

**CAN INVESTOR'S SENTIMENT FROM FORUM POSTS PREDICT  
BITCOIN RETURN?**

by  
AYŞE GÜL CANBAZ

Submitted to the Sabancı Graduate Business School  
in partial fulfilment of  
the requirements for the degree of Master of Science in Business Analytics

Sabancı University  
August 2020

CAN INVESTOR'S SENTIMENT FROM FORUM POSTS PREDICT  
BITCOIN RETURN?

Approved by:

Prof. Abdullah Daşçı  
(Thesis co-supervisor)



Assist. Prof. Ali Doruk Günaydın  
(Thesis co-supervisor)



Assoc. Prof. Ayşe Kocabıyıköğlü



Assoc. Prof. Kemal Kılıç



Assist. Prof. Mustafa Hayri Tongarlık



Date of Approval: August 31, 2020



AYŞE GÜL CANBAZ 2020 ©

All Rights Reserved

## ABSTRACT

CAN INVESTOR'S SENTIMENT FROM FORUM POSTS PREDICT BITCOIN  
RETURN?

AYŞE GÜL CANBAZ

BUSINESS ANALYTICS M.Sc. THESIS, AUGUST 2020

Thesis Supervisors: Prof. Abdullah Dağcı, Assist. Prof. Ali Doruk Günaydın

Keywords: Digital currency, Bitcoin, Sentiment Analysis, Vector Autoregressive  
Model

This study aims to investigate if the investors' sentiment expressed on content-specific online forum affects the return of Bitcoin. We use a large dataset consisting of 2.8 million forum posts sourced from "Bitcointalk.org" for a period between Jan 2016 and May 2020. The sentiment is derived daily with the Hu Liu lexical model after a detailed investigation of different lexicons. Using time-series data, we test for the relationship between the investors' sentiment and Bitcoin price along with other financial variables that may inform about the direction of Bitcoin price. Our results show that sentiment derived from online forums do not present an autoregressive relationship with return of Bitcoin.

## ÖZET

### YATIRIMCILARIN FORUM PAYLAŞIMLARINDAKİ DUYGULAR BİTCOİN GETİRİSİNİ TAHMİN EDEBİLİR Mİ?

AYŞE GÜL CANBAZ

İŞ ANALİTİĞİ YÜKSEK LİSANS TEZİ, EYLÜL 2020

Tez Danışmanları: Prof. Dr. Abdullah Daşçı, Dr. Öğretim Üyesi Ali Doruk  
Günaydın

Anahtar Kelimeler: Dijital Para Birimi, Bitcoin, Duygu Analizi, Vektör Otoregresif  
Model

Bu çalışma, yatırımcıların içerik odaklı çevrimiçi forumda ifade ettikleri duygularının Bitcoin fiyatı üzerindeki etkilerini araştırmayı amaçlamaktadır. Ocak 2016 ile Mayıs 2020 arasında bir süre için “Bitcointalk.org” kaynaklı 2.8 milyon forum mesajından oluşan büyük bir veri seti kullanılmıştır. Duygu, farklı sözlüklerin performanslarının ayrıntılı bir karşılaştırılmasının ardından, Hu Liu sözlüğü kullanılarak günlük olarak türetilmiştir. Zaman serisi verileri kullanılarak, yatırımcıların duyguları ve Bitcoin fiyatı arasındaki ilişki test edilmiştir, çalışma Bitcoin fiyatının yönü hakkında bilgi verebilecek diğer finansal değişkenleri de kapsamaktadır. Sonuçlarımız, Bitcoin fiyatlarının çevrimiçi forumlardan elde edilen duygular ile otoregresif bir ilişki sergilemediğini göstermektedir.

## ACKNOWLEDGEMENTS

I would like to express my sincerest gratitude to Prof. Abdullah Daşcı for his patient guidance and valuable support in my thesis. I consider myself a fortunate student who worked under the supervision of Assist. Prof. Ali Doruk Günaydın. Their ideas were invaluable and provided me guidance whenever I struggled throughout the thesis process. Finally, none of this would have been possible without the help and support of all business analytics graduate students and my family.



## TABLE OF CONTENTS

<b>LIST OF TABLES</b> .....	<b>viii</b>
<b>LIST OF FIGURES</b> .....	<b>x</b>
<b>1. INTRODUCTION</b> .....	<b>1</b>
<b>2. LITERATURE REVIEW</b> .....	<b>3</b>
<b>3. SENTIMENT ANALYSIS</b> .....	<b>8</b>
<b>4. COLLECTION &amp; DESCRIPTIVE ANALYSIS OF DATA</b> .....	<b>10</b>
4.1. Bitcointalk.org .....	10
4.2. Financial metrics .....	14
<b>5. METHODOLOGY</b> .....	<b>16</b>
5.1. Sentiment analysis .....	16
5.2. VAR model .....	23
5.3. Analysis of weekly & daily data .....	25
<b>6. RESULTS</b> .....	<b>26</b>
6.1. The result of daily model .....	26
6.2. The result of weekly model .....	30
<b>7. CONCLUSION &amp; FUTURE WORK</b> .....	<b>32</b>
<b>BIBLIOGRAPHY</b> .....	<b>34</b>
<b>APPENDIX A</b> .....	<b>37</b>

## LIST OF TABLES

Table 4.1. Example of forum posts with details crawled .....	12
Table 4.2. Number of posts by board .....	13
Table 4.3. Number of posts per year .....	13
Table 5.1. The number of words in lexicons by category .....	18
Table 5.2. List of negation words in lexical model .....	18
Table 5.3. Example of forum posts with labels by different lexicons .....	20
Table 5.4. Prediction result of sample dataset with different lexical models	21
Table 5.5. Summary of prediction results with the sample dataset.....	22
Table 5.6. The distribution of posts per year by sentiment .....	22
Table 5.7. The distribution of posts by board and sentiment .....	23
Table 5.8. Stationarity tests on daily data .....	24
Table 5.9. Key measures and summary statistics - weekly .....	25
Table 5.10. Key measures and summary statistics - daily .....	25
Table 6.1. Statistical significance (p-values) of VAR model for Bitcoin return - daily .....	27
Table 6.2. Statistical significance (p-values) of VAR model with sentiment score for Bitcoin return - daily .....	28
Table 6.3. Statistical significance (p-values) of VAR model for Bitcoin return - 6 days cumulative .....	29
Table 6.4. Estimates of VAR model for Bitcoin return - 6 days cumulative	29
Table 6.5. Statistical significance (p-values) of VAR model for Bitcoin return - weekly .....	30
Table 6.6. Statistical significance (p-values) of VAR model for Bitcoin return - 2 weeks cumulative .....	30
Table 6.7. Estimates of VAR model for Bitcoin return - 2 weeks cumulative	31
Table 6.8. Statistical significance (p-values) of VAR model for Bitcoin return - 3 weeks cumulative .....	31
Table 6.9. Statistical significance (p-values) of VAR model for Bitcoin return - 4 weeks cumulative .....	31



Table A.1. The distribution of posts per year by sentiment (as per LM+HL lexicon) .....	37
Table A.2. Number of posts by boards (as per LM+HL lexicon) .....	37
Table A.3. Statistical significance (p-values) of VAR model for Bitcoin return - daily .....	38
Table A.4. Statistical significance (p-values) of VAR model for Bitcoin return - daily .....	38
Table A.5. Statistical significance (p-values) of VAR model for Bitcoin return - 6 days cumulative .....	39
Table A.6. Statistical significance (p-values) of VAR model for Bitcoin return - weekly .....	39
Table A.7. Statistical significance (p-values) of VAR model for Bitcoin return - 2 weeks cumulative .....	40
Table A.8. Statistical significance (p-values) of VAR model for Bitcoin return - 3 weeks cumulative .....	40
Table A.9. Statistical significance (p-values) of VAR model for Bitcoin return - 4 weeks cumulative .....	40

## LIST OF FIGURES

Figure 4.1. Map of the forum “bitcointalk.org” .....	11
Figure 4.2. Financial data from Jan-16 to May-20 .....	14
Figure 4.3. Correlation matrix for financial variables .....	15
Figure 5.1. Distribution of monthly number of positive and negative posts between 2-Jan-16 and 22-May-20 .....	23

## 1. INTRODUCTION

The way people interact was changed massively with the broader use of the Internet, consecutively led to the emergence of new exchange platforms. The enhancement of peer-to-peer networks and cryptographic platforms resulted in the emergence of digital currencies such as Bitcoin, Ethereum, and Ripple as the most popular ones.

A digital currency can be defined as a decentralized currency that is not controlled by an authority. It does not have a physical form and can be transferred directly to the people over the network. It is not issued by a bank or a government and could be created through a process called “mining”. “Mining” is the only source to increase the amount of digital currency available in the market, indicating the limited availability. Due to the volatility in their prices, cryptocurrencies attract attention from both investors and researchers - Bitcoin’s price increased from zero in 2009 to US\$ 19,891 in December 2017 (highest ever).

Many large companies have been accepting Bitcoin as a legitimate source of payment. For example, Wikipedia has been accepting donations in Bitcoin and Microsoft allows the use of Bitcoin to purchase goods through the Xbox store. As one of the leading travel websites, Expedia integrated a payment solution to start accepting Bitcoin in 2014. The Libra Association will plan to introduce a digital currency, a global, digitally native, reserve-backed cryptocurrency backed by blockchain technology. People will be able to send, receive, spend, and secure their money, enabling a more integrated global financial system. Polasik (2015) stresses the importance of the technology behind the creation of cryptocurrencies and its impact on the e-commerce processes.

Contrary to the fiat currencies that are supplied and controlled by a governmental institution such as central bank, digital currencies are not managed by a legal authority, can be received either through the peer-to-peer exchange or through the mining application, using a publicly known algorithm. While traditional fiat currencies are priced through market conditions – demand and supply in the market, the supply of digital currencies are almost fixed, total cryptocurrency volume in the market

depends on the mining efforts. The demand for cryptocurrencies is driven not by macroeconomic developments of an underlying economy but by the expectations of people about its future price as there is no profit (like interest or dividend). Since there is no underlying economy behind the pricing of cryptocurrencies, investor's expectations about the future trends become more important to foresee the price developments in the cryptocurrency market. Therefore, it necessitates a better understanding of the investor's sentiment. It is the users' believe which affect the value of the currency and therefore, understanding users' motivation for their activities in Bitcoin market may inform about the future fluctuations in the cryptocurrency prices.

