Leopold-Franzens-Universität Innsbruck
Faculty of Business and Management

Department of Banking and Finance

Diplomarbeit

Zur Erreichung des akademischen Titels

Magister der Sozial- und Wirtschaftswissenschaften (Mag.rer.soc.oec.)

Does investor interest for Bitcoin impact the cryptocurrency markets?

An empirical investigation into Bitcoin, Ripple and Ether.

by

Philipp Steinlechner

(Matrikelnummer: 01415201)

Submission date: December 14, 2018

Supervisor: Mag. Alexander Kupfer, PhD

Preface

Since this thesis marks the end of my study at the University of Innsbruck, it is time to say thanks for the support during the last four years. I am very grateful for the support of my parents Elisabeth and Klaus, who helped me both financially and emotionally and were alongside my academic path at all times. At the same time, I want to thank my late grandfather Hans, who was a supporter since I was little, and his generosity had a huge impact in that I could enjoy my exchange year in New Orleans.

At the same time, I want to say thanks to my supervisor on this thesis, Alexander, who helped me in my academic struggles throughout 2018. Furthermore, I am grateful for the flexibility of my current employer Swarovski, who made it possible for me to facilitate both job and university.

At last, I am grateful for my friends, who helped in my journey from despair to the final completion of this thesis and I hope that it gives them strength that they can achieve the same. I am also looking forward to celebrating my graduation together with my sister Nina, who finishes her studies together with mine.

Table of Contents

1 Introduction 1

2 Theoretical Background 4

2.1 Cryptocurrencies 4

2.2 Bitcoin 14

2.3 Value of cryptocurrencies 16

2.4 Research Question 21

3 Data 23

3.1 Observation period & Location 23

3.2 Data sources 24

3.3 Descriptive Statistics 33

4 Methodology 45

4.1 Logarithmic series 45

4.2 Return Volatility 45

4.3 Stationarity 48

4.4 VAR 51

5 Results 55

5.1 Interest Variables Interdependencies 56

5.2 Bitcoin – Interest Variables 57

5.3 Ether – Interest Variables 61

5.4 Ripple – Interest Variables 62

5.5 Cryptocurrency Interdependencies 65

5.6 Robustness 68

6 Conclusion 70

Appendix A: Further Material 73

A.1 Bitcoin Trading Volume (references 2.1.3.2) 73

A.2 Google Trends graph (references 3.2.2.1) 74

A.3 Python script to calculate the daily volatility of prices 75

A.4 Python script to retrieve the number of Reddit posts 76

A.5 Overview of Correlation between cryptocurrencies 77

A.6 Descriptive Statistics for Logarithmic Series 78

A.7 VAR Output for Model 4-2 79

A.8 Impulse Response Functions with Wikipedia instead of Google 81
A.9 Impulse Response Functions for Ether Monetary Trading Volume 82
A.10 Impulse Response Functions for Ripple Monetary Trading Volume 83
A.11 Impulse Response Functions for Robustness Variables 84
7 References 86

List of Abbreviations

e.g.	for example
i.e.	in other words
US	United States
BTC	Bitcoin
DJI	Dow Jones Index
VIX	Volatility Index
CNY	Chinese Renminbi
JPY	Japanese Yen
EUR	Euro
API	Application Programming Interface
URL	Uniform Resource Locator
JSON	JavaScript Object Notation
Q	Quartal
VAR	Vector AutoRegression
ROI	Return on Investment
RSS	Residual Sum of Squares
vs.	versus

List of Figures

Figure 2-1 Centralized vs. Peer-to-Peer network (Bieg, 2018) 5

Figure 2-2 Example of a Bitcoin Transaction (Blockchain, 2018) 6

Figure 2-3 Evolution of Cryptocurrency Market Capitalization (logarithmic scale) (CoinMarketCap, 2018f) 9

Figure 2-4 Monthly Trading Volume of Bitcoin by Currency 01/2011 – 09/2018 (Bitcoinity.org, 2018b) 10

Figure 2-5 Monthly Trading Volume of Bitcoin by Currency 02/2017 – 09/2018 (based on data from Bitcoinity.org) 11

Figure 2-6 Distribution of Currencies 02/2017 – 09/2018 (based on data from Bitcoinity.org) . 11

Figure 2-7 Overview of Bitcoin Transaction Process (Crosby et al., 2016) 16

Figure 2-8 Development of Bitcoin price and Google searches (own creation) 19

Figure 3-1 Development of the Bitcoin Price (from bitcoinity.org) 23

Figure 3-2 Example of a GET-request to Twitter's API using Postman (own creation) 30

Figure 3-3 Example of API response 31

Figure 3-4 Bitcoin price evolution with Boxplot 34

Figure 3-5 Ether price evolution with Boxplot 35

Figure 3-6 Ripple price evolution with Boxplot 36

Figure 3-7 Comparison of price development (normalized scale) 37

Figure 3-8 Google search evolution with Boxplot 38

Figure 3-9 Tweet volume evolution with Boxplot 39

Figure 3-10 Reddit posts evolution with Boxplot 39

Figure 3-11 Comparison of interest proxies' development (normalized scale) 40

Figure 3-12 Google Search volume vs. Wikipedia page views 42

Figure 3-13 Number of tweets containing 'BTC' vs. 'EPL' 42

Figure 3-14 Reddit posts: r/TaylorSwift vs. r/Bitcoin 43

Figure 4-1 Comparison of 3-, 5- and 7-day Return Volatility 47

Figure 5-1 Example of Impulse Response Function 55

Figure 5-2 Impulse Response Functions for **Model 4-2** 56

Figure 5-3 Impulse Response Function for Bitcoin price changes (**Model 4-3**.A) 57

Figure 5-4 Impulse Response Functions for Bitcoin Trading Volume (BTC) changes (**Model 4-3**.B) 58

Figure 5-5 Impulse Response Functions for Bitcoin Trading Volume (USD) changes (**Model 4-3**.C)........58

Figure 5-6 LBTC_TRADE_BTC_D (blue) vs. LBTC_TRADE_USD_D (red)........59

Figure 5-7 Impulse Response Functions for Bitcoin Intra-day Volatility changes (**Model 4-3**.D) 59

Figure 5-8 Impulse Response Function for Bitcoin Return Volatility (**Model 4-3**.E)........60

