2012). "Does herding affect volatility? Implications for the Spanish stock market." Quantitative Finance **12**(2): 311-327.

Borovkova, S. (2011). "News analytics for energy futures." Available at SSRN 1719582.

Borovkova, S. and D. Mahakena (2015). "News, volatility and jumps: the case of natural gas futures." Quantitative Finance **15**(7): 1217-1242.

Brands, S. A. (1997). Secure cryptographic methods for electronic transfer of information, Google Patents.

Brown, G. W. and M. T. Cliff (2004). "Investor sentiment and the near-term stock market." Journal of empirical finance **11**(1): 1-27.

Brown, G. W. and M. T. Cliff (2005). "Investor sentiment and asset valuation." The Journal of business **78**(2): 405-440.

Camenisch, J., et al. (2005). Compact e-cash. Annual International Conference on the Theory and Applications of Cryptographic Techniques, Springer.

Chaum, D. (1983). Blind signatures for untraceable payments. Advances in cryptology, Springer.

Chaum, D., et al. (1988). Untraceable electronic cash. Conference on the Theory and Application of Cryptography, Springer.

Chaum, D. and T. P. Pedersen (1992). Wallet databases with observers. Annual International Cryptology Conference, Springer.

Cheah, E.-T. and J. Fry (2015). "Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin." Economics Letters **130**: 32-36.

Chen, G. m., et al. (2001). "The dynamic relation between stock returns, trading volume, and volatility." Financial Review **36**(3): 153-174.

Cheung, A., et al. (2015). "Crypto-currency bubbles: an application of the Phillips–Shi–Yu (2013) methodology on Mt. Gox bitcoin prices." Applied Economics **47**(23): 2348-2358.

Ciaian, P., et al. (2016). "The economics of BitCoin price formation." Applied Economics **48**(19): 1799-1815.

Clark, P. K. (1973). "A subordinated stochastic process model with finite variance for speculative prices." Econometrica: Journal of the Econometric Society: 135-155.

Comiskey, E. E., et al. (1987). "Dispersion of expectations and trading volume." Journal of Business Finance & Accounting **14**: 229-239.

Copeland, T. E. (1976). "A model of asset trading under the assumption of sequential information arrival." the Journal of Finance **31**(4): 1149-1168.

Cornell, B. (1981). "The relationship between volume and price variability in futures markets." Journal of Futures Markets **1**(3): 303-316.

Crouch, R. L. (1970). "The volume of transactions and price changes on the New York Stock Exchange." Financial Analysts Journal **26**(4): 104-109.

Da, Z., et al. (2014). "The sum of all FEARS investor sentiment and asset prices." The Review of Financial Studies **28**(1): 1-32.

Davies, G. (2010). History of money, University of Wales Press.

Dwork, C. and M. Naor (1992). Pricing via processing or combatting junk mail. Annual International Cryptology Conference, Springer.

Dyhrberg, A. H. (2016). "Bitcoin, gold and the dollar–A GARCH volatility analysis." Finance Research Letters **16**: 85-92.

Edmans, A., et al. (2007). "Sports sentiment and stock returns." the Journal of Finance **62**(4): 1967-1998.

Epps, T. W. (1975). "Security price changes and transaction volumes: Theory and evidence." The American Economic Review **65**(4): 586-597.

Epps, T. W. (1977). "Security price changes and transaction volumes: Some additional evidence." Journal of Financial and quantitative Analysis **12**(1): 141-146.

Epps, T. W. and M. L. Epps (1976). "The stochastic dependence of security price changes and transaction volumes: Implications for the mixture-of-distributions hypothesis." Econometrica: Journal of the Econometric Society: 305-321.

Esqueda, O. A., et al. (2015). "The linkage between the US “fear index” and ADR premiums under non-frictionless stock markets." Journal of Economics and Finance **39**(3): 541-556.

Ferguson, N. (1993). Single term off-line coins. Workshop on the Theory and Application of of Cryptographic Techniques, Springer.

Gao, L. and S. Süss (2015). "Market sentiment in commodity futures returns." Journal of empirical finance **33**: 84-103.

Garcia, D., et al. (2014). "The digital traces of bubbles: feedback cycles between socio-economic signals in the Bitcoin economy." Journal of the Royal Society Interface **11**(99): 20140623.

Gervais, S., et al. (2001). "The high-volume return premium." the Journal of Finance **56**(3): 877-919.

Glaser, F., et al. (2014). "Bitcoin-asset or currency? revealing users' hidden intentions." Revealing Users' Hidden Intentions (April 15, 2014). ECIS.

Goetzmann, W. N., et al. (2014). "Weather-induced mood, institutional investors, and stock returns." The Review of Financial Studies **28**(1): 73-111.

Grammatikos, T. and A. Saunders (1986). "Futures price variability: A test of maturity and volume effects." Journal of Business: 319-330.

Griffith, J., et al. (2019). "Emotions in the Stock Market." Journal of Behavioral Finance: 1-15.

Hagiwara, M. and M. A. Herce (1999). "Endogenous exchange rate volatility, trading volume and interest rate differentials in a model of portfolio selection." Review of International Economics **7**(2): 202-218.

Harris, L. and E. Gurel (1986). "Price and volume effects associated with changes in the S&P 500 list: New evidence for the existence of price pressures." the Journal of Finance **41**(4): 815-829.

Hiemstra, C. and J. D. Jones (1994). "Testing for linear and nonlinear Granger causality in the stock price-volume relation." the Journal of Finance **49**(5): 1639-1664.

Hirshleifer, D. and T. Shumway (2003). "Good day sunshine: Stock returns and the weather." the Journal of Finance **58**(3): 1009-1032.

Huang, B.-N. and C.-W. Yang (2001). "An empirical investigation of trading volume and return volatility of the Taiwan Stock Market." Global Finance Journal **12**(1): 55-77.

Jain, P. C. and G.-H. Joh (1988). "The dependence between hourly prices and trading volume." Journal of Financial and quantitative Analysis **23**(3): 269-283.

Jakobsson, M. and A. Juels (1999). Proofs of work and bread pudding protocols. Secure Information Networks, Springer**:** 258-272.

Karabulut, Y. (2013). Can facebook predict stock market activity? AFA 2013 San Diego Meetings Paper.

Karpoff, J. M. (1987). "The relation between price changes and trading volume: A survey." Journal of Financial and quantitative Analysis **22**(1): 109-126.

Kearney, C. and S. Liu (2014). "Textual sentiment in finance: A survey of methods and models." International Review of Financial Analysis **33**: 171-185.

Krafft, P. M., et al. (2018). An experimental study of cryptocurrency market dynamics. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, ACM.

Kumar, S. (2017). "Revisiting the price-volume relationship: a cross-currency evidence." International Journal of Managerial Finance **13**(1): 91-104.

Kurov, A. (2010). "Investor sentiment and the stock market's reaction to monetary policy." Journal of Banking & Finance **34**(1): 139-149.