The sentiment analysis is a rapidly developing field of natural language processing (NLP) and is now widely used in economic and financial research. Social media has been a source of investor's emotions to predict the stock market returns in finance literature. Zhang et al. (2011) predicted stock market indicators, NASDAQ, S&P 500, Dow Jones, and VIX measuring hope and fear in tweets each day. They found a significant correlation and concluded that twitter for emotional outbursts of any kind gives a good prediction of the market will be doing the next day. Bollen et al. (2011) found that the accuracy of DIJA prediction could significantly be improved by the inclusion of public mood. Using two mood tracking tools, namely OpinionFinder (positive vs. negative mood) and Google-Profile Mood States (calm, alert, sure, vital, kind, and happy), they measured the ability to predict changes of DIJA closing values.

In this study, we aim to understand whether sentiment from the Bitcoin forum impacts the return of Bitcoin, which has the highest market capitalization among other cryptocurrencies. Using lexical techniques, we derived the investor sentiment from their posts in a Bitcoin-specific online forum. To understand the impact, we run vector autoregressive models for a period of Jan-16 to May-20. Our results indicate that the sentiment derived from online forum do not offer predictive information for the changes in Bitcoin price. The impact is not visible in daily or weekly observations and boards in the forum do not present any difference.

This thesis is organized as follows: Chapter 2 reviews the existing literature on predicting financial metrics using the sentiment. In Chapter 3, we provide general methodologies for performing sentiment analysis. In Chapter 4, we describe our data collection, cleaning, and structuring procedures. Chapter 5 presents the methodology used in this study. Chapter 6 discusses the predictive analysis models and their results, followed by the conclusion in Chapter 7.

## 2. LITERATURE REVIEW

In this section, we will summarize the studies on predicting financial indicators - mainly the cryptocurrency price movements- using sentiment analysis. Stock market prediction using sentiment analysis has been a research area for academia for a long time while cryptocurrencies are new to the field.

Bollen et al. (2010) investigated whether public mood states from daily tweets measured by OpinionFinder and Google-Profile of Mood States (GPOMS) could predict the changes in the Dow Jones Industrial Average (DIJA). Through Granger Causality Analysis and neural networks, they achieved an accuracy of 86% in predicting daily changes in the closing value of DIJA.

Zhang et al. (2011) predicted stock market indicators by analyzing Twitter posts. They filtered Twitter posts by words indicating “hope” and “worry”. They found that the percentage of emotional tweets significantly negatively correlated with Dow Jones, NASDAQ, and S&P500 but positively correlated with VIX.

Smailović et al. (2012) used volume and sentiment polarity of Apple financial tweets to predict future movements in Apple stock price. Granger causality analysis revealed that future prices could be predicted for two days. For sentiment analysis, they used the dataset from Stanford University containing 1.6m labeled tweets where positive and negative emoticons were used as labels.

Wang et al. (2014) developed a sentiment analysis tool using datasets from SeekingAlpha and Stocktwits to analyze the historical stock performance. They applied Loughran and McDonald dictionary for SeekingAlpha and were able to get 85% on the validation set with the lexical model while “bullish, bearish” labels in the Stocktwits dataset was used as training data in a machine learning model. With the Support Vector Machines (SVM) model applied to the Stocktwits dataset, they achieved an accuracy of 76.2% on the validation dataset.

Li and Shah (2017) developed a finance domain lexicon with the aim to include different contextual meaning of words such that many terms in the financial context

have different meanings than those in other domains or sources. For example, terms such as “long”, “short”, “put”, and “call” have special meanings in the stock market context. They used the Stocktwits dataset, posts with bullish or bearish label to build Learning Sentiment Lexicon and Sentiment Oriented Word Embeddings (SOWE).

Studies in this field differ in i) the cryptocurrency that the study aims to predict, ii) the independent variables that they use for sentiment analysis -Twitter data, or microblogging. While some researcher constructs their studies on only Bitcoin, some are using other cryptocurrencies such as Ethereum or Ripple which have highest market capitalization after Bitcoin. There are also studies comparing the predictability of different cryptocurrencies. We have encountered only one study which incorporates other altcoins with lower market capitalization.

Valencia et al. (2019) applied their study to 4 cryptocurrencies -Bitcoin, Ethereum, Ripple, and Litecoin. They evaluated and compared the performance of three prediction models: Multi-Layer Perceptrons (MLPs), Support Vector Machines (SVM), and Random Forest (RF) using social data, market data, or both. Social data includes raw tweets from Twitter. Market data includes closing/opening price, highest/lowest price, and volume in the period sourced for selected cryptocurrencies. They concluded that at least with one model they can predict the direction of the market movements. Litecoin was the most predictable market, having the highest precision score, followed by Bitcoin and Ripple. Twitter has different predictive power in terms of explaining changes in cryptocurrencies while it can not used as the only source of data in any model.

Kim et al. (2016) crawled user comments and replies on online cryptocurrency communities for Bitcoin, Ethereum, and Ripple, and labelled user comments based on VADER algorithm. They found that positive user comments significantly affected price fluctuations of Bitcoin whereas negative user comments influence the price of Ethereum and Ripple. User opinions could predict the fluctuation in prices in 6-7 days. The predicted results were more precise for Bitcoin. This result was justified considering the higher amount and the greater activity of Bitcoin investors among other cryptocurrencies.

In a later paper, Kim et al. (2017) included Google trends data and Wikipedia usage besides Bitcoin-related posts in an online forum (bitcointalk.org). They aimed to conclude more general features rather than a polarity such as positive & negative in sentiment analysis. Using concept building methods, they generated 10 concepts: mining, transaction, Silkroad, security, illegal, blockchain, altcoin, wallet, china, investment. The Granger Causality test showed that the concept ‘China’ indicates

strong cause-effect relationship with bitcoin price while the concepts ‘Blockchain’, ‘Altcoin’, and ‘Transaction’ have with Bitcoin transaction count.

Kjærland et al. (2018) examined the relationship between volume and Bitcoin’s price as well as political incidents’ impact on the price. News regarding legal concerns about Bitcoin in the United States and China, tax decisions regarding Bitcoin investments in the United States and the EU, the shutdown of the Silk Road, and the Cyprus banking crisis are some examples of political incidents. Using the autoregressive distributed lag model, they showed that Bitcoin’s price has a significant relationship to the variables Google search, volume, positive & negative shocks. The volume of Bitcoin has a significant negative relationship with Bitcoin price while Google search and Bitcoin price have a significant positive relationship. Political incidents and government releases regarding Bitcoin affect Bitcoin’s price when the news is announced while negative news shocks are significant at a higher level than positive shocks.

Kaminski (2014) using Twitter data concluded that Twitter is Bitcoin’s virtual trading floor, emotionally reflecting its trading dynamics. Additionally, he found that there is no statistical significance for Twitter signals as a predictor of Bitcoin’s close price, intraday spread, or intraday return. On the contrary, higher trading volumes Granger cause more signals of uncertainty within a 24 to 72-hour timeframe.

Kristoufek (2014) studied the relationship between Bitcoin price and search queries on Google Trends and Wikipedia for the period from May 1, 2011 and Jun 30, 2013 (total of 788 observations). He found a bidirectional relationship: Not only the price influences search queries, but also the search query impacts the price. Additionally, he concluded bubble-curst behavior; if the prices are high, the increasing interest leads to higher prices and vice versa.

Matta et al. (2015) investigated if the spread of Bitcoin’s price is related to the volumes of tweets or web search media results. Using “SentiStrength” which estimates the polarity as positive/negative sentiments of a text, they classified tweets posted between Jan- March 2015 (c.2 million tweets). In their model, they used independent variables: # of tweets, # of tweets with positive mood, and google search trend. They showed that a positive mood could predict the bitcoins price almost 3-4 days in advance. Google Trends data have a cross-correlation value of 0.64 with a time lag 0 days, showing it is a strong predictor.

Hernandez et al. (2014) studied whether Bitcoin users are less sociable based on the language used by the user and their social connections on Twitter. They concluded that Bitcoin followers are less likely to mention family, friends, religion, sex, and

emotion-related words in their tweets, and have a significantly less social connection to other users on the site. Their study has implications for distinct behavioral features of Bitcoin users, hence the future of the currency.

Xie et al. (2017) worked with i) forum postings from bitcointalk.org (average fraction of negative words in postings in economics, speculation, trading sections on the forum), ii) articles from FACTIVA for traditional media sentiment, iii) volatility (calculated as the sum of squared daily returns during the previous calendar month). They found that information related to fundamentals (i.e. economics section) predicts only long-term price changes and that information related to speculative topics (i.e. speculation, trading sections) predicts both long and short-term price changes. Using several user features that show the activity level of users, they found that the inactive users provide more information for future bitcoin movements than active users and their predictive power differs in the type of section used for the analysis.

Steinert and Herff (2018) collected data set containing price and social media activity for 181 altcoins for a timeframe of 72 days with 426k tweets. After initial descriptive analysis, they included altcoins that were referred by tweets on at least 10% of all days, so the total # of altcoins decreased from 181 to 131. They applied ordinary least squares regression analysis using altcoin return as dependent variable, they found statistically significant data for at least one-time lag for altcoins “Bitcoindark, Ethereum, Purevidz, Steem dollars, Voxels”.

Mai et al. (2018) analyzed the effects of users with different levels of activity - The active users who contribute most content (the vocal minority) and relatively inactive users who contribute less content (the silent majority). Most users belong to the silent majority, however the vocal minority constituting small portion generated most of the content. They also investigated how messages from twitter and the Internet forum affect the Bitcoin market differently. They found that i) social media metrics significantly affect future bitcoin prices, such that increased positive (negative) sentiments indicate higher (lower) future Bitcoin prices, ii) the predictive power of social data depends mostly on the information derived from content created by silent majority rather than the one created by the vocal minority, and iii) user-generated content from online community rather than from Twitter, has a stronger impact fluctuations in Bitcoin price.