Figure 5-9 Impulse Response Function for Ether Price changes (**Model 4-4**.A)........61

Figure 5-10 Impulse Response Functions for Ether Trading Volume (ETH) changes (**Model 4-4**.B)........61

Figure 5-11 Impulse Response Functions for Ether Return Volatility (**Model 4-4**.E) 62

Figure 5-12 Impulse Response Functions for Ripple price changes (**Model 4-5**.A) 63

Figure 5-13 Impulse Response Functions for Ripple Trading Volume changes (XRP) (**Model 4-5**.B)........63

Figure 5-14 Impulse Response Functions for Ripple Intra-Day Volatility changes (**Model 4-5**.D) 64

Figure 5-15 Impulse Response Functions for Ripple Return Volatility (**Model 4-5**.E) 64

Figure 5-16 Impulse Response Functions for Cryptocurrency Price changes (**Model 4-6**.A).....65

Figure 5-17 Impulse Response Functions for Cryptocurrency Trade Volume changes (**Model 4-6**.B)........66

Figure 5-18 Impulse Response Functions for Cryptocurrency Trade Volume (USD) changes (**Model 4-6**.C) 66

Figure 5-19 Impulse Response Functions for Cryptocurrency Intra-Day Volatility changes (**Model 4-6**.D) 67

Figure 5-20 Impulse Response Functions for Cryptocurrency Return Volatility (**Model 4-6**.E) 67

Figure 5-21 Impulse Response Functions for Trading Volume changes 68

List of Tables

Table 2-1 Overview of largest cryptocurrencies (CoinMarketCap, 2018f).................................... 8

Table 2-2 Overview of largest exchanges (CoinMarketCap, 2018e).. 9

Table 3-1 Overview of Variable Names... 33

Table 3-2 Descriptive Statistics of Bitcoin.. 34

Table 3-3 Descriptive Statistics of Ether.. 35

Table 3-4 Descriptive Statistics of Ripple... 36

Table 3-5 Descriptive Statistics for Interest Proxies.. 38

Table 3-6 Correlation between Interest Proxies and Crypto-Markets.. 41

Table 4-1 Descriptive Statistics for Return series... 46

Table 4-2 Descriptive Statistics for Return Volatilities.. 48

Table 4-3 Results of Stationarity Analysis... 50

List of Equations

Equation 3-1 Bitcoin Price 25

Equation 3-2 Sample Standard Deviation 25

Equation 3-3 Daily Trading Volume Bitcoin 27

Equation 3-4 Daily Monetary Trading Volume 27

Equation 3-5 Daily Trading Volume Ether 27

Equation 3-6 Daily Trading Volume Ripple 27

Equation 4-1 Logarithmic Transformation 45

Equation 4-2 Return Calculation 46

Equation 4-3 Return Volatility 47

Equation 4-4 Three-day Return Volatility 47

Equation 4-5 First Differences Calculation 50

1 Introduction

Bitcoin has found its way to mainstream media with its stunning price surge, making many people wonder about the actual value of the cryptocurrency. For stocks, the price of one share equals the value of a fracture of the corporation, which shares an investor bought. For bonds, the price equals the coupon that is paid for lending the money in addition to the amount of money lent. But with Bitcoin, there is no enterprise that generates revenue by selling products or services or a government, which is capable of repaying its debt, thus the question of the value of Bitcoin remains.

As Bitcoin is labelled as a crypto*currency*, it is reasonable to assume that concepts of value generation can be borrowed from other currencies. There, the value is determined by a central authority that regulates supply in a way that the price of the currency satisfies certain requirements, e.g. price stability. For Bitcoin, these dynamics are uniquely different as its concept is based on decentralization thus no central authority, which regulates the supply, exists. Instead the supply is determined by a publicly known generation algorithm, that defines the rate at which Bitcoins are created. A process, by which the supply is reduced or increased outside of this algorithm does not exist for Bitcoin, thus the supply is fixed.

This fixed inelastic supply side of Bitcoin, thus, shifts the focus to the demand side, i.e. the investors of Bitcoin. Investors, who decide to exchange their income for assets, in theory, follow a rational decision process, represented by the 'homo oeconomicus' – the economic human – who bases his decision on information, cost-benefit calculations, risk and expectations in a way that is subjectively optimal. For example, investors, who are deciding whether to buy Amazon stocks will only do so if they think that in the future the stock price of Amazon will increase based on certain expectations about the corporation itself, about the stock market or the US economy in general.

While investors have shown that their decisions are not always as rational as theoretically assumed and this process is also driven by other factors, such as emotions, the homo oeconomicus is still considered as the theoretical foundation for investor decisions.

For an investor to make a rational decision, there must therefore exist a theoretical foundation. However, for Bitcoin no concept exists that allows for the determination of a fundamental value. There are no guaranteed pay-offs and expectations about the future of the Bitcoin price are nothing more than sheer hopes.

This has drastic consequences for institutional investors and analysts: they leave the Bitcoin market as they do not have a theoretical foundation to justify their investment decisions. The remaining market participants are private investors and speculative traders. These investors' decisions are driven mostly by the attractiveness of the asset, like products with strong brand names and heavy advertising are more likely to be bought, because consumers are more aware of these products.

Researchers have shown that due to the abundance of different stocks and the resulting scarcity of time for information search on all stocks, those that grab the investors' attention experience high abnormal trading volume and returns. Simply put, of all opportunities available to investors, those opportunities that attract their attention are more likely to be considered and thus more likely to be chosen, while less attractive options are more likely to be ignored. This proves to be especially applicable to retail investors, who are less experienced and have less access to professional tools than institutional investors (Barber & Odean, 2007).

Since the Bitcoin market lacks institutional investors, and thus consists mostly of private investors, this idea has been applied on the Bitcoin economy: the more attractive Bitcoin becomes to investors, the more interested they are and the more they buy Bitcoin, which leads to increasing prices.

The question then is, how this attractiveness respectively the consequent interest of investors can be quantitively measured. Google search behavior has proven to be indicative of larger trends that are otherwise hard to observe: In 2009, researchers at Google Inc. together with the Centers for Disease Control and Prevention in the Atlanta, USA, have shown that by using Google Search data they can "accurately estimate the current level of weekly influenza activity in each region of the United States" (Ginsberg et al., p. 1), which can help reduce the impact of influenza (Ginsberg et al., 2009).