Lauricella, T. (2011). "Pivot Point: Investors Lose Faith in Stocks." Wall Street Journal.

Luu, J. C. and M. Martens (2003). "Testing the mixture-of-distributions hypothesis using "realized" volatility." Journal of Futures Markets: Futures, Options, and Other Derivative Products **23**(7): 661-679.

Malliaris, A. G. and J. L. Urrutia (1998). "Volume and price relationships: hypotheses and testing for agricultural futures." Journal of Futures Markets **18**(1): 53-72.

Morgan, I. G. (1976). "Stock prices and heteroscedasticity." The Journal of business **49**(4): 496-508.

Mougoué, M. and R. Aggarwal (2011). "Trading volume and exchange rate volatility: Evidence for the sequential arrival of information hypothesis." Journal of Banking & Finance **35**(10): 2690-2703.

Najand, M. and K. Yung (1991). "A GARCH examination of the relationship between volume and price variability in futures markets." Journal of Futures Markets **11**(5): 613-621.

Nakamoto, S. (2008). "Bitcoin: A peer-to-peer electronic cash system."

Nakamoto, S. (2009). Bitcoin: A Peer-to-Peer Electronic Cash System. Whitepaper (2008).

Okamoto, T. and K. Ohta (1991). Universal electronic cash. Annual international cryptology conference, Springer.

Peterson, R. (2013). "Thomson Reuters MarketPsych Indices (TRMI) White Paper." Inside the Mind of the Market.

Poyser, O. (2018). "Herding behavior in cryptocurrency markets." arXiv preprint arXiv:1806.11348.

Rashid, A. (2007). "Stock prices and trading volume: An assessment for linear and nonlinear Granger causality." Journal of Asian Economics **18**(4): 595-612.

Rogalski, R. J. (1978). "The dependence of prices and volume." The Review of Economics and Statistics: 268-274.

Rubbaniy, G., et al. (2014). "Do fear indices help predict stock returns?" Quantitative Finance **14**(5): 831-847.

Rutledge, D. J. (1979). Trading volume and price variability: New evidence on the price effects of speculation.

Simon, D. P. and R. A. Wiggins III (2001). "S&P futures returns and contrary sentiment indicators." Journal of Futures Markets: Futures, Options, and Other Derivative Products **21**(5): 447-462.

Smales, L. A. (2015). "Asymmetric volatility response to news sentiment in gold futures." Journal of International Financial Markets, Institutions and Money **34**: 161-172.

Stalling, W. (2003). "Network Security Essentials: Applications and Standards Prentice Hall." New Jersey.

Sun, L., et al. (2016). "Stock return predictability and investor sentiment: A high-frequency perspective." Journal of Banking & Finance **73**: 147-164.

Tauchen, G. E. and M. Pitts (1983). "The price variability-volume relationship on speculative markets." Econometrica: Journal of the Econometric Society: 485-505.

Tetlock, P. C. (2007). "Giving content to investor sentiment: The role of media in the stock market." the Journal of Finance **62**(3): 1139-1168.

Verma, R., et al. (2008). "The impact of rational and irrational sentiments of individual and institutional investors on DJIA and S&P500 index returns." Applied Financial Economics **18**(16): 1303-1317.

Verma, R. and G. Soydemir (2009). "The impact of individual and institutional investor sentiment on the market price of risk." The Quarterly Review of Economics and Finance **49**(3): 1129-1145.

Verma, R. and P. Verma (2007). "Noise trading and stock market volatility." Journal of Multinational Financial Management **17**(3): 231-243.

Wang, Y.-H., et al. (2006). "The relationships between sentiment, returns and volatility." International Journal of Forecasting **22**(1): 109-123.

Westerfield, R. (1977). "The distribution of common stock price changes: An application of transactions time and subordinated stochastic models." Journal of Financial and quantitative Analysis **12**(5): 743-765.

Yelowitz, A. and M. Wilson (2015). "Characteristics of Bitcoin users: an analysis of Google search data." Applied Economics Letters **22**(13): 1030-1036.

Yermack, D. (2015). Is Bitcoin a real currency? An economic appraisal. Handbook of digital currency, Elsevier**:** 31-43.

Ying, C. C. (1966). "Stock market prices and volumes of sales." Econometrica (pre-1986) **34**(3): 676.

Table 1. Summary statistics - Spot

Panel A: Transactions Frequency

	Price	Volume	Return	|Return|	Uptick	Downtick	Sametick
Count	50423880	50423880	50423880	50423880	50423880	50423880	50423880
Mean	6221.80	0.28	0.00	0.00	0.25	0.25	0.50
Std	5092.88	1.10	0.00	0.00	0.43	0.44	0.50
Min	109.87	0.00	-0.45	0.00	0.00	0.00	0.00
Median	6403.73	0.03	0.00	0.00	0.00	0.00	0.00
Max	19891.99	89.98	0.77	0.77	1.00	1.00	1.00

Panel B: 5 minutes Frequency

	Price	Volume	Return	|Return|	Uptick	Downtick
Count	381308	381308	381308	381308	381308	381308
Mean	2967.50	37.64	0.00	0.00	0.50	0.44
Std	3852.84	69.21	0.00	0.00	0.50	0.50
Min	111.89	0.00	-0.53	0.00	0.00	0.00
Median	742.33	19.17	0.00	0.00	0.00	0.00
Max	19891.98	2682.82	0.49	0.53	1.00	1.00

Panel C: Hourly Frequency

	Price	Volume	Return	|Return|	Uptick	Downtick
Count	31989	31989	31989	31989	31989	31989
Mean	2953.17	448.63	0.00	0.00	0.53	0.47
Std	3847.43	649.78	0.01	0.01	0.50	0.50
Min	113.06	0.02	-0.44	0.00	0.00	0.00
Median	737.96	288.25	0.00	0.00	1.00	0.00
Max	19847.11	31505.46	0.28	0.44	1.00	1.00

This table shows a summary of descriptive statistics on our sample data on the spot market. Price is the close price at the end of each time period (transaction, 5 minutes, or one hour). Volume is the total volume traded during the time period. Return is return during the period calculated as the log of close price of the period minus the log of open price of the period (for transactions frequency calculated as log of price minus log of price for previous transaction). |Return| is magnitude of return, calculated as the absolute value of Return. Uptick is a dummy variable equal to 1 if Return > 0 and 0 otherwise. Downtick is a dummy variable equal to 1 if Return < 0 and 0 otherwise.