Garcia and Schweitzer (2015) used Reddit posts and Google News to predict future Bitcoin value. They used VADER, TextBlob, and Flair lexical models to analyze sentiment and performed neural networks comparing different models, using either historical price data, sentiment data, or both. Their result indicates that the best model was the one that includes all parameters.



Abraham et al. (2018) , using tweets and Google Trends data, were able to accurately predict the direction of price movements in the Bitcoin and Ethereum markets. They found that tweet volume rather than tweet sentiment (analyzed through VADER) provides better information to predict future price movements.



### 3. SENTIMENT ANALYSIS

Microblogging websites have evolved to become a source of information in various fields (Agarwal et al., 2011). Microblogs reflect real-time opinions of people, showing their posts real-time messages about their opinions on a variety of topics. Considering that behavioral economics tells us emotions can profoundly affect individual behavior and decision making (Akerlof & Schiller, 2010), microblogging websites attract researchers due to the large amount of information they contain, as a source of user's opinion and hence, sentiment analysis (Pak & Paroubek, 2010).

Sentiment analysis, or opinion mining, is a study area in the field of natural language processing that analyzes people's opinions, sentiments, evaluations, attitudes, and emotions through the computational calculation of subjectivity in a text (Hutto & Gilbert, 2014). Sentiment analysis is widely used in social media as it allows us to obtain an overview of general opinion behind a specific topic. It is a classification problem (positive, negative, neutral) or a rating problem (valence measure).

In general, sentiment analysis techniques can be divided into two; lexicon-based methods and machine learning methods while hybrid approaches are also used. For example, Popowich and Moghaddam (2010) used a Naïve Bayes classifier to identify the polarity of an adjective while computing similarity values of adjectives through the WordNet lexicon.

Machine learning (ML) approaches use previously labeled data to learn sentiment related features of the text and predict the sentiment of newly encountered data. The machine learning model is trained on the dataset where textual context was linked with a sentiment label assigned by humans. ML approaches can leverage linear classifiers or deep learning models to automatically learn sentiment features for words and entire texts and learn how to derive a sentiment score for the whole text (Shapiro et al., 2017).

Machine learning techniques are widely used for opinion mining from Twitter. Pak and Paroubek (2010) aimed to build a sentiment classifier that can determine positive, negative, and neutral sentiment for documents. They constructed a simple

binary classifier that used n-gram and POS features and trained on instances that had been annotated according to the existence of positive and negative emoticons.

Scientists may prefer machine learning approaches over the lexicon-based approaches as i) they give better accuracy with a great volume of data, and ii) creating and validating a comprehensive lexicon is both labor and time-intensive (Hutto & Gilbert, 2014). However, it is problematic in applications where several different domains, languages, and text data involved as models have to be trained for each one (Maynard et al. 2012). ML approaches have the following shortcomings: i) they require labeled data to train, ii) training data set do not always inquire several features especially in the context of short social media postings, and iii) in terms of computational requirement, they are more expensive.

Lexical approaches aim to match the word to sentiment by using a lexicon/dictionary. The selected method depends on the pre-defined list of words-called lexicon or dictionary, with labels or assigned scores indicating negativity or positivity. Generally, the score is between -1 and +1 such that -1 indicates negativity, +1 indicates positivity. Some methods (such as VADER) have valence measure besides the three categories of positive/negative/neutral sentiment. There is no consensus on which dictionary performs best. It depends on following i) which feature space the dictionary was created on, and ii) which content it will be used for.

Word matching is the basic principle of the method. Additionally, negation, lemmatization, and part-of-speech methods can improve accuracy. The advantage of the lexical approaches over other methods lies in that one does not need to train the model with labeled data. However, they have shortcomings such that i) they ignore general sentiment intensity for features within the lexicon, and ii) they are time consuming as building a lexicon acquires a new set of human validated lexical features. When lexical methods are used, the result of sentiment analysis is heavily impacted by the chosen lexicon.

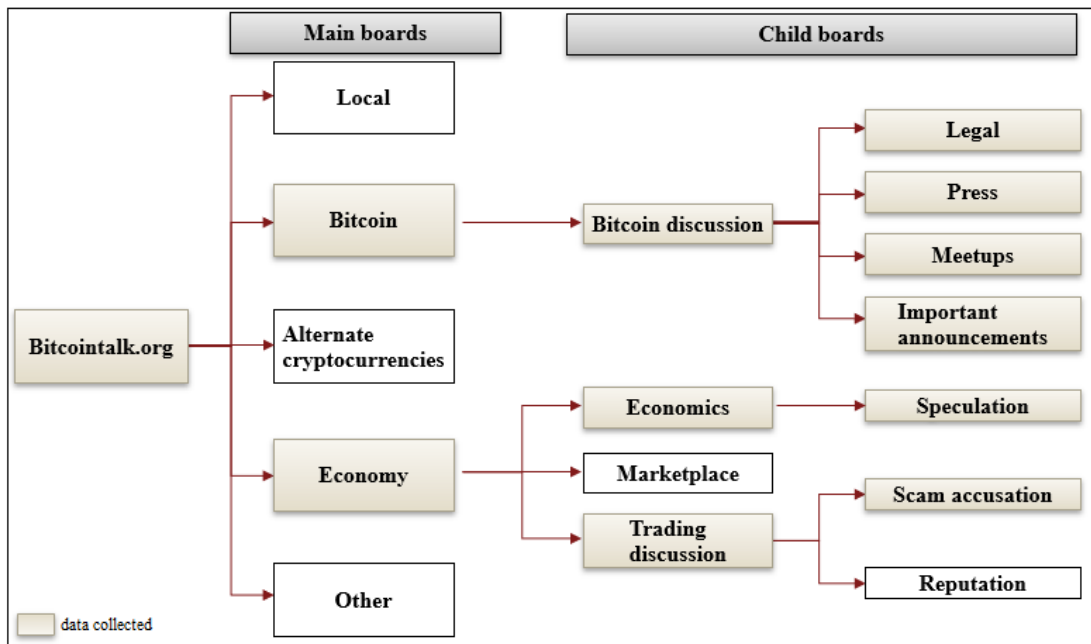
## 4. COLLECTION & DESCRIPTIVE ANALYSIS OF DATA

To analyze what impact Bitcoin return, we have incorporated several data sources. The first input we used was the user's threads from the forum "bitcointalk.org". Each post was assessed for a sentiment score. We have downloaded Bitcoin price along with other financial variables, the S&P500 index, VIX index, and gold price sourced from Yahoo!Finance. In this chapter, we aim to present the data collection process, data sources used, and data cleaning & transformation processes. Additionally, we provide a descriptive analysis of the forum data along with the financial measures.

### 4.1 Bitcointalk.org

We collected forum threads from "bitcointalk.org" which is the main online community where people share their ideas about bitcoin, blockchain technology, and cryptocurrencies. The forum was initiated by Satoshi Nakamoto, the inventor of Bitcoin in 2009. It is one of the most active forums in the cryptocurrency domain. It has five mainboards with several child boards that are originated from them. We have assessed the boards based on their content and decided to crawl data from "Bitcoin" and "Economy" mainboards. We excluded the "Local" mainboard because it contains discussions on Bitcoin in local languages such as Chinese, German, etc. Additionally, we did not crawl threads from the "Alternate currencies" that contains discussion in other cryptocurrencies. The mainboard "Other" was also excluded since it contains posts on about the forum itself and off-topics. In Figure 4.1, we have shown the main and child boards chosen for this study.

Figure 4.1 Map of the forum “bitcointalk.org”



This figure maps the sections of the forum in our study. Please note that the mainboards which were excluded from the analysis also have child boards. We did not show them for simplicity.

Comments and relevant replies posted by users were crawled from each selected board. The overlapping text or replies quoting previous comments and replies were excluded. We have crawled the following details regarding each thread:

- the time when each comment and reply was posted,
- the author of the thread and their membership status including activity level and merit,
- the subject of the post

We have shown a randomly chosen sample from the forum data in Table 4.1.

Table 4.1 Example of forum posts with details crawled

Author	Post	Subject	Hour	Membership level	Offline activity	Merit	Date	Board
shulio	What is about new Chinese legislation? How long will it take to them to close all websites concerning ICO ? Who knows?	List of court cases, complaints, regulatory actions, etc.	05:15:41 PM	Legendary	1,540	1,016	9/4/2017	Legal
Harriti	After the strong growth of Bitcoin, I believe it will soon collapse again. The price of BTC is sideway to wait for more fries to buy bitcoin, then they will kick down the price of bitcoin again. Bitcoin halving is usually like that, we should set a lower buy price for bitcoin and wait. Buying BTC now is a pretty risky decision.	BTC Price might drop now?!	05:02:14 AM	Sr. Member	602	251	5/17/2020	Speculation
duts_bg	The Bitcoin is a native response that gives the life , to the degenerated financial system. He returns money to their natural role as a medium of exchange. It takes the authority of politicians to artificially manipulate the economy.	Inflation and Deflation of Price and Money Supply	01:23:11 PM	Full Member	167	100	5/6/2016	Economics

The data from the forum was extracted through R-Studio and saved in a csv file. The regex was applied in Alteryx to have a clean and structured dataset, i.e. ads, double spaces were removed. The challenge was to scrap the data in bulk. Since the website has access limit per minute, crawling attempts over this limit were denied by the website automatically. It implied the process that we crawled manually for every 20 pages in each child board. The crawling period took over one month.

The forum posts were crawled for the period between Jan 2, 2016 and May 22, 2020. The data was traced back to 2016 specifically in order to include two halving periods in the study. A total of 2.8m unique threads were collected for this period. 2018 was the most active year of the forum, in line with the popularity of Bitcoin and with an increase in price and transaction volume. Users were mostly active in the “Discussion” board which was followed by “Speculation”.

Table 4.2 Number of posts by board

	# of posts	Total %
Discussion	1,400,372	49.5%
Speculation	505,923	17.9%
Trading	442,450	15.6%
Economics	376,826	13.3%
Press	64,536	2.3%
Legal	34,616	1.2%
Meetups	3,388	0.1%
<b>Total</b>	<b>2,828,111</b>	<b>100.0%</b>

This table presents the number of posts by boards, which we have crawled for the period between 2-Jan-2016 and 22-May-2020.