Triggered by this idea of Google search behavior revealing underlying trends, many studies were conducted by researchers who investigated the impact of search behavior on the stock market and found striking correlations between the search volume of stocks and their corresponding prices, their volatility and their trading volume.

Due to the confirmation of Google Search as a proxy for investor interest, other sources such as Wikipedia and news reports were investigated with similar outcomes.

Expanding the field from information search to information sharing, researchers explored the possibilities of Social Media as a means for information sharing. Twitter, with its fast-paced instantaneous characteristics, established itself as a valid indicator for a topic's relevance and with the later inclusion of sentimental analysis even as indicator of public mood.

At first, researchers focused mostly on the stock market and tried to find another way of determining stock prices aside from established concepts such as the Capital Asset Pricing Model. With the emergence of cryptocurrencies in the late 2000s, many analyses used the insights garnered from the stock market and tried to apply them on the cryptocurrency market. Bitcoin, as the best known and most used cryptocurrency, received most of the researchers' attention, which led to many studies on the connection between various characteristics of Bitcoin and the different proxies for investors' interest.

Most prominently, the link between Google search volume for the term Bitcoin and the Bitcoin price was studied, later broadening this approach by additionally investigating Bitcoin's volatility and trading volume. Other sources for investor interest included Wikipedia, Twitter, Facebook, news, surveys and forums.

Led by the introduction of many new cryptocurrencies, the researching community also started to investigate other cryptocurrencies. The first to reach public attention was Ripple in 2012, followed by Ether[1] in 2015, and both started to attract the attention of researchers.

While the existing literature was very thorough, it was also very specific and narrow: It either focused on Bitcoin with all its facets but neglected other cryptocurrencies. Or it investigated all three cryptocurrencies, but only analyzed one metric, like the price and neglected other characteristics such as the trading volume or the volatility. Furthermore, the dynamics between Bitcoin,

[1] Note that Ether is the actual cryptocurrency while Ethereum is the protocol used like Blockchain is for Bitcoin. However, Ether is often mistakenly referred to as Ethereum (Ethereum Foundation, 2018a). From here on, Ether will be used when referring to the actual cryptocurrency, while Ethereum will only be used if the protocol is meant.

Ether and Ripple have not been analyzed thoroughly yet. This is an interesting aspect of cryptocurrencies as the price development of Bitcoin may affect the price development of Ether and Ripple. If Bitcoin's price increases sharply, it is likely that other cryptocurrencies follow, because Bitcoin is perceived as a representative for all cryptocurrencies.

Thus, the proposition of this thesis is to:

- study the relationship between the Bitcoin price, its trading volume and its volatility and different proxies for investor interest, specifically
    - Google Search behavior,
    - Twitter and
    - Reddit,
- include Ether and Ripple and check if they react differently to investor interest,
- search for interdependencies among the interest proxies themselves and
- analyze the dynamics between Bitcoin, Ether and Ripple.

In the following chapter 2 Theoretical Background, a more in-depth analysis of the existing literature and the theoretical background for this thesis will be given followed by chapter 3 Data, where together with the chapter 4 Methodology, the empirical foundation will be explained. In the last two chapters Results (5) and Conclusion (6), the outcome of the empirical analysis will be analyzed and discussed.

2 Theoretical Background

This chapter gives an introduction into cryptocurrencies and the characteristics that make them so unique and disruptive. Then, the history of cryptocurrencies is reviewed and an overview of their market and regulations is given. Subsequently, Bitcoin and Blockchain will be explained in more detail. Then, the chapter concludes with the exploration into the value of cryptocurrencies.

2.1 Cryptocurrencies

In the Merriam-Webster dictionary, a cryptocurrency is defined as "*any form of currency that only exists digitally, that usually has no central issuing or regulating authority but instead uses a decentralized system to record transactions and manage the issuance of new units, and that relies on cryptography to prevent counterfeiting and fraudulent transactions*" (Webster, 2018). This extensive definition highlights different characteristics that are unique to cryptocurrencies:

1. It is a currency that only exists digitally, i.e. it does not exist in physical form like other currencies, which usually are available both in digital form (book money) and physical form (cash). While cash consists of notes and coins and is used for over-the-counter transactions, book money is a form of digital money in a way that it is transferred from one bank account to another directly. Thus, digital money disposes of the need to withdraw the amount on one account and depositing it in another. Cryptocurrencies do not have notes or coins, which means that the only way of transfer is digitally.
2. Cryptocurrencies generally do not have a central issuing or regulating authority unlike regular currencies where an institution, normally a central bank, manages a country's currency, money supply and interest rates.
3. To the contrary, a cryptocurrency uses a decentralized system. Such systems enable accountability and regulation of the money supply. More specifically, the supply of cryptocurrencies, i.e. the generation of new coins, is determined according to a generation algorithm, which is defined by computer code (Nakamoto, 2008).
   Among the different cryptocurrencies, the decentralized system varies: Bitcoin uses the Blockchain technology (Nakamoto, 2008), Ripple utilizes a similarly functioning system called XRP (Ripple, 2018) while Ether uses a modified version of Blockchain focusing on Smart Contracts (Ethereum Foundation, 2018b).
   Regular currencies use banks, which accomplish this facilitator role for transactions and are responsible for recording and storing them, constituting a centralized system.
4. To establish security, these decentralized systems use cryptography to prevent irregularities while at the same time ensuring that the parties involved remain anonymous (Nakamoto, 2008). As there is no central authority to oversee transactions, verify the transaction parties and ensure that the money is transferred correctly, this part needs to be accomplished by the system itself.
   If transactions were invalid, decentralized anonymous transactions became problematic because the system does not have access over the personal information of either party and thus cannot settle the dispute, i.e. these transactions are irreversible (Crosby, Pattanayak, Verma, & Kalyanaraman, 2016). Accordingly, the system must make sure that transactions are correct.

This is achieved partly via cryptography: a verification process invented to ensure that transactions are only executed if they are valid. This process, simplified, uses digital signatures and a peer-to-peer network to oversee transaction.
Other nodes within the network only give the green light to those transactions that are valid. All these nodes combined constitute the third party to a transaction like the bank is in centralized systems (Nakamoto, 2008).