Table 2. Summary statistics – Futures

Panel A: Transactions Frequency

	Price	Volume	Return	|Return|	Uptick	Downtick	Sametick
Count	262542	262542	262542	262542	262542	262542	262542
Mean	8694.75	1.17	-0.00	0.00	0.28	0.28	0.44
Std	2013.00	0.53	0.00	0.00	0.45	0.45	0.50
Min	5755.00	1.00	-0.07	0.00	0.00	0.00	0.00
Median	8435.00	1.00	0.00	0.00	0.00	0.00	0.00
Max	19770.00	23.00	0.07	0.07	1.00	1.00	1.00

Panel B: 5 minutes Frequency

	Price	Volume	Return	|Return|	Uptick	Downtick
Count	30055	30055	30055	30055	30055	30055
Mean	9255.72	10.18	0.00	0.00	0.34	0.34
Std	2558.45	17.02	0.00	0.00	0.47	0.47
Min	5775.00	1.00	-0.06	0.00	0.00	0.00
Median	8625.00	5.00	0.00	0.00	0.00	0.00
Max	19770.00	412.00	0.06	0.06	1.00	1.00

Panel C: Hourly Frequency

	Price	Volume	Return	|Return|	Uptick	Downtick
Count	3127	3127	3127	3127	3127	3127
Mean	9484.19	97.81	-0.00	0.01	0.47	0.48
Std	2726.74	117.54	0.01	0.01	0.50	0.50
Min	5830.00	1.00	-0.11	0.00	0.00	0.00
Median	8760.00	62.00	0.00	0.00	0.00	0.00
Max	19715.00	1470.00	0.11	0.11	1.00	1.00

This table shows a summary of descriptive statistics on our sample data on the futures market. Price is the close price at the end of each time period (transaction, 5 minutes, or one hour). Volume is the total volume traded during the time period. Return is return during the period calculated as the log of close price of the period minus the log of open price of the period (for transactions frequency calculated as log of price minus log of price for previous transaction). |Return| is magnitude of return, calculated as the absolute value of Return. Uptick is a dummy variable equal to 1 if Return > 0 and 0 otherwise. Downtick is a dummy variable equal to 1 if Return < 0 and 0 otherwise.

Table 3. Vector Auto Regression Results - Spot

Panel A: Transactions Frequency					
Dependent Variable	Constant	Volume(-1)	Return(-1)	Volume(-2)	Return(-2)
Volume	0.23***	0.10***	-0.83***	0.08***	-0.71***
	(0)	(0)	(0.13)	(0)	(0.13)
Return	0.0000***	0.0000***	-0.5200***	0.0000***	-0.2319***
	(0)	(0)	(0.0001)	(0)	(0.0001)
Panel B: 5 Minutes Frequency					
DV	Constant	Volume(-1)	Return(-1)	Volume(-2)	Return(-2)
Volume	8.47***	0.52***	-22.28***	0.26***	1.41
	(0.09)	(0)	(4.65)	(0)	(4.65)
Return	0	0.0000**	0.0133***	0	0.0085***
	(0)	(0)	(0.0016)	(0)	(0.0016)
Panel C: Hourly Frequency					
DV	Constant	Volume(-1)	Return(-1)	Volume(-2)	Return(-2)
Volume	93.25***	0.70***	-94.82**	0.05***	18.45
	(2.93)	(0.01)	(41.43)	(0.01)	(41.43)
Return	0.0004	0	-0.0007	0	-0.0015
	(0.0004)	(0)	(0.0056)	(0)	(0.0056)

This table show the results for vector auto regression in the spot market. The effect of up to two lags of Volume and Return on each other is being tested. Each variable is regressed on an intercept, two lags of itself, and two lags of the other variable. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 4. Vector Auto Regression Results - Futures

Panel A: Transactions Frequency					
Dependent Variable	Constant	Volume(-1)	Return(-1)	Volume(-2)	Return(-2)
Volume	0.87***	0.09***	-1.46**	0.05***	-0.58
	(0)	(0)	(0.66)	(0)	(0.66)
Return	0	0	0	0	-0.01***
	(0)	(0)	(0)	(0)	(0)
Panel B: 5 Minutes Frequency					
DV	Constant	Volume(-1)	Return(-1)	Volume(-2)	Return(-2)
Volume	3.38***	0.38***	-130.74	0.10***	-94.33***
	(0.11)	(0.01)	(25.33)	(0.01)	(25.35)
Return	-0.00**	0	-0.01**	0	-0.01
	(0)	(0)	(0.01)	(0)	(0.01)
Panel C: Hourly Frequency					
DV	Constant	Volume(-1)	Return(-1)	Volume(-2)	Return(-2)
Volume	31.66***	0.4***	-394.30*	0.062***	-75.57
	(2.62)	(0.02)	(135.09)	(0.02)	(135.16)
Return	0	0	-0.01	0	-0.08***
	(0)	(0)	(0.02)	(0)	(0.02)

This table show the results for vector auto regression in the futures market. The effect of up to two lags of Volume and Return on each other is being tested. Each variable is regressed on an intercept, two lags of itself, and two lags of the other variable. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 5. Granger Causality Test Results

	Transactions	5 minutes	Hourly
Panel A: Spot			
Lag 1	9.30***	40.04***	4.55**
	(0.00)	(0.00)	(0.03)
Lag 2	29.42***	113.03***	2.70*
	(0.00)	(0.00)	(0.07)
Panel B: Futures			
	Transactions	5 minutes	Hourly
Lag 1	4.578**	19.111***	6.820***
	(0.032)	(0.000)	(0.009)
Lag 2	2.673*	17.090***	4.016**
	(0.069)	(0.000)	(0.018)

This table shows the results of Granger causality test. We are examining to see whether lags of Return do have statistical significance in explaining variation in Volume. The first two lags of Return have been included in the regression. Reported numbers are F-statistics and p-values are reported in the parentheses. Null hypothesis is that the first two lags of Return have no statistically significant influence on Volume. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 6. Specifications I and II Regression Results

Panel A: HC1 Robust Errors - Spot						
	Transactions		5 Minutes		Hourly	
Intercept	0.29***	0.28***	37.11***	37.57***	287.51***	448.57***
Return		-0.61***		-380.95***		-2073.39***
|Return|	-0.15		4273.31***		32870***	
Panel B: GARCH (1, 1) - Spot						
Intercept	0.1305***	0.1305***	11.417***	19.0303***	274.4634***	263.5037***
Return		-0.0911		-700.931***		-24488***
|Return|	-0.3378***		11228***		20637***	
ARCH0	0.0083***	0.0083***	47.831***	86.0496***	330078***	9516***
ARCH1	0.0981***	0.0981***	0.1951***	0.2369***	0.3365***	0.1837***
GARCH1	0.9432***	0.9432***	0.8328***	0.7947***	0	0.7855***
Panel C: HC1 Robust Errors - Futures						
	Transactions		5 Minutes		Hourly	
Intercept	1.17***	1.17***	6.44***	10.18***	79.10***	97.7366***
Return		-1.23		-22.38		-256.12
|Return|	-4.05***		2194.79***		2342.35***	
Panel D: GARCH (1, 1) - Futures						
Intercept	1.1239***	1.12***	3.5819***	6.0244***	35.2214***	64.4945***
Return		0.05		5.7446		-53.0782
|Return|	-6.5837***		1477***		2058***	
ARCH0	0.009***	0.0009***	32.9594***	37.1902***	1197***	5301***
ARCH1	0.0179***	0.0178***	0.4884***	0.5108***	0.536***	0.7136***
GARCH1	0.9804***	0.9805***	0.5539***	0.5607***	0.5426***	0.0851***