Table 4.3 Number of posts per year

	# of posts	Total %
2016	386,574	13.7%
2017	770,845	27.3%
2018	1,225,594	43.3%
2019	353,788	12.5%
2020	91,310	3.2%
<b>Total</b>	<b>2,828,111</b>	<b>100.0%</b>

This table presents the number of posts by year. Please note that 2020 figures are until 22-May.

## 4.2 Financial metrics

We have downloaded Bitcoin price, volume, and transaction volume from Bitstamp Ltd, the top Bitcoin exchange platform. The Bitcoin-related financial data was available on a 1-hour basis.

Additionally, we downloaded the S&P500 index, stock market volatility (VIX index from Chicago Board Options Exchange), COMEX gold price sourced from Yahoo!Finance. Since Bitcoin is traded 24-hour basis, we set the open-close price based on NY time. Financial data were collected for the period between Jan 2, 2016 and May 22, 2020 in line with the forum data collection. In Figure 4.2, we showed the trend of bitcoin price (on the left axis), S&P500 daily close index, and Gold daily close price (on the secondary axis). The Bitcoin price was very volatile especially in 2018 when it peaked at US\$18K.

Figure 4.2 Financial data from Jan-16 to May-20



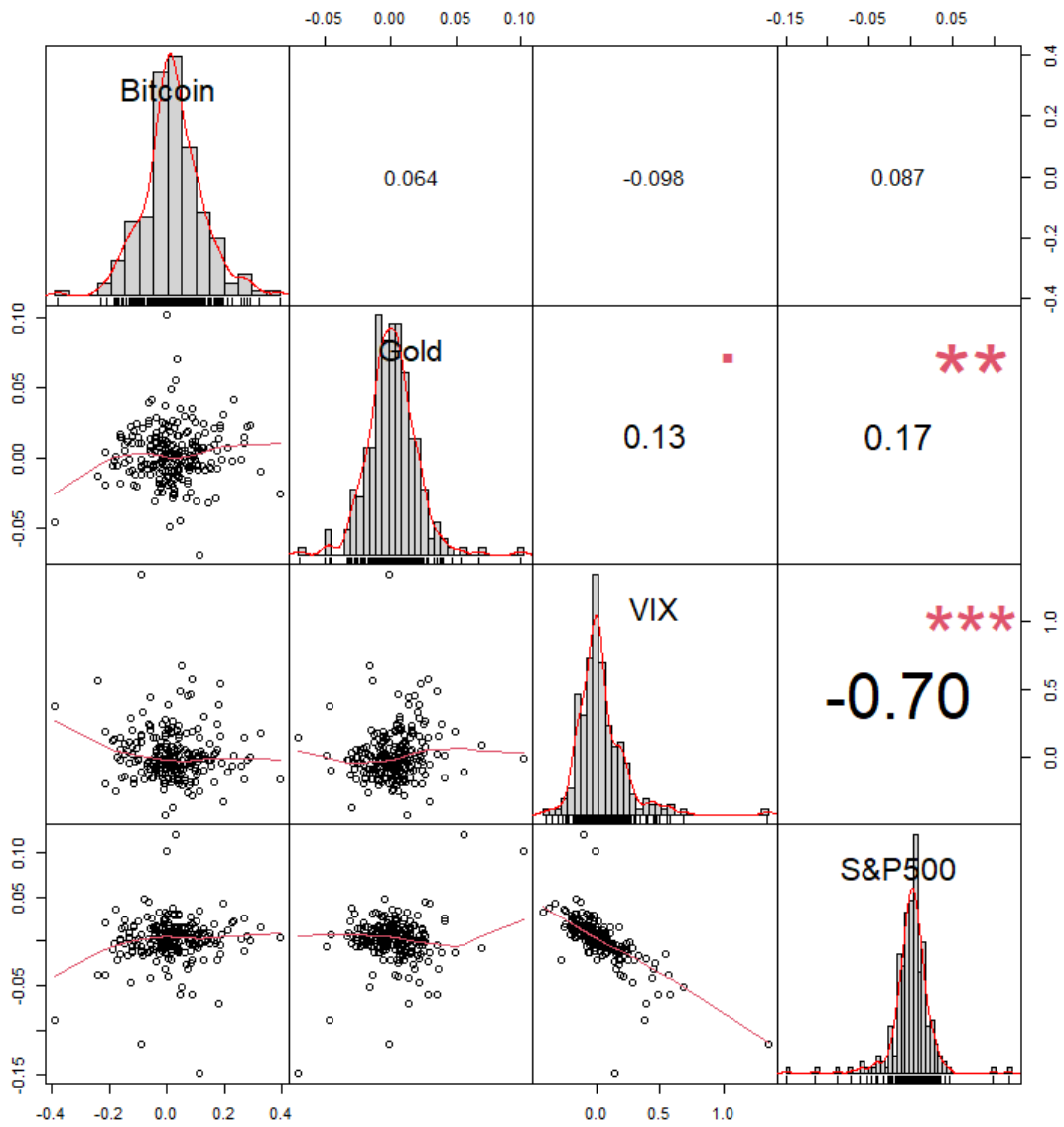
This figure shows the evolution of Bitcoin daily close price, S&P500 daily close index and gold daily close price. Since Bitcoin was traded on 24-hour basis, NY EDT 12:00 am was chosen as close price.

\* S&P 500 and gold daily close price are represented on the secondary axis.

In Figure 4.3, we have presented the correlation matrix of financial variables. The distribution of each variable is shown on the diagonal. The bivariate scatter plots with a fitted line are displayed on the bottom of the diagonal. The value of the correlation plus the significance level as stars are on the top of the diagonal. The correlation coefficients indicate that the Bitcoin price is positively correlated with the Gold price and S&P500 index, but the correlation is negative with VIX index, as expected. However, all three correlation are not strong. Additionally, VIX and S&P500 indices show significant negative correlation.



Figure 4.3 Correlation matrix for financial variables



## 5. METHODOLOGY

### 5.1 Sentiment analysis

In this study, we have decided to follow the lexical-based approach for two reasons. Firstly, a labeled dataset for forum post was not available for the training dataset. Secondly, previous research suggests that machine learning methods do not have significant advantages over lexical approaches (Hutto & Gilbert, 2014). Previous studies examined several lexicons such as LIWC, VADER, SentiStrength, Textblob, Loughran & McDonald, Hu & Liu. Steinert and Herff (2018), Kim et al. (2016), Prajapati (2020), Valencia et al. (2019), Garcia et al. (2015) used VADER in the sentiment analysis part of their studies. Shapiro and Wilson (2017) applied the Loughran&McDonald lexicon to FOMC meetings transcripts. Wang et al. (2014) applied Loughran McDonald to the SeekingAlpha dataset. Matta et al. (2015) modeled “Sentistrength” to the Twitter dataset.

Previous research was mostly either on short-text like tweets or long-texts like news articles. Each text source has its own features, therefore the lexicon to be applied should be chosen accordingly. In our case, forum posts are not always as short as tweets while they contain slang language of social media. Considering the previous studies on different mediums, we have applied five lexicons to the forum dataset to compare their performance.

- **Loughran and McDonald (will be referred as LM):** Loughran and McDonald (2011) used a large sample of SEC 10-Ks filings to create a finance-specific word list, arguing that the Harvard lexicon considers some words as negative while they are neutral when used finance domain. For example, the word “liability” is generally neutral in finance context but negative in daily language. Their lexicon is now widely used in finance literature. It contains 2355

negative words and 354 positive words. Shapiro and Wilson (2019) used the LM dictionary to derive the tone of language in FOMC transcripts to estimate the FOMC loss function, including the implicit inflation target. In the study to understand the relationship between news-based measures of sentiment and GDP, Fraiberger (2016) created a combined dictionary from Loughran and McDonald (2011) for finance context and Young and Soroka (2012) for political context. Mai et al. (2018) applied the LM dictionary to posts from online community to analyze impact of sentiments on price.

- **VADER** was another extensively used method. It is an open-source tool developed by Hutto and Gilbert in 2014. It is a simple rule-based model for general sentiment analysis that is specifically attuned to sentiment in the social media context. Their algorithm incorporates several features impacting intensity such as slang language, punctuation, capitalization, degree modifiers, the conjunction “but”, negation flips by examining trigrams, being preceded by a degree modifier such as “very, extremely, slightly,” etc. VADER provides positive and negative valence scores as well as a normalized, weighted compound score for each text. In this study, we used the compound score. We classified the post as positive if this score is greater than 0, as negative if less than 0, and neutral if 0.
- **Hu Liu** lexicon was developed using online movie reviews where the movie reviews are assigned as positive/negative by the reviewers themselves. Although it is not a finance/economics domain-specific lexicon, Shapiro and Sudhof applying it to financial newspapers achieved close results with a finance-specific lexicon on the same dataset.
- **SentiWordNet**: Sebastiani and Esuli developed Sentiwordnet, a lexical resource for opinion mining. SentiWordNet assigns to each word one of three sentiment scores: positivity, negativity, objectivity.
- **Harvard General Inquirer** : The GI dictionary was one of the earliest and prevalent valence lexicons, consisting of 3,626 words labeled positive or negative. It is meant to be a general English language lexicon. It has categories such as “words of pleasure”, “motivation-related words”, “cognitive orientation”, “two large valence categories” etc. We have used the words in “two large valence categories” for the positive and negative outlook.

In Table 5.1, we summarized the word count in each dictionary that we have examined. The table shows the count of positive / negative / neutral words in each lexicon. "LM + HL" represents the combined lexicon of Loughran McDonald and

Hu Liu lexicons.

Table 5.1 The number of words in lexicons by category

	Loughran & McDonald	Hu Liu	LM + HL	Sentiword	Harvard GI
Positive	354	2024	2178	11029	1915
Negative	2355	4837	6350	8898	2291
Neutral	0	13	13	166	0
<b>Total</b>	<b>2,709</b>	<b>6,874</b>	<b>8,541</b>	<b>20,093</b>	<b>4,206</b>

VADER algorithm itself deals with negation but other lexicons do not account for negated words (i.e. “I don’t feel well” and “I feel well” both will be labeled as positive.) To cope with this problem, we divided the posts into bigrams. Then we reversed the sentiment polarity of words preceded by the negation word in a bigram (i.e. if the word “good” is preceded by the word “not” – bigram “ not good”- it is counted as negative, while it was originally labeled as positive in the lexicon.) For this operation, we used negation words defined as in “‘qdapDictionaries’” library of R-Studio. The list of 24 negation words is shown in Table 5.2.