2.1.1 Characteristics

Cryptocurrencies are inherent to certain characteristics, which earned them their disruptive and innovative reputation. Their decentralized architecture and anonymity are properties that challenge the current system, in which big corporations and institution have come under scrutiny (Harrington, 2017). In times where Facebook data was used to influence voters in elections (Cadwalladr & Graham-Harrison, 2018) and banks still recovered from their damaged reputation caused by the financial crisis in 2007, Bitcoin and Co. suddenly became an alternative to the existing system (Naughton, 2013).

2.1.1.1 Decentralization

The concept of decentralization of Bitcoin is unique and fundamentally different to centralized systems. With regular currencies, there exists a central authority that regulates the currency, as for example the European Central Bank (ECB) does for the Euro or the Federal Reserve (FED) does for the US Dollar. Furthermore, trusted third parties must keep proof of every transaction stored as otherwise no proof of the transaction could be validated in case of a dispute. These centralized entities pose risks not only in that they are more prone to attacks from illegitimate parties due to their concentrated nature (single point of failure), but also in their influenceability in terms of policy decisions from major players in the market (Böhme, Christin, Edelman, & Moore, 2015).

To better understand the difference between a centralized and a decentralized system, it is helpful to imagine the download of software: The software can be either downloaded directly from the software provider's server (centralized system) or it can be downloaded from different users, who have already downloaded it onto their personal computers (decentralized system). Then, these computers constitute servers, multiplying the server number by the number of users in the network.

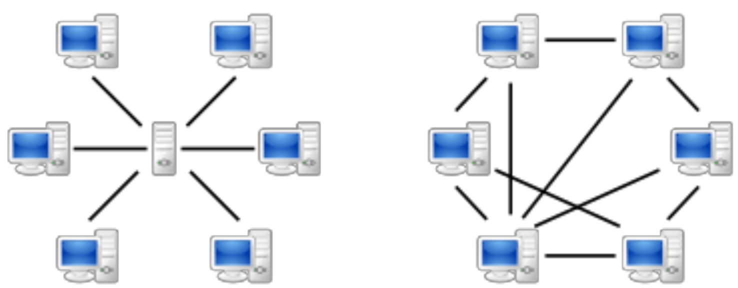


Figure 2-1 *Centralized vs. Peer-to-Peer network (Bieg, 2018)*

Such decentralized systems are also called peer-to-peer (P2P) networks, because they directly connect users (peers) within a network and utilize network effects (Schollmeier, 2001).

In **Figure 2-1**, the central server on the left side symbolizes the central bank, through which all transactions run. The computers on the right side are analogous to the users within the cryptocurrency network, who are interconnected and communicate directly without the need of a central party.

One major disadvantage of centralized systems is that they have a single point of failure, i.e. if, for whatever reason, the central party is unavailable, the entire system becomes unavailable as all connecting nodes are suddenly unconnected. If one node in the decentralized system becomes unavailable, the connection to that party breaks down, but the remaining parties can still communicate with each other. Thus, a decentralized system is more resilient to attacks and failure than centralized systems (Böhme et al., 2015; Crosby et al., 2016).

Another advantage of decentralized systems is that, by concept, they are more time- and cost-efficient. Central authorities and third parties require a pay-off for their services and add another step in the transaction process as there is a back and forth between buyer, seller and central authority until the transaction is approved (Crosby et al., 2016).

2.1.1.2 Anonymity

To understand why cryptocurrencies are anonymous, it is necessary to first take a closer look at how transaction in decentralized systems are handled. If someone wants to trade cryptocurrencies, the first step is to create a wallet. A wallet, analogous to the wallet that contains bills, coins and cards, stores a holder's crypto-coins and information necessary to trade. A wallet holds the following properties:

- The **wallet address** is the unique identifier of the wallet and contains between 26-35 characters. The address is used to route transactions from one wallet to another and replaces the names of wallet holders (Bitcoin.com, 2017). In other words, if money is sent from wallet A to wallet B, the transaction will only record wallet A's and wallet B's address:

Figure 2-2 *Example of a Bitcoin Transaction (Blockchain, 2018)*

- The **public and the private key** of the wallet are used to generate a digital signature between sender and receiver. This process establishes cryptographic proof and protects a transaction.

To create a wallet, the only information that is needed is an email address and a password. Thus, the wallet is not linked to the personal information of the wallet holder and the real identity of users is not revealed, i.e. they remain anonymous (Crosby et al., 2016).

The wallet addresses of transactions, however, are visible to everybody as transaction are publicly available (see **Figure 2-2**). Thus, crypto-transactions are also called pseudonymous as the real identity is not revealed but the wallet address of the holder is not secret (Bitcoin.com, 2017).

2.1.2 History

Decentralization and anonymity were also the main drivers in the initial conception of cryptocurrencies. In 1982, David Chaum envisioned a currency, which was anonymous and disposed of the need to transmit credit card information in electronic payments. Chaum's main concern was the traceability of electronic payments as transactions would pass through banks. Therefore, they had access to the personal information of the payee and the usage of the transaction This led him to create a software called eCash, which stored a user's money and enabled the user to anonymously spend the money at vendors who accepted eCash. Personal information was encrypted using blind signatures, which were only readable with the corresponding public key, thus constituting the first cryptographic currency (Chaum, 1983).

From its conception in 1982, it took 13 more years until its realization through Chaum's corporation Digicash in 1995. Despite success, the company went bankrupt in 1998 as eCash was not able to establish itself next to regular credit cards (Pitta, 1999).

In the same year, a concept called proof-of-work was first conceptualized for cryptocurrencies. Wei Dai, a computer engineer from the University of Washington, published a paper about "b-money", which was intended to be an anonymous distributed electronic cash system. b-money was followed by Nick Szabo's "bit gold". The difference between b-money, bit gold and Chaum's eCash was that now other participants of the network would solve cryptographic puzzles to verify and time-stamp new coins to avoid the problem of double-spending, i.e. copying and pasting already existing data (Chohan, 2017).

In a very simplified explanation, this concept of solving computational puzzles is called proof-of-work and established itself as key element in the underlying decentralized systems of cryptocurrencies, e.g. Blockchain for Bitcoin. Although bit gold was never implemented, it is considered as the precursor of Bitcoin (Moskov, 2018), which was introduced, ten years later, in 2008.