This table shows the regression results for specifications I and II (I: $Volume_t = \alpha + \beta Return_t + \varepsilon_t$; II: $Volume_t = \alpha + \beta |Return|_t + \varepsilon_t$). For each frequency, results of specification I are reported on the right and the results of specification II are reported on the left hand side. In panels B and D, GARCH (1, 1) effects have been added to the regression. This implies running an additional regression on the residuals (ε_t) from either specification. The GARCH (1, 1) regression is $h_t = \theta + \mu h_{t-1} + \vartheta \varepsilon_{t-1}^2$ where h_t is given by $\varepsilon_t = \sqrt{h_t}.\omega_t$ and ω_t is i.i.d with zero mean and unit variance. ARCH0 shows the value of θ, ARCH1 that of μ, and GARCH1 that of ϑ. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 7. Bruesch-Pagan Heteroskedasticity Test Results

	Transactions	5 Minutes	Hourly
Panel A: Spot			
Lagrange Multiplier	0.07	44844.60	416.38
P-value	0.79	0.00	0.00
F-statistic	0.07	50808.24	421.84
P-value	0.79	0.00	0.00
Panel B: Futures			
Lagrange Multiplier	1.96	915.12	60.57
P-value	0.16	0.00	0.00
F-statistic	1.96	943.79	61.73
P-value	0.16	0.00	0.00

This table shows the results from Bruesch-Pagan heteroscedasticity test. This is to test the correlation between error terms from specification II (II: $Volume_t = \alpha + \beta|Return|_t + \varepsilon_t$) with the independent variable |Return|. Both the Lagrange multiplier statistic and the F-statistic along with their corresponding p-values are reported.

Table 8. Specification III Regression Results

	Transactions		5 Minutes		Hourly	
Panel A: Ordinary Least Squares - Spot						
Intercept	0.28***	0.2454***	29.76***	8.41***	271.95***	57.41***
Up	0.00***	-0.0039***				
Down	0.02***	0.0139***	4.90***	3.55***	31.94	13.62
Up Return	-0.43***	-3.7857***	3997.14***	2802.88***	34350***	18820***
Down Return	-0.56***	-3.9587***	4510.25***	3531.66***	31590***	21830***
Volume(-1)		0.1085***		0.69***		0.73***
Up(-1)		0.0049***				
Down(-1)		0.0181***		-0.92**		-31.95***
Up Return(-1)		2.4909***		-1501.95***		-7632.97***
Down Return(-1)		4.8368***		-1476.29***		-2764.73***
Panel B: GARCH(1, 1)						
	Transactions		5 Minutes		Hourly	
Intercept	0.1284***	0.1277***	11.8994***	11.4072***	223.8684***	188.3981***
Up	-0.0137***	-0.0131***				
Down	0.0242***	0.0219***	0.2232***	0.1302***	80.8003***	68.1301***
Up Return	-0.197***	-1.0047***	10442***	9397***	31272***	34834***
Down Return	-0.6782***	-2.0937***	13585***	12266***	14960***	23883***
Up(-1)		-0.00826***				
Down(-1)		0.0138***		-2.8316***		-166.0027***
Up Return(-1)		0.3555***		928.1957***		7178***
Down Return(-1)		1.5157***		5200***		53534***
ARCH0	0.0083***	0.0084***	19.4648***	60.0285***	326663***	259872***
ARCH1	0.0987***	0.0993***	0.1416***	0.2382***	0.422***	0.3808***
GARCH1	0.943***	0.9427***	0.8886***	0.7979***	0	0.001021***
Panel C: Ordinary Least Squares - Futures						
	Transactions		5 Minutes		Hourly	
Intercept	1.14***	1.03***	5.63***	2.18***	72.59***	18.12***
Up	0.07***	0.07***				
Down	0.07***	0.07***	2.94***	1.65***	14.28**	11.65**
Up Return	-16.80***	-11.54***	2289.69***	2008.18***	2482.84***	3208.72***
Down Return	-14.90***	-9.20***	1888.60***	1858.87***	2132.27***	3106.38***
Volume(-1)		0.11***		0.47***		0.6009***
Up(-1)		-0.03***				
Down(-1)		-0.02***		0.34		4.84
Up Return(-1)		-12.00***		-514.23***		-1708.91***
Down Return(-1)		-10.04***		-420.68***		-1348.41***

Panel D: GARCH(1, 1) - Futures						
	Transactions		5 Minutes		Hourly	
Intercept	1.1077***	1.125***	3.4095***	2.9863***	31.7174***	40.4731***
Up	0.0305***	0.0353***				
Down	0.0341***	0.0371***	0.6959***	0.6956***	11.0608***	11.1482***
Up Return	-10.0987***	-8.6528***	1533***	1418***	2306***	2472***
Down Return	-11.4576***	-8.9295***	1378***	1339***	1544***	1851***
Up(-1)		-0.0213***				
Down(-1)		-0.0202***		0.7147***		3.1405
Up Return(-1)		-9.3558***		216.2469***		-550.9031**
Down Return(-1)		-9.117***		449.9601***		-547.9787*
ARCH0	0.0009***	0.0009***	33.2023***	32.6085***	1229***	4173***
ARCH1	0.0179***	0.0179***	0.4869***	0.4853***	0.5265***	0.7133***
GARCH1	0.9804***	0.9806***	0.5534***	0.559***	0.5422***	0.1373***

This table shows regression results for specification III ($Volume_t = \alpha + \beta . down_t + \gamma . up_t \times Return_t + \delta . down_t \times |Return|_t + \varepsilon_t$) with and without inclusion of the first lag of all independent variables. Up is the Uptick dummy, Down is the Downtick dummy, Up Return is the interaction of Uptick with |Return|, Down Return is the interaction of Downtick with |Return|. For each frequency, results of specification III regression are reported on the left hand side and results of the same regression but with first lags of all variables included in the equation are reported on the right hand side. In panels B and D GARCH (1, 1) effect has been controlled for. This implies running an additional regression on the residuals (ε_t) from either specification. The GARCH (1, 1) regression is $h_t = \theta + \mu h_{t-1} + \vartheta \varepsilon_{t-1}^2$ where h_t is given by $\varepsilon_t = \sqrt{h_t} . \omega_t$ and ω_t is i.i.d with zero mean and unit variance. ARCH0 shows the value of θ, ARCH1 that of μ, and GARCH1 that of ϑ. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 9. Summary Statistics