Table 5.2 List of negation words in lexical model

ain’t	don’t	neither	shan’t
aren’t	hasn’t	never	shouldn’t
can’t	haven’t	no	wasn’t
couldn’t	isn’t	nobody	weren’t
didn’t	mightn’t	nor	won’t
doesn’t	mustn’t	not	wouldn’t

By lexical models, we get number of positive, negative and neutral words for each post. Thereafter, we used the established method of calculating sentiment polarity of the post (Twedt and Rees, 2012; Kearney and Liu, 2014) :

$$Sentiment\ score = \frac{N_{i,pos} - N_{i,neg}}{N_{i,pos} + N_{i,neg}}$$

$$Daily\ sentiment\ score_t = \frac{\sum Sentiment\ score_i}{\sum Number\ of\ posts_i}$$

We had a sample of 1000 posts from our dataset, labelled by us as positive, negative or neutral. Using 1000 randomly selected posts from different boards in the forum,

we evaluate the performance of the alternative lexicons. We applied VADER method with the cranR package “getVaderRuleBasedSentiment”.

The result from the sample dataset was summarized in Table 5.4 and 5.5. The accuracy measures in the table represent the ability of the lexicon to classify text correctly into the discrete categories. The results in the table show that the predictive accuracy of Hu Liu and LM lexical models are similar, each dominated the other lexical models and VADER. Since LM is a domain-specific lexicon, the better accuracy was expected. Hu Liu, on the other hand, gains from being a larger lexicon compared to LM and Harvard GI and having a feature space movie reviews with labels by "reviewer". While the accuracy result of LM was slightly worse than the Hu Liu lexical model, its classification performance was impacted by the lexicon size, it could not classify 23% of posts (meaning that 23% of posts have words which the LM lexicon does not contain). Next, we combined LM and Hu Liu, finance-specific lexicon, and the best performing lexicon to check whether the performance can be improved by combining lexicons. For words included in both lexicons, LM was preferred. The total lexicon size increased to 8541 words. As shown in the table, accuracy performance increased as compared to the LM lexical model, the total classified number of posts increased as compared to both Hu Liu and LM lexical models. However, in terms of accuracy, it performed worse than Hu Liu. The accuracy of SentiWordNet was still better than random-over 50% but less than LM and Hu Liu lexical models. Surprisingly since many previous studies applied VADER as a recognized method, VADER performed worst among all (33% accuracy). One explanation may be that previous work was mostly on Twitter or Stocktwits, short-length social media text, for which VADER was attuned.

Table 5.3 Example of forum posts with labels by different lexicons

Post	Human rater	Loughran McDonald	Hu Liu	LM + HL	Sentiword	VADER	Harvard GI
This is still a far cry from the adoption that we wanted to see from Amazon. Just some company (and a shady one) decided to create a browser extension and tie it to your account to be able to shop in Amazon. Really scary, and I wouldn't even used and download that extension. You don't want to be caught in the middle with high fees and the likelihood that your account can be compromised.	negative	positive	negative	negative	positive	negative	neutral
Not all people liked bitcoin even though they use this forum but don't mean they will supporting cryptocurrency but for us who really support bitcoin i think whatever they talking bad or negative about bitcoin but i think the real cryptocurrency users will ignoring what they said because we believe cryptocurrency community will bigger than now	positive	negative	positive	neutral	negative	negative	positive

Table 5.4 Prediction result of sample dataset with different lexical models

		Loughran McDonald			Hu Liu			LM + HL		
		Predicted label			Predicted label			Predicted label		
Actual label		positive	negative	neutral	positive	negative	neutral	positive	negative	neutral
	positive	63%	23%	14%	79%	13%	8%	76%	13%	10%
	negative	23%	65%	13%	33%	57%	10%	31%	56%	13%
neutral	38%	47%	15%	29%	24%	47%	25%	32%	44%	
		Sentiword			VADER			Harvard General Inquirer		
		Predicted label			Predicted label			Predicted label		
Actual label		positive	negative	neutral	positive	negative	neutral	positive	negative	neutral
	positive	74%	19%	7%	39%	19%	43%	64%	29%	7%
	negative	61%	28%	12%	37%	22%	41%	22%	68%	10%
neutral	62%	22%	16%	45%	20%	43%	29%	50%	21%	

The overall classification result was summarized in Table 5.5. Based on the result we had on the sample dataset, we have applied the Hu Liu lexicon model to the forum dataset.

Table 5.5 Summary of prediction results with the sample dataset

	Loughran Hu & Liu McDonald	LM + HL	Senti word	VADER	Harvard GI	
Accuracy	62%	69%	67%	52%	33%	61%
# of posts not classified	231	84	66	9	-	1
# of posts classified	769	916	934	912	1,000	994
# of posts true classified	473	631	624	477	333	603
# of posts false classified	296	285	310	435	667	391

The forum dataset consisted of 2,828,111 posts of which 2,556,272 Hu Liu lexical-model were able to label (90% of total post). As shown on the table 5.6, 58% of posts were labeled as positive in 2016 and 2017. The positive posts' share decreased slightly to 56% in 2018 and further decreased to 52% in 2020 while negative posts share increased in the same period. Considering that our crawling period comprises the Covid-19 period, a significant decrease in positivity in 2020 was expected.

Table 5.6 The distribution of posts per year by sentiment

	Positive	Negative	Neutral
2016	58%	29%	14%
2017	58%	29%	13%
2018	56%	30%	14%
2019	56%	31%	14%
2020	52%	35%	13%
<b>Total</b>	<b>57%</b>	<b>30%</b>	<b>14%</b>

This table presents the number of posts by year. Please note that 2020 figures are until 22-May.

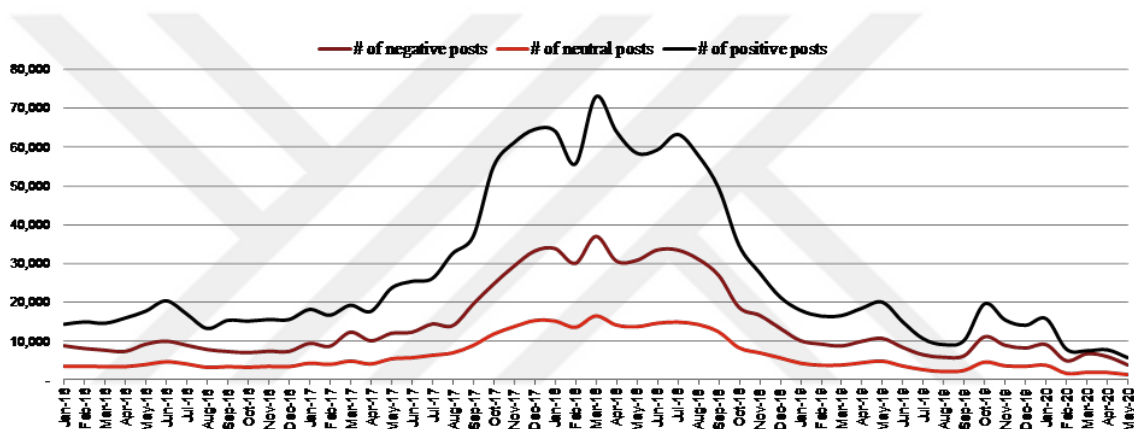
The classification by boards was shown in Table 5.7. The positive post count has the highest percentage among all boards while the distribution changes by the board. For example, 60% of the posts on the board "Trading" were classified as positive while the ratio dropped to 45% in the board "Legal" which could be explained with all the discussions about banning cryptocurrencies or recognizing them as legitimate.



Table 5.7 The distribution of posts by board and sentiment

	Positive	Negative	Neutral
Discussion	57%	29%	13%
Speculation	52%	33%	15%
Trading	60%	27%	13%
Economics	58%	28%	13%
Press	51%	36%	13%
Legal	45%	40%	14%
Meetups	85%	9%	6%
<b>Total</b>	<b>57%</b>	<b>30%</b>	<b>14%</b>

Figure 5.1 Distribution of monthly number of positive and negative posts between 2-Jan-16 and 22-May-20



In addition to Hu Liu lexicon, we have performed analysis with the combined lexicon of Hu Liu and Loughran & McDonald. The respective results were summarized in Appendix.

## 5.2 VAR model

In this study we aim to investigate whether there is a predictive relationship between activities in the forum and Bitcoin returns, dependent variable. To get better insight in the drivers of Bitcoin return we also add control variables such as gold return, VIX index and S&P500 return. To capture lagged effects of independent variables, we have chosen to estimate vector autoregressive models (VAR). By using VAR, the models may show how the different variables in earlier periods (in lags) affect

Bitcoin return.

Vector autoregression is a standard procedure for analyzing relationship between multiple series (Sims & Lütkepohl) . In a case of the pair of series  $x_t$  and  $y_t$ , the vector autoregression of order  $t$  (VAR( $t$ )) is written as

$$\Delta x_t = \alpha_1 + \sum_{k=0}^n \beta_{1k} \Delta x_{t_i} + \sum_{k=0}^n \gamma_{1k} \Delta y_{t_i} + \epsilon_{1t}$$

$$\Delta y_t = \alpha_2 + \sum_{k=0}^n \beta_{2k} \Delta y_{t_i} + \sum_{k=0}^n \gamma_{2k} \Delta x_{t_i} + \epsilon_{2t}$$

with possibly correlated disturbances  $\epsilon_{1t}$  and  $\epsilon_{2t}$  and lag selected according to an information criterion, such as the Akaike Information Criterion (AIC), Hannan-Quinn Information Criterion (HQIC) and Schwarz Information Criterion (SIC). We have chosen optimal lag with AIC in this study.