Satoshi Nakamoto conceptualized the first fully functional decentralized cryptocurrency, called Bitcoin (Nakamoto, 2008). Bitcoin used the groundwork of Chaum, Dai and Szabo but introduced SHA-256 hash functions as proof-of-work puzzles.

Bitcoin's decentralized features come from the underlying Blockchain technology: Through a distributed ledger technology, it allows transactions to be distributed across all users instead of storing them all in one central authority (Nakamoto, 2008).

In 2011, Litecoin was released, which was the first cryptocurrency to use a different hash function but was also based on Blockchain (McMillan, 2013). There were also successful attempts to use a different technology than Blockchain, most prominently Ripple, which utilizes a distributed ledger technology called XRP Ledger (Chase & MacBrough, 2018).

With the many new cryptocurrencies, which were conceptually closely related to Bitcoin, a term to refer to these alternative cryptocurrencies was created – altcoins. As of October 2018, more than 2000 different altcoins existed, with a combined market capitalization of over 200 billion US Dollar (CoinMarketCap, 2018d).

2.1.3 Market

2034 different cryptocurrencies equal an average of more than 200 new altcoins a year – or close to 19 per month – since the initial introduction of Bitcoin in 2008. However, of all these cryptocurrencies only a few are notable if typical market metrics such as market capitalization, market share and trading volume are considered. On October 5th, 2018, the following were the ten most notable based on market capitalization:

#	Icon	Name	Market Capitalization	Price	24h Trading Volume	Market Share
1		Bitcoin	$113,304,872,002	$6,555.98	$3,348,217,522	51.79%
2		Ether	$22,799,908,340	$223.25	$1,295,288,562	10.42%
3		Ripple	$20,654,637,985	$0.52	$487,118,933	9.44%
4		BTC Cash	$8,967,378,277	$516.46	$299,793,373	4.10%
5		EOS	$5,212,973,957	$5.75	$646,671,820	2.38%
6		Stellar	$4,576,198,899	$0.24	$36,511,173	2.09%
7		Litecoin	$3,421,861,020	$58.58	$321,192,429	1.56%
8		Tether	$2,785,958,399	$0.99	$1,956,421,402	1.27%
9		Cardano	$2,114,759,946.00	$0.08	$29,054,391	0.97%
10		Monero	$1,872,702,979	$113.99	$18,769,329	0.86%
Σ			$185,711,251,804			84.89%

Table 2-1 *Overview of largest cryptocurrencies (CoinMarketCap, 2018f)*

With an estimated total market capitalization of all 2034 cryptocurrencies of close to 219 billion US Dollar (October 5th, 2018), the ten largest by market capitalization, as depicted in Table 2-1, make up almost 85% of the overall market.

As one would expect, the most prominent cryptocurrency, Bitcoin, takes more than half of the total market with a market capitalization of about 113 billion US Dollar, distantly followed by Ether and Ripple, which both take approximately ten percent of the market. Ranked behind the major three is Bitcoin Cash, which is a modification of Bitcoin itself introduced in 2017. From there on downwards, no altcoin has a market share over two and a half percent, indicating that the four leaders own more than three quarters of the cryptocurrency market.

As seen in **Figure 2-3**, the market itself has evolved from approximately one billion US dollar in July 2013, to its highest point of close to one trillion US dollar at the end of 2017 and now settles at around 200 billion US dollar:

***Figure 2-3** Evolution of Cryptocurrency Market Capitalization (logarithmic scale) (CoinMarketCap, 2018f)*

2.1.3.1 Exchanges

Cryptocurrency exchanges act as market makers for the trading of cryptocurrencies and other assets, e.g. users may wish to trade US dollar for Bitcoin. Exchanges are an alternative to the execution of payments directly from one party to another, as it helps to connect a buyer and a seller. In return, they receive a transaction fee as commission, usually the bid-ask spread of the transaction like market makers on stock exchanges (Moore, 2018). Cryptocurrency exchanges are usually internet based and do not offer a physical locality. (Hansen, 2018)

Of all exchanges, Mt. Gox, a Japan based exchange, received particularly much attention, because it was the largest exchange before it shut down its operations in 2014 because of a large theft of Bitcoins (Wolf & Flitter, 2014).

#	Exchange	Country	Launch	Markets	24h Volume	30d Volume
1	Binance	Japan	07/2017	386	$1,149,051,911	$31,565,228,160
2	Bitfinex	Hong Kong	10/2012	86	$677,728,008	$13,055,581,296
3	OKEx	Samoa	01/2014	505	$648,787,065	$19,531,349,712
4	Huobi	Hong Kong	09/2013	281	$437,533,125	$16,227,085,974
5	ZB.COM	Singapore	11/2017	93	$454,113,569	$13,347,055,520
14	Kraken	USA	07/2011	66	$163,545,638	$3,269,633,076
25	Bitstamp	EU	07/2011	14	$94,650,129	$2,420,255,502

***Table 2-2** Overview of largest exchanges (CoinMarketCap, 2018e)*

As of October 2018, the market share among the different exchanges is similar to the market structure of the different cryptocurrencies: Binance, a Japan based exchange, is with 20% market share the major player in the exchange market. Other notable exchanges are Bitfinex (6%), Kraken (4%), Huobi (4%) and Bitstamp (1%) (CryptoCoinCharts, 2018).

2.1.3.2 Countries and currencies

Cryptocurrencies can be exchanged for regular currencies like US Dollar (USD), the Euro (EUR) or the Renminbi (CNY), but they are also exchanged for other cryptocurrencies. Thus, like for other currencies, there are many different exchange rates.

It is reasonable to assume that the exchange rates to national currencies are linked to the origin of the buyer. It is more likely that an investor from the United States will exchange US Dollar than the Euro. Thus, the trading volume of certain exchange rates give an insight into which countries trade cryptocurrencies the most, which is helpful because the identity of investors is unknown due to the anonymity (pseudonimity) of cryptocurrencies.

Bitcoin is the longest existing, most prominent and most traded cryptocurrency, and therefore gives the best insight into the evolution of cryptocurrency usage by country and currency. Furthermore, Ripple's and Ether's data availability does not match that of Bitcoin thus the following overview will be focused on Bitcoin.