Panel A: Spot market								
	Return	Volume	Volatility	Sentiment	Optimism	Joy	Fear	Gloom
Count	7319	7320	7320	7320	7320	7320	7320	7320
Mean	0	465	14.5	-0.01	0.02	0.01	0.01	0.02
Std	0	586	17	0.05	0.02	0	0	0.01
Min	-0.01	0	0	-0.31	-0.17	0	-0.01	-0
Median	0	290	9.37	-0	0.02	0.01	0.01	0.02
Max	0.02	10097	303	0.32	0.11	0.09	0.09	0.1
Panel B: Futures Market								
	Return	Volume	Volatility	Sentiment	Optimism	Joy	Fear	Gloom
Count	2900	2901	2901	2901	2901	2901	2901	2901
Mean	0	87	14.8	-0	0.02	0.01	0.01	0.02
Std	0	126	21.1	0.04	0.02	0	0	0.01
Min	-0.07	0	0	-0.31	-0.17	0	-0.01	-0
Median	0	52	9.79	-0	0.02	0.01	0.01	0.02
Max	0.07	1470	299	0.32	0.1	0.09	0.09	0.1

This table shows a summary of the descriptive statistics for our variables of interest. Return is calculated as log of close price minus the log of open price. Volume is volume of trade over each time period. Volatility is the standard deviation of transaction frequency Return over each time period. Sentiment is the overall sentiment measure that reflects overall positive attitudes net of negative attitudes. Optimism is the weight of optimism mentioned, net of pessimism. Joy shows mentions of the word “joy” and its equivalents. Fear reflects mentions of the word “fear” and its equivalents. Gloom shows mentions of the word “gloom” and its equivalents.

Table 10. Correlation Matrix

Panel A: Spot market

	Return	Volume	Volatility	Sentiment	Optimism	Joy	Fear	Gloom
Return	1.00	0.01	0.00	0.02	0.00	0.02	0.00	-0.03
Volume		1.00	0.71	-0.25	-0.24	0.16	0.04	0.18
Volatility			1.00	-0.12	-0.13	0.09	0.10	0.09
Sentiment				1.00	0.69	0.08	-0.12	-0.40
Optimism					1.00	-0.06	-0.08	-0.36
Joy						1.00	-0.05	0.05
Fear							1.00	0.20
Gloom								1.00

Panel B: Futures Market

	Return	Volume	Volatility	Sentiment	Optimism	Joy	Fear	Gloom
Return	1.00	-0.01	-0.02	0.04	0.05	0.01	0.04	-0.01
Volume		1.00	0.57	-0.06	-0.05	0.06	-0.03	0.02
Volatility			1.00	-0.11	-0.12	0.06	-0.01	0.07
Sentiment				1.00	0.64	0.18	-0.20	-0.37
Optimism					1.00	0.00	-0.13	-0.33
Joy						1.00	-0.05	-0.02
Fear							1.00	0.25
Gloom								1.00

This table shows the correlation matrix for all of our variables of interest. Return is calculated as log of close price minus the log of open price. Volume is volume of trade over each time period. Volatility is the standard deviation of transaction frequency Return over each time period. Sentiment is the overall sentiment measure that reflects overall positive attitudes net of negative attitudes. Optimism is the weight of optimism mentioned, net of pessimism. Joy shows mentions of the word "joy" and its equivalents. Fear reflects mentions of the word "fear" and its equivalents. Gloom shows mentions of the word "gloom" and its equivalents. All of the correlation figures among our sentiment measures are statistically significant at the 1% level.

Table 11. Contemporaneous Regression Results

	Model 1	Model 2	Model 3
Panel A: Spot Market – Simple OLS			
Intercept	7.49E-05	8.45E-06	5.97E-05
Sentiment	0.0007*	0.0006*	
Optimism	-0.0015*		-0.0004
Joy	0.0054*		0.0064**
Fear	0.0016		-0.0043***
Gloom	-0.0037**		0.0014
Panel B: Spot Market – GARCH (1, 1)			
	Model 1	Model 2	Model 3
Intercept	0.0001	0.0000	0.0001
Sentiment	-0.0002	-0.0001	
Optimism	-0.0001		-0.0004
Joy	0.0020		0.0017
Fear	-0.0043*		-0.0042*
Gloom	-0.0012		-0.0010
ARCH0	1.05E-08***	1.10E-08***	1.05E-08***
ARCH1	0.036***	0.0369***	0.0359***
GARCH1	0.9574***	0.9564***	0.9576***
Panel C: Futures Market – Simple OLS			
	Model 1	Model 2	Model 3
Intercept	-0.0004	0.0001**	-0.0004*
Sentiment	0.0015	0.0027**	
Optimism	0.0066*		0.0086***
Joy	0.0085		0.011
Fear	0.0321**		0.0311*
Gloom	0.0005		-0.0007
Panel D: Futures Market – GARCH (1, 1)			
	Model 1	Model 2	Model 3
Intercept	-0.0003	0.0001	-0.0004
Sentiment	0.0014	0.0026	
Optimism	0.0066		0.0085
Joy	0.0085		0.0110
Fear	0.0324*		0.0315**
Gloom	-0.0001		-0.0012
ARCH0	9.4586E-06***	0.0000	9.4609E-06***
ARCH1	0.0000	0.0000	0.0000
GARCH1	0.0000	0.0001	0.0000

This table shows the results of regressing Returns on our sentiment measures of interest. Specifications for the three models are Model 1: $Return_t = \alpha + \beta \cdot Sentiment_t + \gamma \cdot Optimism_t + \delta \cdot Joy_t + \theta \cdot Fear_t + \mu \cdot Gloom_t + \varepsilon_t$; Model 2: $Return_t = \alpha + \beta \cdot Sentiment_t + \varepsilon_t$; Model 3: $Return_t = \alpha + \gamma \cdot Optimism_t + \delta \cdot Joy_t + \theta \cdot Fear_t + \mu \cdot Gloom_t + \varepsilon_t$. Return is calculated as log of close price minus the log of open price. Sentiment is the overall sentiment measure that reflects overall positive attitudes net of negative attitudes. Optimism is the weight of optimism mentioned, net of pessimism. Joy shows mentions of the word "joy" and its equivalents. Fear reflects mentions of the word "fear" and its equivalents. Gloom shows mentions of the word "gloom" and its equivalents. Panels B and D incorporate GARCH (1, 1) correction. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 12. Causal Regression Results