VAR model requires that series  $x_t$  and  $y_t$ , are  $I(1)$ , their first differences  $\Delta x_t$  and  $\Delta y_t$  are  $I(0)$  and thus stationary so that the system can be easily estimated using either the ordinary least squares or maximum likelihood procedures. For testing stationarity, we utilize the Augmented Dickey-Fuller test (ADF). ADF has a null hypothesis of a unit root against the alternative of no unit root. Table 5.8 shows the result of stationarity test on daily dataset, indicates that all variables are stationary, and VAR model can be applied.

Table 5.8 Stationarity tests on daily data

Variable	ADF	p-value
Bitcoin daily change	(11.214)	<0.01
Gold daily change	(12.443)	<0.01
S&P 500 index daily change	(10.765)	<0.01
VIX index daily change	(12.544)	<0.01
difference in daily # of negative posts	(15.791)	<0.01
difference in daily # of positive posts	(16.098)	<0.01
difference in daily sentiment score	(15.157)	<0.01

### 5.3 Analysis of weekly & daily data

In our thesis, we firstly examined models with data on weekly basis. In our time-frame, we have 229 weeks. We examine models which include weekly and daily observations for return in Bitcoin as well as changes in the number of negative posts and positive posts relative to previous week. Lastly, we included other financial metrics such as weekly/daily return in Gold, weekly change in the S&P500 and VIX indices. Table 5.9 summarizes the statistics of weekly observations. Additional in Table 5.10, daily observations could be traced.

Table 5.9 Key measures and summary statistics - weekly

Variable	Definition	Mean	SD	Median	Min	Max
$BTC_t$	Bitcoin weekly return	0.02	0.11	0.02	(0.39)	0.40
$NEG_t$	$\Delta$ in # of negative posts	(4)	465	(10)	(2,137)	1,426
$POS_t$	$\Delta$ in # of positive posts	(8)	695	5	(3,736)	2,182
$S\&P500_t$	Weekly return in S&P500 index	0.00	0.02	0.00	(0.15)	0.12
$VIX_t$	Weekly return in VIX index	0.02	0.19	(0.02)	(0.43)	1.35
$GOLD_t$	Weekly return in Gold price	0.00	0.02	0.00	(0.07)	0.10
$Sent_t$	Difference in weekly average sentiment score	(0.00)	0.04	0.00	(0.14)	0.10

Table 5.10 Key measures and summary statistics - daily

Variable	Definition	Mean	SD	Median	Min	Max
$BTC_k$	Bitcoin daily return	0.00	0.04	0.00	(0.36)	0.30
$NEG_k$	$\Delta$ in # of negative posts	(0)	82	(2)	(554)	500
$POS_k$	$\Delta$ in # of positive posts	(0)	130	(1)	(588)	926
$S\&P500_k$	Daily return in S&P500 index	0.00	0.01	0.00	(0.12)	0.09
$VIX_k$	Daily return in VIX index	0.00	0.08	0.00	(0.26)	1.16
$GOLD_k$	Daily return in Gold price	0.00	0.01	0.00	(0.05)	0.08
$Sent_k$	Difference in daily average sentiment score	(0.00)	0.05	0.00	(0.21)	0.17

## 6. RESULTS

### 6.1 The result of daily model

We have summarized the p-values of the daily model in Table 6.1. The model includes daily bitcoin return, and the social media variables, the difference in the number of negative posts count and positive posts count as well as other financial control variables. We selected the model with lag length  $k = 16$ , according to the Akaike information criteria. We expected to have those days with increases in the number of positive posts tend to precede days with an increase in Bitcoin return, but our model suggests that increases/decreases in the number of positive posts do not exhibit a strong autoregressive relationship. While the number of positive posts was not significant in any lag, we observed a significant relationship in lag=3 for the changes in the number of negative posts however the impact did not last in the next lags. For other financial variables, we have observed no significant relationship which lasts for more than one lag.

Table 6.1 Statistical significance (p-values) of VAR model for Bitcoin return - daily

Time lag	Daily return in Bitcoin price	$\Delta$ in # of negative posts	$\Delta$ in # of positive posts	Daily return in S&P500 index	Daily return in VIX index	Daily return in Gold price
1 days	0.0063**	0.4233	0.2536	0.2412	0.0173*	0.5984
2 days	0.3357	0.2357	0.1521	0.2345	0.2419	0.343
3 days	0.0726	0.0317*	0.1662	0.616	0.5566	0.8771
4 days	0.4303	0.3738	0.6697	0.8098	0.0616.	0.0487*
5 days	0.93	0.2138	0.646	0.5246	0.5003	0.3113
6 days	0.058	0.2947	0.5053	0.0041**	0.0001***	0.1969
7 days	0.3826	0.4404	0.2539	0.5686	0.3591	0.0918.
8 days	0.8321	0.7288	0.9028	0.5621	0.9218	0.2712
9 days	0.0515	0.7575	0.9753	0.9777	0.2062	0.2261
10 days	0.4217	0.366	0.7911	0.2091	0.512	0.0632.
11 days	0.6216	0.7383	0.1174	0.1752	0.2569	0.6853
12 days	0.1192	0.874	0.052	0.8768	0.5076	0.0569.
13 days	0.0699	0.2317	0.4121	0.2794	0.643	0.3544
14 days	0.5652	0.8917	0.7126	0.8833	0.4742	0.3922
15 days	0.5509	0.1583	0.2186	0.7927	0.3792	0.2975
16 days	0.4474	0.5425	0.177	0.7101	0.2601	0.3168

Significance codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Observing no statistically significant relationship between changes in negative/positive post counts and Bitcoin return, we further analyzed whether changing forum data metric would change results in VAR model. Instead of in # of negative posts and in # of positive posts, we have included the difference in average daily sentiment score. The results indicated that there is no statistically significant relationship between average sentiment score and Bitcoin return in any lag of 13 days (Table 6.2).

Table 6.2 Statistical significance (p-values) of VAR model with sentiment score for Bitcoin return - daily

Time lag	Daily return in Bitcoin price	$\Delta$ in daily sentiment score	Daily return in S&P500 index	Daily return in VIX index	Daily return in Gold price
1 days	0.0069**	0.1046	0.2309	0.0277*	0.6563
2 days	0.3253	0.1569	0.0899	0.1209	0.3336
3 days	0.1434	0.5533	0.7083	0.5342	0.9712
4 days	0.8334	0.5412	0.8568	0.0747	0.0301*
5 days	0.6614	0.9393	0.5346	0.5132	0.3825
6 days	0.0333*	0.3356	0.0042**	0.0013**	0.2057
7 days	0.4067	0.1947	0.7963	0.5577	0.07415
8 days	0.7962	0.5695	0.7076	0.7201	0.1754
9 days	0.074	0.6231	0.941	0.2121	0.4283
10 days	0.2669	0.8978	0.2346	0.5855	0.0675
11 days	0.7551	0.3042	0.0826	0.0792	0.745
12 days	0.1396	0.1287	0.6779	0.2412	0.0642
13 days	0.1012	0.3666	0.3531	0.804	0.3486

Additional to forum data metrics such as changes in the number of positive/negative data or changes in daily sentiment score, we examined whether the number of negative/positive words has any predictability in Bitcoin return. The result indicated that changes in the number of negative/positive words have no lagging effect on the Bitcoin return.

After investigating daily models, we have examined whether forum data has any predictability in terms of cumulative return in Bitcoin. We have modeled returns from 2 days to 8 days cumulative. In Tables 6.3 and 6.4, we have summarized statistical significance and estimates respectively. One more negative forum post is associated with a decrease in bitcoin return 10 days ahead by 2.43 basis points. The impact continues from 4 to 7 days lag however the estimates in 6&7 days lag are positive, contrary to the expected negative relationship. The model also suffers from autocorrelation as a cumulative return in  $day_t$  contains information from the previous 6 days.

Table 6.3 Statistical significance (p-values) of VAR model for Bitcoin return - 6 days cumulative

Time lag	6 days cumulative return in Bitcoin price	$\Delta$ in # of negative posts	$\Delta$ in # of positive posts	Daily return in S&P500 index	Daily return in VIX index	Daily return in Gold price
1 days	< 2e-16***	0.1213	0.2031	0.0958	< 2e-16***	0.3076
2 days	0.0016**	0.8675	0.516	0.3402	0.0341*	0.0759
3 days	0.4802	0.1282	0.1648	0.2693	0.2599	0.7981
4 days	0.2241	0.0151*	0.0219*	0.456	0.3794	0.2417
5 days	0.3903	0.0062**	0.0124*	0.6194	0.5182	0.0636
6 days	< 2e-16***	0.0489*	0.6203	0.0376*	0.119	0.7341
7 days	< 2e-16***	0.0397*	0.7672	0.0579	0.8404	0.1928
8 days	0.07946.	0.7096	0.9234	0.4738	0.6828	0.7515
9 days	0.2944	0.6014	0.5942	0.8665	0.4242	0.2889
10 days	0.6555	0.8086	0.9362	0.776	0.9341	0.4615
11 days	0.4867	0.8272	0.8427	0.2181	0.3939	0.4183
12 days	< 2e-16***	0.1338	0.4991	0.7899	0.0960	0.2412
13 days	< 2e-16***	0.9172	0.0675	0.9504	0.0890	0.5935
14 days	0.5729	0.1792	0.8931	0.3441	0.5279	0.3112

Table 6.4 Estimates of VAR model for Bitcoin return - 6 days cumulative

Time lag	6 days cumulative return in Bitcoin price	$\Delta$ in # of negative posts	$\Delta$ in # of positive posts	Daily return in S&P500 index	Daily return in VIX index	Daily return in Gold price
1 days	<b>33.67</b>	1.55	(1.27)	1.67	<b>4.22</b>	(1.02)
2 days	<b>3.17</b>	0.17	0.65	(0.95)	<b>(2.12)</b>	(1.78)
3 days	0.71	(1.52)	1.39	(1.11)	(1.13)	(0.26)
4 days	(1.22)	<b>(2.43)</b>	<b>2.3</b>	(0.75)	(0.88)	1.17
5 days	(0.86)	<b>(2.74)</b>	<b>2.5</b>	(0.5)	(0.65)	1.86
6 days	<b>(19.64)</b>	<b>1.97</b>	0.5	<b>2.08</b>	1.56	(0.34)
7 days	<b>15.49</b>	<b>2.06</b>	(0.3)	(1.9)	0.2	(1.3)
8 days	1.76	(0.37)	(0.1)	(0.72)	(0.41)	0.32
9 days	(1.05)	(0.52)	0.53	(0.17)	0.8	1.06
10 days	0.45	(0.24)	0.08	(0.29)	(0.08)	0.74
11 days	(0.7)	0.22	0.2	1.23	(0.85)	0.81
12 days	<b>(9.77)</b>	1.5	(0.68)	0.27	(1.67)	(1.17)
13 days	<b>9.26</b>	(0.1)	(1.83)	(0.06)	1.7	(0.53)
14 days	(0.56)	(1.34)	(0.13)	0.95	0.63	(1.01)

## 6.2 The result of weekly model

Table 6.5 Statistical significance (p-values) of VAR model for Bitcoin return - weekly

Time lag	Weekly return in Bitcoin price	$\Delta$ in # of negative posts	$\Delta$ in # of positive posts	Weekly return in S&P500 index	Weekly return in VIX index	Weekly return in Gold price
1 week	0.1019	0.7195	0.1734	0.7343	0.5386	0.226
2 weeks	0.7669	0.7475	0.656	0.9269	0.3803	0.3078
3 weeks	0.9916	0.6383	0.658	0.2784	0.0696	0.4595

We have derived weekly data using daily data modeled in Section 6.1. We aimed to understand whether the activity in the forum has a long-lasting impact that could not be captured in daily models. Table 6.5 shows the results from weekly predictions. With the optimal lag=3, none of the variables show statistical significance.