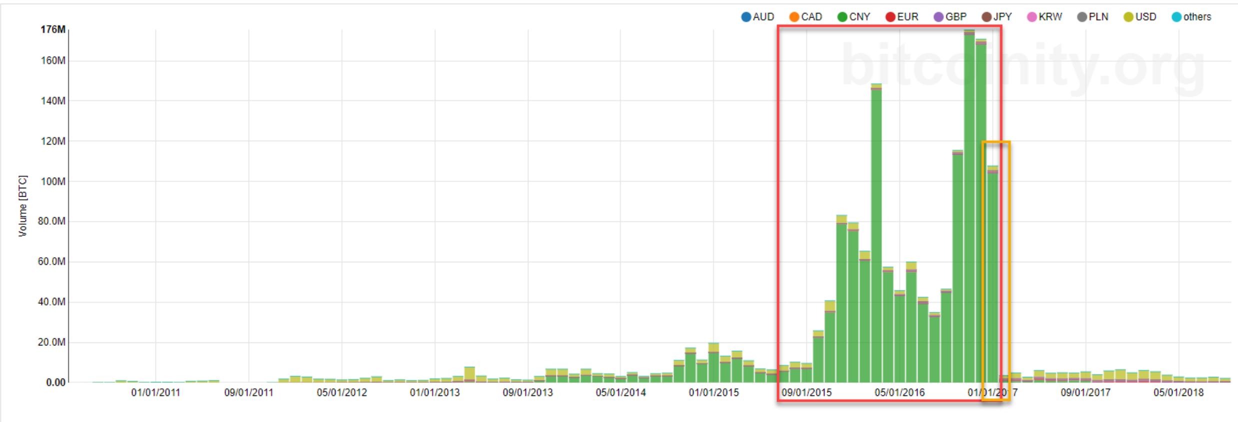


Figure 2-4 *Monthly Trading Volume of Bitcoin by Currency 01/2011 – 09/2018 (Bitcoinity.org, 2018b)*

As highlighted in red in **Figure 2-4**[2], the trading volume of Bitcoin has reached remarkable heights between 2015 and 2017 where Bitcoin was traded with a volume of up to 170 million Bitcoins a month. Furthermore, there is a sharp decline between January and February of 2017 (highlighted in orange). During January 2015 and January 2017, 94% of the trading volume came from the exchange of Bitcoin for the Chinese Renminbi (CNY) (Bitcoinity.org, 2018b).

Both phenomena are linked to the same source: The major exchanges in China, OkCoin, BTCC and Huobi did not charge fees until January 24th, 2017, which meant that until then, users were able to buy and sell Bitcoins free of charge. Thus, the incredible trading volumes were mostly

[2] For better readability of **Figure 2-4**, turn to Appendix (A.1), where the figure is depicted in larger dimensions.

generated by the Chinese market. After the exchanges started to charge fees for their services, the trading volume sank rapidly (Bovaird, 2017; Parker, 2017).

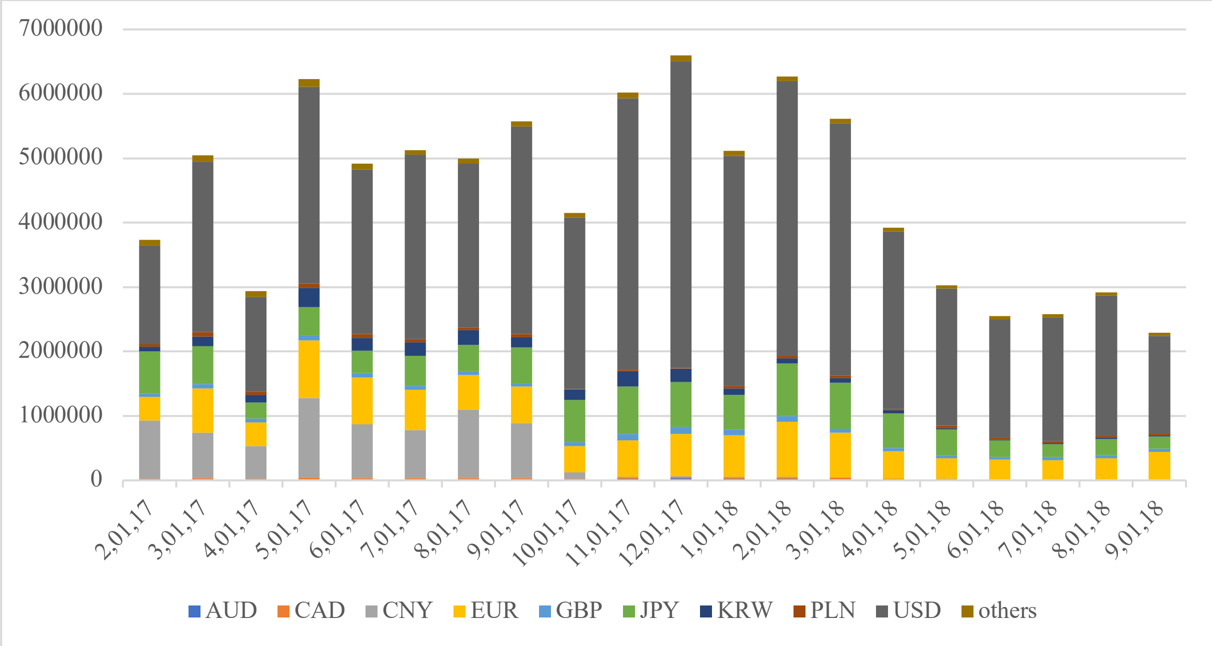


Figure 2-5 *Monthly Trading Volume of Bitcoin by Currency 02/2017 – 09/2018 (based on data from Bitcoinity.org)*

From February 2017 onwards, the distribution of currencies changed drastically, as seen in **Figure 2-5**.

The continued decrease in trading in CNY ended in November 2017, where the CNY numbers became negligible. This time, the reason for the stop of trading was China's central bank, which opened investigation into the major Chinese exchanges due to alleged market irregularities and money laundering among other issues in January 2017 and led to the closure of the major crypto-currency exchanges in September 2017 (Chaparro, 2017; Kelly, 2017).