	Model 1	Model 2	Model 3
Panel A: Spot Market – Simple OLS			
Intercept	1.96E-05	9.24E-06	-4.58E-08
Return	-0.0244	-0.0236	-0.0238
Sentiment	0.0009**	0.0006*	
Optimism	-0.0012		0.0002
Joy	0.0006		0.0019
Fear	0.0097		0.0095
Gloom	-0.0024		-0.0032**
Panel B: Spot Market – GARCH (1, 1)			
	Model 1	Model 2	Model 3
Intercept	9.83E-06	1.48E-05	1.02E-05
Return	-0.0174	-0.0249	-0.0211
Sentiment	-0.0001	-0.0001	
Optimism	-0.0002		-0.0003
Joy	0.0010		0.0009
Fear	0.009987***		0.009992***
Gloom	-0.002846**		-0.002783**
ARCH0	1.08E-08***	1.05E-08***	1.05E-08***
ARCH1	0.0368***	0.0359***	0.0361***
GARCH1	0.9564***	9.58E-01***	9.57E-01***
Panel C: Futures Market – Simple OLS			
	Model 1	Model 2	Model 3
Intercept	0.0002	0.0001**	0.0002
Return	-0.0033	-0.0031	-0.003
sentiment	0.0019	0.0027**	
optimism	0.0013		0.0041
joy	0.0008		0.0042
gloom	0.0021		0.0008
fear	-0.0064		-0.0079
Panel D: Futures Market: GARCH (1, 1)			
	Model 1	Model 2	Model 3
Intercept	0.000202	0.000127**	0.000163
sentiment	-0.0037	-0.0032	-0.0034
optimism	0.002032	0.002746*	0.004206
joy	0.001362		
fear	0.000662		0.004192
gloom	0.001422		-0.000029
ARCH0	-0.005133		-0.00672
ARCH1	9.49E-06	9.49E-06	9.49E-06
GARCH1	0	0	0

This table shows the results of regressing Returns on first lags of our sentiment measures of interest. Specifications for the three models are Model 1: $Return_t = \alpha + \beta \cdot Sentiment_{t-1} + \gamma \cdot Optimism_{t-1} + \delta \cdot Joy_{t-1} + \theta \cdot Fear_{t-1} + \mu \cdot Gloom_{t-1} + \varepsilon_t$; Model 2: $Return_t = \alpha + \beta \cdot Sentiment_{t-1} + \varepsilon_t$; Model 3: $Return_t = \alpha + \gamma \cdot Optimism_{t-1} + \delta \cdot Joy_{t-1} + \theta \cdot Fear_{t-1} + \mu \cdot Gloom_{t-1} + \varepsilon_t$. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 13. Vector Autoregression Results – Return

Panel A: Spot Market			
	Model 1	Model 2	Model 3
Const.	-0.000001	0.00001	-0.000014
L1.Return	-0.024305**	-0.023807**	-0.024144**
L1.Sentiment	0.000978**	0.000587*	
L1.Optimism	-0.001552*		-0.00035
L1.Joy	-0.000087		0.001361
L1.Gloom	-0.00185		-0.002467
L1.Fear	0.009953***		0.009358***
L2.Return	-0.006902	-0.006142	-0.006561
L2.Sentiment	-0.000512	0.000116	
L2.Optimism	0.001566*		0.001137
L2.Joy	0.003476		0.002814
L2.Gloom	-0.001646		-0.001464
L2.Fear	-0.000171		0.000328
Panel B: Futures Market			
	Model 1	Model 2	Model 3
Const.	-0.000144	0.000128**	-0.000182
L1.Return	-0.003089	-0.003542	-0.002892
L1.Sentiment	0.001596	0.001948	
L1.Optimism	-0.000187**		0.001931
L1.Joy	-0.000432**		0.002082
L1.Gloom	-0.002243		-0.003244
L1.Fear	-0.009103		-0.010203
L2.Return	0.00033**	0.002543	0.000317
L2.Sentiment	0.000096*	0.001623	
L2.Optimism	0.008063		0.008535**
L2.Joy	0.001887		0.003103
L2.Gloom	0.009609		0.009465
L2.Fear	0.010743		0.010269

This table show the results for vector auto regression for Return. The effect of up to two lags of our five sentiment measures on Return is being tested. Return is regressed on an intercept, two lags of itself, and two lags of the other five sentiment variables. Return is calculated as log of close price minus the log of open price. Sentiment is the overall sentiment measure that reflects overall positive attitudes net of negative attitudes. Optimism is the weight of optimism mentioned, net of pessimism. Joy shows mentions of the word "joy" and its equivalents. Fear reflects mentions of the word "fear" and its equivalents. Gloom shows mentions of the word "gloom" and its equivalents. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 14. Vector Autoregression Results - Volume

	Model 1	Model 2	Model 3
Panel A: Spot Market			
Const.	125.532313***	185.372363***	143.564318***
L1.Volume	0.4639***	0.46844***	0.466667***
L1.Sentiment	-394.026367**	-677.374863***	
L1.Optimism	-652.349773		-1279.642077***
L1.Joy	1493.34816		826.797747
L1.Gloom	3922.609453***		3920.305047***
L1.Fear	499.587468		865.312443
L2.Volume	0.109181***	0.115325***	0.112827***
L2.Sentiment	-359.687506**	-607.545096***	
L2.Optimism	-588.362964		-1170.57244***
L2.Joy	1208.555041		541.030543
L2.Gloom	-1870.674467		-1845.611161
L2.Fear	1753.833856**		2076.567661***
Panel B: Futures Market			
	Model 1	Model 2	Model 3
Const.	43.585183***	31.018731***	43.663359***
L1.Volume	0.525649***	0.524875***	0.52566***
L1.Sentiment	-0.540836	1.023273	
L1.Optimism	-55.0924		-56.521962
L1.Joy	-109.975107		-112.057411
L1.Gloom	42.951454		42.822478
L1.Fear	-437.90717*		-436.428851*
L2.Volume	0.118664***	0.118522***	0.118705***
L2.Sentiment	-3.301364	-23.276702	
L2.Optimism	-29.277264		-33.741514
L2.Joy	-556.066279		-561.853361
L2.Gloom	-391.497563		-389.133966
L2.Fear	324.426571		326.613945

This table show the results for vector auto regression for Volume. The effect of up to two lags of our five sentiment measures on Volume is being tested. Volume is regressed on an intercept, two lags of itself, and two lags of the other five sentiment variables. Volume is calculated as the trading volume in each time period. Sentiment is the overall sentiment measure that reflects overall positive attitudes net of negative attitudes. Optimism is the weight of optimism mentioned, net of pessimism. Joy shows mentions of the word “joy” and its equivalents. Fear reflects mentions of the word “fear” and its equivalents. Gloom shows mentions of the word “gloom” and its equivalents. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 15. Vector Autoregression Results - Volatility

Panel A: Spot Market			
	Model 1	Model 2	Model 3
Const.	6.103666***	7.304511***	6.220301***
L1.Volume	0.268904***	0.275819***	0.269051***
L1.Sentiment	-2.253206	-11.348669**	
L1.Optimism	-20.381605		-24.193634**
L1.Joy	226.935441***		226.72123***
L1.Gloom	-25.789509		-29.211567
L1.Fear	0.931893		3.215295
L2.Volume	0.203944***	0.210658***	0.204211***
L2.Sentiment	-2.418278	-9.940322**	
L2.Optimism	-26.52024***		-30.452988***
L2.Joy	74.05034		74.331492
L2.Gloom	28.741822		24.725986
L2.Fear	8.839939		11.075759
Panel B: Futures Market			
	Model 1	Model 2	Model 3
Const.	9.263753***	7.191269***	9.607736***
L1.Volume	0.289658***	0.292165***	0.29053***
L1.Sentiment	-7.673856	-9.923711	
L1.Optimism	-22.696997		-34.64406
L1.Joy	-7.277582		-3.459214
L1.Gloom	84.823215		68.866444
L1.Fear	-40.38488		-32.533091
L2.Volume	0.214541***	0.217784***	0.215834***
L2.Sentiment	-9.215756	-21.295162**	
L2.Optimism	-46.5545*		-60.1736***
L2.Joy	-37.611608		-30.599646
L2.Gloom	-29.180643		-49.61898
L2.Fear	12.58447		20.279511