We have also applied a model with a cumulative return in 2 weeks. Although the p-values for negative posts show significance at 1% level, the coefficients are positive, indicating a positive impact when the number of negative posts increases.

Table 6.6 Statistical significance (p-values) of VAR model for Bitcoin return - 2 weeks cumulative

Time lag	2 weeks cumulative return in Bitcoin price	$\Delta$ in # of negative posts	$\Delta$ in # of positive posts	Weekly return in S&P500 index	Weekly return in VIX index	Weekly return in Gold price
1 week	< 2e-16 ***	0.4851	0.9617	0.8278	0.9015	0.8808
2 weeks	< 2e-16 ***	0.0013**	0.4714	0.9714	0.6587	0.1974
3 weeks	0.0010***	0.0088**	0.1482	0.4543	0.3303	0.5815



Table 6.7 Estimates of VAR model for Bitcoin return - 2 weeks cumulative

Time lag	2 weeks cumulative return in Bitcoin price	$\Delta$ in # of negative posts	$\Delta$ in # of positive posts	Weekly return in S&P500 index	Weekly return in VIX index	Weekly return in Gold price
1 week	<b>0.8274</b>	1.9e-05	0.0000	(0.1179)	0.0084	0.0712
2 weeks	<b>(0.4891)</b>	<b>8.8e-05</b>	(0.0000)	0.0196	0.0303	(0.6186)
3 weeks	<b>0.2445</b>	<b>6.9e-05</b>	(0.0000)	(0.4040)	(0.0664)	0.2662

Table 6.8 Statistical significance (p-values) of VAR model for Bitcoin return - 3 weeks cumulative

Time lag	3 weeks cumulative return in Bitcoin price	$\Delta$ in # of negative posts	$\Delta$ in # of positive posts	Weekly return in S&P500 index	Weekly return in VIX index	Weekly return in Gold price
1 week	<2e-16***	0.3675	0.9123	0.6069	0.8005	0.9433
2 weeks	0.5239	0.0499*	0.9899	0.3839	0.9711	0.2544
3 weeks	1.8e-07***	0.0273*	0.7584	0.0342*	0.0146*	0.8891
4 weeks	4.2e-07***	0.3071	0.6574	0.4731	0.4646	0.7516

Table 6.9 Statistical significance (p-values) of VAR model for Bitcoin return - 4 weeks cumulative

Time lag	4 weeks cumulative return in Bitcoin price	$\Delta$ in # of negative posts	$\Delta$ in # of positive posts	Weekly return in S&P500 index	Weekly return in VIX index	Weekly return in Gold price
1 week	< 2e-16 ***	0.383	0.8729	0.9286	0.7191	0.9043
2 weeks	0.4886	0.5415	0.0524	0.2365	0.3573	0.7153
3 weeks	0.0500	0.181	0.9178	0.455	0.132	0.9725

## 7. CONCLUSION & FUTURE WORK

The cryptocurrencies have many offerings such as breaking global financial barriers, lowering transactions cost, and enabling a faster way of peer-to-peer transactions. The technology behind cryptocurrencies, the blockchain promises a great contribution to the financial innovation ecosystem. Understanding the dynamics of cryptocurrencies will be more crucial than now.

In our study, we have sought to quantify the dynamic relationship between investors' sentiment and the value of Bitcoin. To the best of our knowledge, we provide the most comprehensive study to date with regards to the time frame, which contains second and third Bitcoin halving in 2016 and 2020. Since the pricing dynamic of Bitcoin is different than the fiat currencies, we were expecting that investors' sentiment provides valuable information about the trading of Bitcoin. Additionally, we thought that the relationship would be more obvious in a content-specific online forum. However, our data revealed no such relationship between investors' sentiment sourced from forum activity and Bitcoin return.

Our results raise questions about the characteristics of Bitcoin users and the quality of the information sourced from the online forum. Hernandez et al. (2014) found that Bitcoin followers are less likely to mention social relations and emotion-related words in their tweets and demonstrate less social connection. Mai et al. (2018) found that content created by silent majority provides more information than the one by vocal minority who creates most of the content in the online forum. If real investors who have a greater impact on the transaction volume and price, are the ones who show little social interaction, the information that we have captured from the forum may have no valuable implications.

Another consideration may be about the effects of real and fake news, speculative information on the Bitcoin prices. In a market with limited official sources, it is hard to distinguish real information from a fake one, which leads to increased noise in the data.

We may suggest several recommendations for future research. Concerning the

methodology, more dimensions than positive/negative polarity could be incorporated into the model. Those dimensions may relate to the characteristics of Bitcoin users and enable the researcher to size/to differentiate the value of the information created in online platforms. It would be interesting to test whether the type of social platform offers any difference in terms of sentiment. Lastly, the proposed method could be applied to other cryptocurrencies to understand whether there is any characteristic difference in the investor base of different cryptocurrencies.



## BIBLIOGRAPHY

- Abraham, J., Higdon, D., Nelson, J., & Ibarra, J. (2018). Cryptocurrency price prediction using tweet volumes and sentiment analysis. *SMU Data Science Review*, 1(3), 1.
- Agarwal, A. & Sabharwal, J. (2012). End-to-end sentiment analysis of twitter data. In *Proceedings of the Workshop on Information Extraction and Entity Analytics on Social Media Data*, (pp. 39–44).
- Akerlof, G. A. & Shiller, R. J. (2010). *Animal spirits: How human psychology drives the economy, and why it matters for global capitalism*. Princeton university press.
- Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of computational science*, 2(1), 1–8.
- Esuli, A. & Sebastiani, F. (2007). Sentiwordnet: a high-coverage lexical resource for opinion mining. *Evaluation*, 17(1), 26.
- Fraiberger, S. P. (2016). News sentiment and cross-country fluctuations. *Available at SSRN 2730429*.
- Garcia, D. & Schweitzer, F. (2015). Social signals and algorithmic trading of bitcoin. *Royal Society open science*, 2(9), 150288.
- Gilbert, C. & Eric, H. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Eighth International Conference on Weblogs and Social Media (ICWSM-14)*. Available at (20/04/16) <http://comp.social.gatech.edu/papers/icwsm14.vader.hutto.pdf>, volume 81, (pp.82).
- Hernandez, I., Bashir, M., Jeon, G., & Bohr, J. (2014). Are bitcoin users less sociable? an analysis of users' language and social connections on twitter. In *International Conference on Human-Computer Interaction*, (pp. 26–31). Springer.
- Hu, M. & Liu, B. (2004). Mining opinion features in customer reviews. In *AAAI*, volume 4, (pp. 755–760).
- Kaminski, J. (2014). Nowcasting the bitcoin market with twitter signals. *arXiv preprint arXiv:1406.7577*.
- Kearney, C. & Liu, S. (2014). Textual sentiment in finance: A survey of methods and models. *International Review of Financial Analysis*, 33, 171–185.
- Kim, Y. B., Kim, J. G., Kim, W., Im, J. H., Kim, T. H., Kang, S. J., & Kim, C. H. (2016). Predicting fluctuations in cryptocurrency transactions based on user comments and replies. *PloS one*, 11(8), e0161197.
- Kim, Y. B., Lee, J., Park, N., Choo, J., Kim, J.-H., & Kim, C. H. (2017). When bitcoin encounters information in an online forum: Using text mining to analyse user opinions and predict value fluctuation. *PloS one*, 12(5), e0177630.
- Kjærland, F., Khazal, A., Krogstad, E. A., Nordstrøm, F. B., & Oust, A. (2018). An analysis of bitcoin's price dynamics. *Journal of Risk and Financial Management*, 11(4), 63.
- Kristoufek, L. (2013). Bitcoin meets google trends and wikipedia: Quantifying the relationship between phenomena of the internet era. *Scientific reports*, 3(1), 1–7.
- Li, Q. & Shah, S. (2017). Learning stock market sentiment lexicon and sentiment-