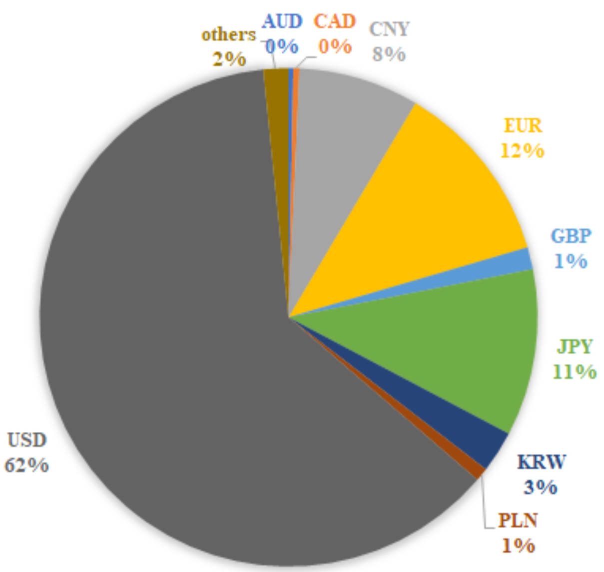


Figure 2-6 *Distribution of Currencies 02/2017 – 09/2018 (based on data from Bitcoinity.org)*

The distribution of currencies in this period saw a sharp increase in USD, which accounted for almost two third of trading volume. This steep increase was only relative as overall trading volume fell sharply from the peaks in December 2016 (~ 170 million BTC) to an average of 4.5 million BTC between 2/2017 and 9/2018 and the remaining trading was mostly done in USD.

The Euro accounted for twelve percent of trading volume, which was relatively stable over the period, while the Japanese Yen (JPY) noted a temporary increase after the exchange closures in China.

Chinese investors, who were not able to trade Bitcoin on Chinese Exchanges anymore started to trade on Japanese exchanges, where Bitcoin was previously recognized as a legal form of payment by the Japanese government.

The eight percent of trading volume between Bitcoin and the Chinese Renminbi are the average over the whole period, but if split into two periods at the regulatory restrictions in September 2017, CNY accounted for about 18 percent before and close to zero percent after September 2017.

2.1.3.3 Market participants and Usage

Due to the anonymity of users, statistics on users and the usage of cryptocurrencies are rare and researchers mostly use heuristics in their analyses. One approach, heavily dependent on the willingness and truthfulness of volunteers, is to ask Bitcoin users to answer a questionnaire. Bohr and Bashir (2014) used a publicly available dataset from 2013, which recorded close to 1200 participants.

While the authors warn about the validity of data, they found that anonymity and freedom are major drivers for investors. Additionally, the usage of Bitcoin for the acquisition of illegal goods correlated positively with the accumulation of Bitcoins among users.

Another approach is to link Google search data of terms representative for a certain clientele with the search volume for the term “Bitcoin”. Yelowitz and Wilson (2015) try to link the terms “Computer Science” (representative for computer programming enthusiasts), “Silk Road” (illegal activity), “Free Market” (Libertarians) and “Make Money” (Speculative Investors) to Bitcoin and found that only the first two are significantly positively linked, thus strengthening the finding of illegal activity as a driver for cryptocurrency users.

In 2018, Foley, Karlsen, and Putninš further investigate this dark side of Bitcoin. The researchers conclude that a significant amount of trading – around 25% of users and almost half of all transactions – are associated with illicit motives.

While the intended purpose of cryptocurrency as conceptualized by Nakamoto was as a payment instrument, the incredible price developments of cryptocurrencies has led many users to hold them as investment opportunities. Researchers found that more than 50 percent of all accounts that held Bitcoin were dormant, indicating that users held Bitcoins to earn returns from increasing prices (Ron & Shamir, 2013).

Thus, on the one hand, cryptocurrencies serve the purpose of a payment instrument, where anonymity and independence are the major drivers for users. On the other hand, cryptocurrencies are used as investments, where users expect to earn unusually high returns.

2.1.4 Regulation

The popularity of cryptocurrencies has called governments across the world to action. Aside from regulations to impede criminality, the taxation of income generated by cryptocurrencies is a key issue for governments.

For reference, the one trillion mark in market capitalization, which was almost hit by the overall cryptocurrency market at the end of 2017, is the same mark that was crossed by Apple and Amazon (Streitfeld, 2018), the biggest corporations worldwide, thus the need for regulation grew.

From an investor's perspective, governmental regulations influence the way cryptocurrencies are traded and could even put a stop to cryptocurrencies. Therefore, the introduction of policies and laws are watched closely by the respective markets. Just the announcement of potential changes impacts price volatility (Bovaird, 2018).

Aside from the economic impact of cryptocurrencies, the usage of cryptocurrencies for illegal transactions (as mentioned in 2.1.3.3) became an issue: Via Silk Road, an online black market mostly used to buy and sell drugs using an anonymous browser and Bitcoin payments, an estimated 15 million US dollar were exchanged in 2012 (Christin, 2013). Although it was closed down by the Federal Bureau of Investigation in 2013 (Greenberg, 2013), the problem was all but solved. Governments across the world are still in search of a way to deal with the novice that are crypto-currencies.

2.1.4.1 Asia & Russia

The Asian markets, especially China, Japan and South Korea, were major drivers for the initial success of cryptocurrencies, mainly Bitcoin (see 2.1.3.2). China became the major player in mining – the process of creating crypto coins – due to its comparative advantage in electricity cost (Huang, 2018). China also dominated the overall market as Chinese investors' investment opportunities are bound by national regulations and cryptocurrencies offered a new investment option for them (Popper, 2016).

As Chinese official were not pleased with the cryptocurrency phenomena and the affection of investors, the government came down hard on cryptocurrencies, especially Bitcoin. In 2014, the first attempt by the Chinese government to curb cryptocurrency trading was to order banks and payment providers to close accounts that were linked to operators of websites which enabled crypto trading. In addition to investigating major exchanges (see 2.1.3.2), Chinese regulators banned initial coin offerings (ICOs) in 2017. ICOs are a means for companies to raise money in return for a new cryptocurrency like an initial public offering (IPO) on stock markets.

In 2018, China continued their aggressive approach against cryptocurrencies and tried to impede the mining process by threatening miner to end their activities. Thus, many Chinese players decided to move their operations outside of China's jurisdiction to Canada, Switzerland, Japan and Malta where regulations are more favorable (Kharpal, 2018; Vaghela & Tan, 2018).

Japan recognized the opportunity handed by the Chinese government and decided to take a more lenient stance on cryptocurrencies: In 2017, Bitcoin was accepted as a legal form of payment and the Japanese government later also allowed exchanges to be registered and considered legal operators. However, several hacks, one of the biggest Japanese exchanges among them, called the Japanese government to action, but regulations on ICOs, speculative trading and security are still pending.