This table show the results for vector auto regression for Volatility. The effect of up to two lags of our five sentiment measures on Volatility is being tested. Volatility is regressed on an intercept, two lags of itself, and two lags of the other five sentiment variables. Volatility is the standard deviation of transaction frequency Return over each time period. Sentiment is the overall sentiment measure that reflects overall positive attitudes net of negative attitudes. Optimism is the weight of optimism mentioned, net of pessimism. Joy shows mentions of the word “joy” and its equivalents. Fear reflects mentions of the word “fear” and its equivalents. Gloom shows mentions of the word “gloom” and its equivalents. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Figure 1. Bitcoin Spot Price

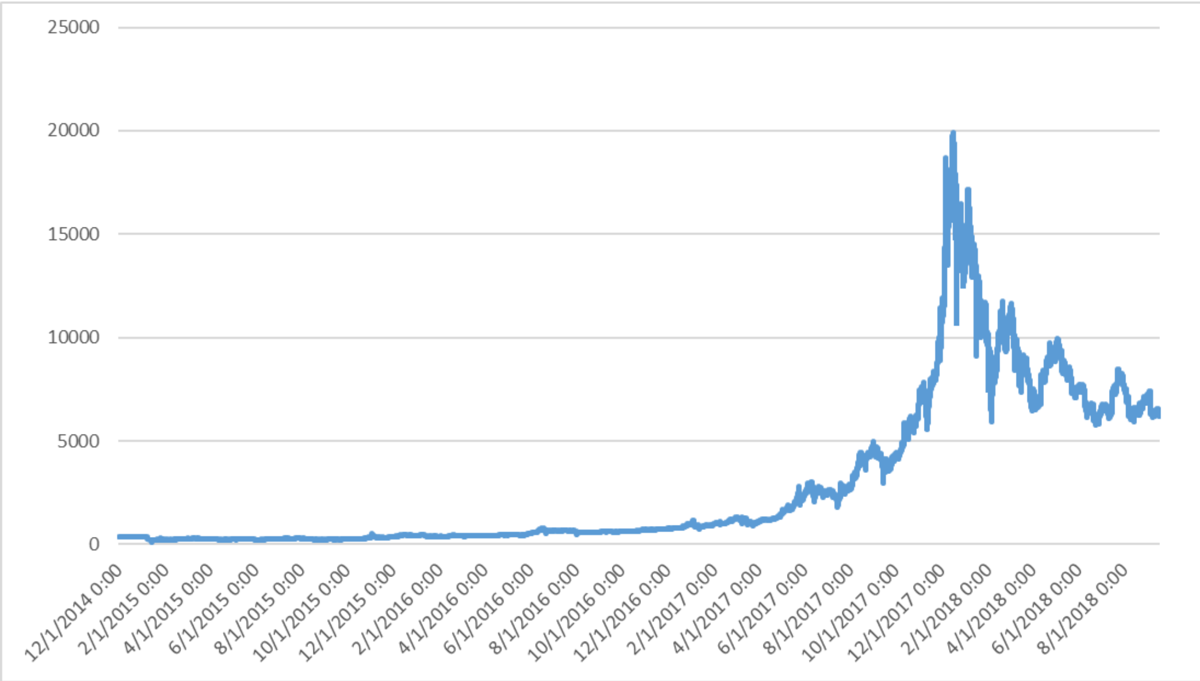


Figure 2: Bitcoin Returns

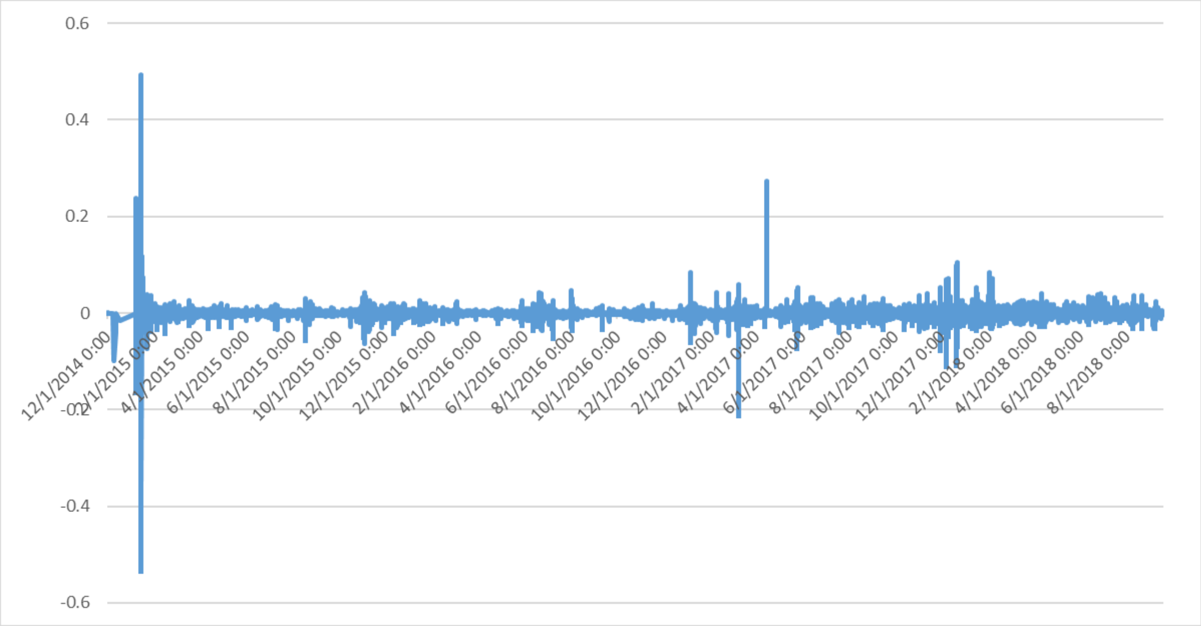


Figure 3: Bitcoin Trading Volume

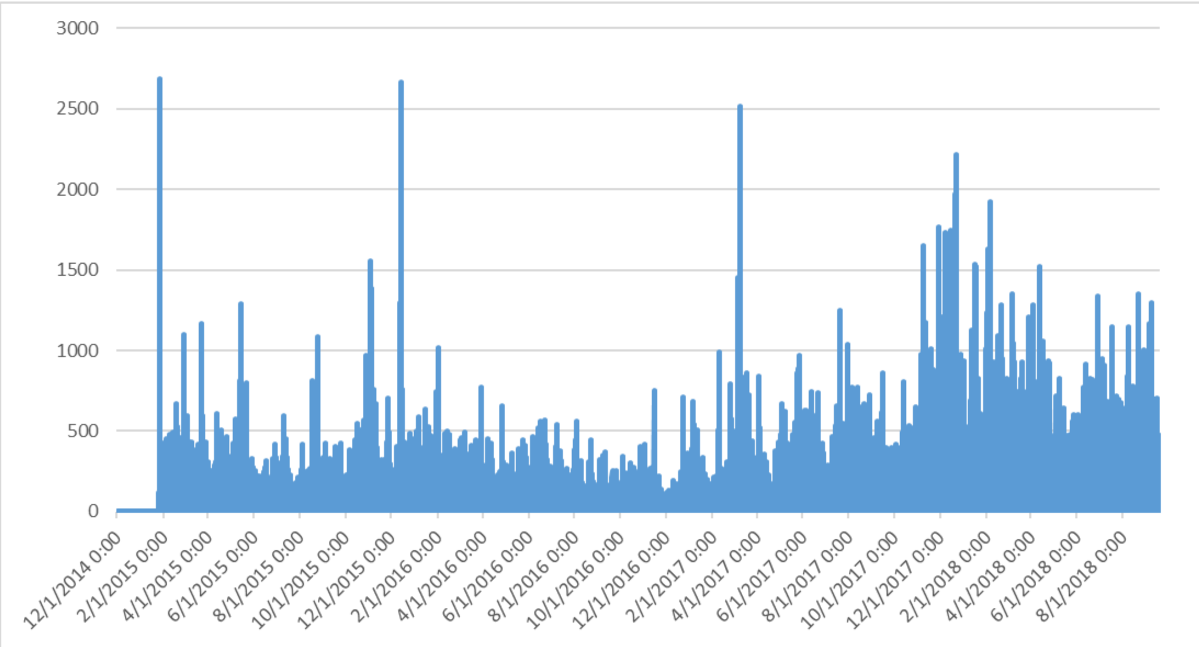
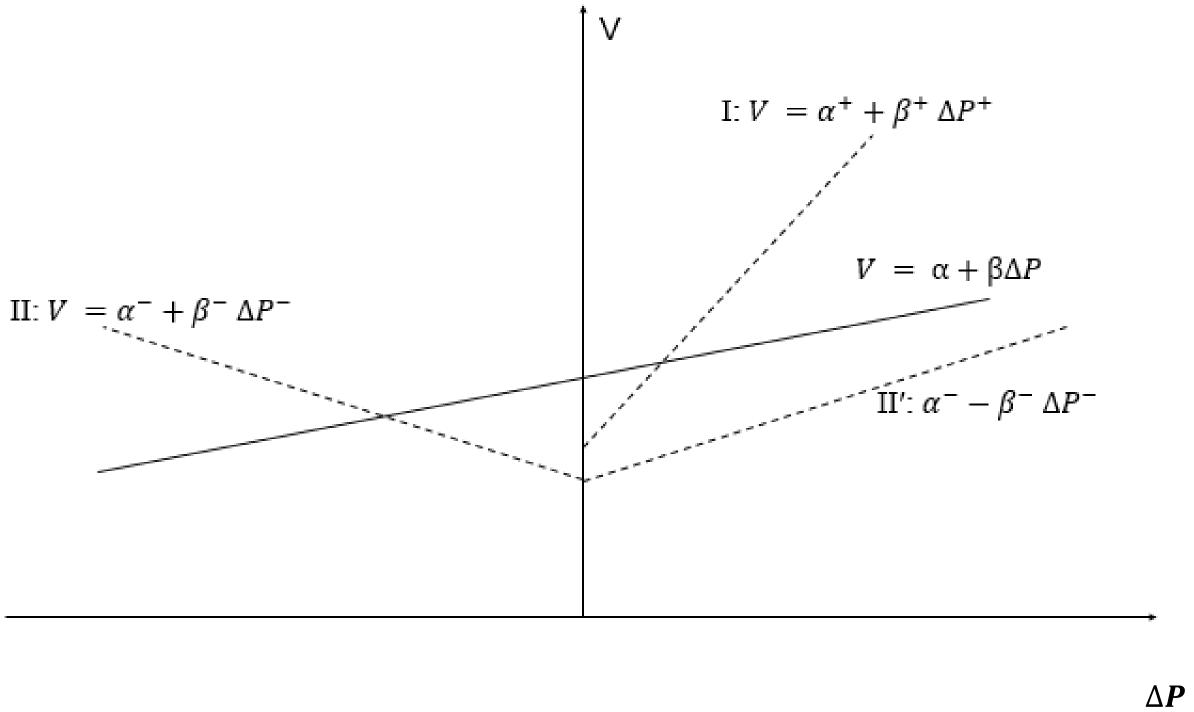


Figure 4. Karpoff's Model



VITA

Mohammad Bayani Khaknejad
Strome College of Business, Old Dominion University, Norfolk, VA 23529
Email: mbaya001@odu.edu
LinkedIn: www.linkedin.com/in/mohammad-bayani

Education

- **Ph.D.**, Finance and International Business, Old Dominion University 2016 - 2019