- oriented word vector from stocktwits. In *Proceedings of the 21st Conference on Computational Natural Language Learning (CoNLL 2017)*, (pp. 301–310).
- Loughran, T. & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54(4), 1187–1230.
- Lütkepohl, H., Krätzig, M., & Phillips, P. C. (2004). *Applied time series econometrics*. Cambridge university press.
- Mai, F., Shan, Z., Bai, Q., Wang, X., & Chiang, R. H. (2018). How does social media impact bitcoin value? a test of the silent majority hypothesis. *Journal of Management Information Systems*, 35(1), 19–52.
- Matta, M., Lunesu, I., & Marchesi, M. (2015). Bitcoin spread prediction using social and web search media. In *UMAP workshops*, (pp. 1–10).
- Maynard, D., Bontcheva, K., & Rout, D. (2012). Challenges in developing opinion mining tools for social media. *Proceedings of the @ NLP can u tag# user-generated content*, 15–22.
- Moghaddam, S. & Popowich, F. (2010). Opinion polarity identification through adjectives. *arXiv preprint arXiv:1011.4623*.
- Pak, A. & Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. In *LREc*, volume 10, (pp. 1320–1326).
- Polasik, M., Piotrowska, A. I., Wisniewski, T. P., Kotkowski, R., & Lightfoot, G. (2015). Price fluctuations and the use of bitcoin: An empirical inquiry. *International Journal of Electronic Commerce*, 20(1), 9–49.
- Prajapati, P. (2020). Predictive analysis of bitcoin price considering social sentiments. *arXiv preprint arXiv:2001.10343*.
- Shapiro, A. H., Sudhof, M., & Wilson, D. (2020). Measuring news sentiment. Federal Reserve Bank of San Francisco.
- Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: journal of the Econometric Society*, 1–48.
- Smailović, J., Grčar, M., Žnidaršič, M., & Lavrač, N. (2012). Sentiment analysis on tweets in a financial domain. In *4th Jožef Stefan International Postgraduate School Students Conference*, volume 1, (pp. 169–175).
- Steinert, L. & Herff, C. (2018). Predicting altcoin returns using social media. *PloS one*, 13(12), e0208119.
- Stone, P. J., Dunphy, D. C., & Smith, M. S. (1966). The general inquirer: A computer approach to content analysis.
- Twedt, B. & Rees, L. (2012). Reading between the lines: An empirical examination of qualitative attributes of financial analysts’ reports. *Journal of Accounting and Public Policy*, 31(1), 1–21.
- Valencia, F., Gómez-Espinosa, A., & Valdés-Aguirre, B. (2019). Price movement prediction of cryptocurrencies using sentiment analysis and machine learning. *Entropy*, 21(6), 589.
- Wang, G., Wang, T., Wang, B., Sambasivan, D., Zhang, Z., Zheng, H., & Zhao, B. Y. (2014). Crowds on wall street: Extracting value from social investing platforms. *arXiv preprint arXiv:1406.1137*.
- Xie, P., Wu, J., & Wu, C. (2017). Social data predictive power comparison across information channels and user groups: evidence from the bitcoin market. *The Journal of Business Inquiry*, 17(1), 41–54.
- Zhang, X., Fuehres, H., & Gloor, P. A. (2011). Predicting stock market indicators through twitter “i hope it is not as bad as i fear”. *Procedia-Social and*

*Behavioral Sciences, 26, 55–62.*



## APPENDIX A

### Results of analysis with Hu Liu + Loughran & McDonald Lexicon

Table A.1 The distribution of posts per year by sentiment (as per LM+HL lexicon)

	Positive	Negative	Neutral
2016	56%	31%	14%
2017	56%	31%	13%
2018	55%	31%	13%
2019	54%	33%	13%
2020	50%	37%	13%
<b>Total</b>	<b>55%</b>	<b>31%</b>	<b>13%</b>

Table A.2 Number of posts by boards (as per LM+HL lexicon)

	Positive	Negative	Neutral	<b>Total</b>
Discussion	710,695	400,831	167,601	<b>1,279,127</b>
Speculation	229,626	159,241	67,717	<b>456,584</b>
Trading	249,147	114,770	53,295	<b>417,212</b>
Economics	203,467	105,622	45,853	<b>354,942</b>
Press	28,608	24,808	7,397	<b>60,813</b>
Legal	12,867	14,242	4,307	<b>31,416</b>
Meetups	2,117	289	158	<b>2,564</b>
<b>Total</b>	<b>1,436,527</b>	<b>819,803</b>	<b>346,328</b>	<b>2,602,658</b>

Table A.3 Statistical significance (p-values) of VAR model for Bitcoin return - daily

Time lag	Daily return in Bitcoin price	in # of negative posts	in # of positive posts	Daily return in S&P500 index	Daily return in VIX index	Daily return in Gold price
1 days	0.0075**	0.5134	0.4507	0.2649	0.0204*	0.6826
2 days	0.469	0.4663	0.2861	0.1401	0.1675	0.3228
3 days	0.1058	0.0744	0.3153	0.6172	0.5977	0.9842
4 days	0.5263	0.4777	0.7313	0.8195	0.0731	0.0396*
5 days	0.9386	0.2091	0.687	0.5544	0.5198	0.2572
6 days	0.048*	0.5483	0.8692	0.0033**	0.0007***	0.1739
7 days	0.5087	0.5545	0.308	0.6075	0.3915	0.0703
8 days	0.955	0.5875	0.7113	0.5747	0.873	0.2256
9 days	0.0578	0.7737	0.8424	0.8715	0.2346	0.2508
10 days	0.3686	0.1844	0.6476	0.312	0.6551	0.045*
11 days	0.5289	0.7215	0.0803	0.1016	0.154	0.7247
12 days	0.1381	0.7884	0.0784	0.8051	0.3881	0.0722
13 days	0.1209	0.5933	0.9621	0.417	0.5342	0.3679
14 days	0.3646	0.4166	0.8348	0.997	0.3762	0.4167

Table A.4 Statistical significance (p-values) of VAR model for Bitcoin return - daily

Time lag	Daily return in Bitcoin price	$\Delta$ in daily sentiment score	Daily return in S&P500 index	Daily return in VIX index	Daily return in Gold price
1 days	0.0064**	0.2567	0.2548	0.0308*	0.6524
2 days	0.4394	0.3017	0.0922	0.1212	0.315
3 days	0.1736	0.8227	0.7525	0.5266	0.9264
4 days	0.8871	0.409	0.8583	0.0774	0.0309*
5 days	0.7041	0.772	0.5108	0.5129	0.3632
6 days	0.0366*	0.6658	0.0052**	0.0016**	0.2022
7 days	0.4774	0.2157	0.7794	0.5405	0.0739
8 days	0.8428	0.8013	0.6933	0.7225	0.1837
9 days	0.0685	0.686	0.9516	0.2232	0.4347
10 days	0.2423	0.6982	0.2418	0.5915	0.0617
11 days	0.7917	0.1979	0.0715	0.0749	0.8042
12 days	0.1159	0.1937	0.7118	0.2651	0.0563
13 days	0.1183	0.5941	0.3688	0.7476	0.3415



Table A.5 Statistical significance (p-values) of VAR model for Bitcoin return - 6 days cumulative

Time lag	6 days cumulative return in Bitcoin price	$\Delta$ in # of negative posts	$\Delta$ in # of positive posts	Daily return in S&P500 index	Daily return in VIX index	Daily return in Gold price
1 days	< 2e-16***	0.2471	0.368	0.0884	<1.62e-5***	0.2648
2 days	0.0027**	0.712	0.5547	0.4111	0.0438	0.0694
3 days	0.5414	0.1922	0.2322	0.2903	0.2556*	0.8231
4 days	0.2785	0.0372*	0.0388*	0.4108	0.3295	0.2217
5 days	0.4544	0.0022**	0.0067**	0.5779	0.5165	0.0651
6 days	< 2e-16***	0.0657	0.5732	0.0352*	0.1075	0.6672
7 days	< 2e-16***	0.041*	0.8056	0.0648	0.7539	0.1733
8 days	0.0677	0.3497	0.8092	0.51	0.7462	0.7595
9 days	0.2644	0.6489	0.6342	0.8225	0.4683	0.259
10 days	0.5234	0.5192	0.6234	0.7055	0.8363	0.4534
11 days	0.413	0.903	0.5503	0.2315	0.3918	0.427
12 days	< 2e-16***	0.2804	0.8143	0.7801	0.1031	0.2651
13 days	< 2e-16***	0.8556	0.0489*	0.928	0.0925	0.6168
14 days	0.7095	0.1717	0.9114	0.3778	0.5435	0.331

Table A.6 Statistical significance (p-values) of VAR model for Bitcoin return - weekly

Time lag	Weekly return in Bitcoin price	$\Delta$ in # of negative posts	$\Delta$ in # of positive posts	Weekly return in S&P500 index	Weekly return in VIX index	Weekly return in Gold price
1 week	0.1146	0.6369	0.1456	0.7446	0.551	0.2176
2 weeks	0.7643	0.6207	0.8044	0.9021	0.3758	0.2996
3 weeks	0.9637	0.5394	0.7555	0.2748	0.0766	0.469

Table A.7 Statistical significance (p-values) of VAR model for Bitcoin return - 2 weeks cumulative

Time lag	2 weeks cumulative return in Bitcoin price	$\Delta$ in # of negative posts	$\Delta$ in # of positive posts	Weekly return in S&P500 index	Weekly return in VIX index	Weekly return in Gold price
1 week	< 2e-16***	0.4122	0.975	0.8513	0.8874	0.8796
2 weeks	1.84e-08***	0.0003***	0.2611	0.9634	0.6523	0.2117
3 weeks	0.0007***	0.0027**	0.0829	0.4594	0.3453	0.5784

Table A.8 Statistical significance (p-values) of VAR model for Bitcoin return - 3 weeks cumulative

Time lag	3 weeks cumulative return in Bitcoin price	$\Delta$ in # of negative posts	$\Delta$ in # of positive posts	Weekly return in S&P500 index	Weekly return in VIX index	Weekly return in Gold price
1 week	< 2e-16***	0.1213	0.6620	0.5423	0.7735	0.9827
2 weeks	0.5925	0.0123*	0.6271	0.4261	0.9690	0.2458
3 weeks	3.62e-07***	0.0546	0.9560	0.0284*	0.0168*	0.9044
4 weeks	7.57e-07***	0.3011	0.7106	0.4636	0.4921	0.7496

Table A.9 Statistical significance (p-values) of VAR model for Bitcoin return - 4 weeks cumulative

Time lag	4 weeks cumulative return in Bitcoin price	$\Delta$ in # of negative posts	$\Delta$ in # of positive posts	Weekly return in S&P500 index	Weekly return in VIX index	Weekly return in Gold price
1 week	< 2e-16***	0.1131	0.4391	0.8914	0.7217	0.8781
2 weeks	0.5576	0.4833	0.0622	0.2331	0.3665	0.6683
3 weeks	0.0557	0.1997	0.993	0.4011	0.1294	0.9867