- **M.A**., Economics, Old Dominion University, 2016 - 2019
- **MBA** with concentration in Finance, Sharif University of Technology, 2012 - 2014
- **B.Sc.** Physics, Sharif University of Technology, 2004 - 2009

Certificates and Awards

- **CFA Level III** Candidate, June 2019
- Member of **Beta Gamma Sigma (ΒΓΣ)** Honor Society, May 2019
- **Bloomberg Market Concepts**, 2016
- **Best Teacher Award**, Allame Helli (NODET), 2008

Work Experience

- **Graduate Research Assistant**, Old Dominion University, 2018-2019
- **Financial Advisor**, Sahifeh Eskan, Tehran, Iran, 2014 – 2016
- **Representative** of Moshaveran Saham Brokerage at 7th International Exhibition of Exchange, Bank and Insurance, Tehran, Iran – June 2014
- **Programmer** of the FOREX Hedging team, Savafarin Brokerage, Tehran, Iran, 2009 – 2010

Teaching Experience

- Instructor**, Introductory Financial Management**, Old Dominion University, Fall 2018
- Instructor, **Introduction to Contemporary Business**, Old Dominion University, Fall 2017 & Spring 2018
- Teacher's Assistant, **Macroeconomics**, Sharif University of Technology, Spring 2013
- Teacher's Assistant, **Special Relativity**, Sharif University of Technology, Spring 2007
- Teacher, **Physics**, Allame Helli High School (NODET), Tehran, Iran, 2005-2009