South Korea, on the other hand, has taken a firmer stance against cryptocurrencies warning about the illegal aspects of crypto trading. Mainly, the South Korean regulators brought in rules that take away part of the anonymity of users, giving away a crucial reason for using cryptocurrencies (Kharpal, 2018).

In Russia, the stance on cryptocurrencies has been shifting back and forth between banning cryptocurrencies altogether and regulating them. In December 2016, the Federal Tax Service released a statement where it vaguely implied that cryptocurrencies are not illegal (Prusakova, 2018).

2.1.4.2 United States, United Kingdom and the European Union

Western countries have taken a more lenient but cautionary approach. The regulators' intention is to warn users of the downsides and risks of using cryptocurrencies. They have, however, not yet come up with clear regulation for trading, mining or ICOs.

The United States' Securities and Exchange Commission (SEC) does not consider Bitcoin as securities (Pisani, 2018), which shifted the focus to other US regulators mainly the US government itself. The drive of institutional investors to get involved in cryptocurrencies has led to the introduction of futures, which follow the price of Bitcoin, but dispose of the need to invest in the currency itself. Thus, the vague regulatory environment of cryptocurrencies can to some extent be circumvented. Furthermore, there were attempts to introduce exchange-traded funds (ETFs) that follow the price of cryptocurrency similar to futures.

The United Kingdom has closely mirrored the actions of US regulators in the way that no regulations have passed yet, but several warnings have been issued by the Financial Conduct Authority (FCA) (Kharpal, 2018).

The view on cryptocurrency regulation of countries within the European Union differs from nation to nation, but the directive from the leaders of the European Union is that a consolidated harmonized approach should be taken. In April 2018, lawmakers have passed a policy that shall prevent money laundering, which is indirectly aimed at cryptocurrencies.

Overall, due to the novelty of cryptocurrencies and their disruptive implications for economies and even societies, regulatory laws and policies of most countries are still vague or non-existent with definitive directives still in contemplation.

2.2 Bitcoin

Bitcoin was the first cryptocurrency to emerge and even though many other altcoins have been introduced, Bitcoin has been mentioned in the same breath with cryptocurrencies. With its market share of over 50 percent (see 2.1.3) and peaks of more than eleven million transactions per day (Bitcoinity.org, 2018c), it is worth taking a closer look at the cryptocurrency that initiated the crypto craze.

Bitcoin was the first decentralized cryptocurrency, which was introduced by a to this day unknown person or consortium of persons by the name of Satoshi Nakamoto in 2008 and became operational in 2009. The cryptocurrency is based upon a peer-to-peer network, which enables sellers and buyers to connect directly without the need of a third party. By using the Blockchain technology, it forgoes the problem of trust in transactions that take place over the internet, i.e. Bitcoin enables trustless transactions. Up until then, such transactions were executed through electronic payments and were verified by trusted third parties such as financial institutions (Nakamoto, 2008).

Bitcoin uses the concept of a distributed ledger, which shares the record of transaction with all involved parties and thus eliminates the need of storing this ledger in one centralized place. Essentially, whenever a new transaction is recorded, it is first checked for correctness and mutual consent and, if approved, added to a chain of all previous transactions. Then, it is distributed to all network users resulting in not one but many points of storage of the ledger (Nakamoto, 2008).

2.2.1 Blockchain

When it comes to Bitcoin, it is important to understand the difference between Bitcoin and Blockchain. Blockchain is the technology behind Bitcoin, which facilitates the transactions between buyer and seller in a decentralized way. It earns its name from its architecture, which consists of multiple blocks that are linked to each other, forming a chain of blocks, i.e. a blockchain (Crosby et al., 2016).

Theoretically, Bitcoin and Blockchain could be separated entirely, meaning that Bitcoin could also be based on a different technology. For example, Bitcoin's transactions could be handled by a bank, which verifies transaction parties and approves the transfer. Then, Bitcoin would be based on a centralized network. At the same time, many other applications use the Blockchain technology as it is a means for the facilitation of any kind of transaction between two parties. In that sense, Bitcoin is one of many applications of the Blockchain technology.

Another application are contracts that are based on Blockchain. They remove the issue of trust between two parties, as the risk of nonperformance of one party can be mitigated. These so-called smart contracts guarantee the fulfillment programmatically and thus dispose of the need of a third party (Christidis & Devetsikiotis, 2016; Peters & Panayi, 2015).

For example, the code to access an apartment will only be released to the tenant if the tenant paid a certain amount of cryptocurrency to the renter. The tenant certainly receives the access code as soon as the rent is paid while the renter's access code is only released as soon as the payment is received. Without risk, the access code can be trusted upon the system before the record of payment, as it will only be stored but only released if the condition is fulfilled. Furthermore, the contracts are safe-guarded by the peer-to-peer network, thus modifications of the conditions by either party is impossible without notice and the consequent cancellation of the transaction.

Other applications of Blockchain include royalty payment and copyright record-keeping in the music industry, IPOs for private companies, the insurance of diamonds, document certification for notary purposes and the secure storage of health records (Crosby et al., 2016).

2.2.2 Trading

What is disruptive and innovative about Bitcoin is the Blockchain technology, because it gives the cryptocurrency the remarkable characteristics as discussed in 2.1.1. Thus, Bitcoin has become so popular because it uses the Blockchain technology and it exploits this unique technology: Blockchain validates, safeguards and preserves Bitcoin transactions (Crosby et al., 2016).

Before users can trade Bitcoin, they must first create a wallet (see 2.1.1.2) and then choose of one of the exchanges (see 2.1.3.1), which trades Bitcoin against the relevant currency. Then, the user must choose the amount of Bitcoin so that the exchange can find the respective trading partner/s, after which the transaction will be pending. Alternatively, the user can also transfer cryptocurrencies directly to a specific trading partner (through the wallet address), without the need for an exchange. At that moment, the transaction process executed through Blockchain is triggered:

Validation is achieved by broadcasting the transaction to the peer-to-peer network, which then approves the transaction if it is valid (step 3 and 4 in **Figure 2-7**). For these steps the public and private keys, as mentioned in 2.1.1.2, become necessary as they provide the digital signature needed to verify the transaction.

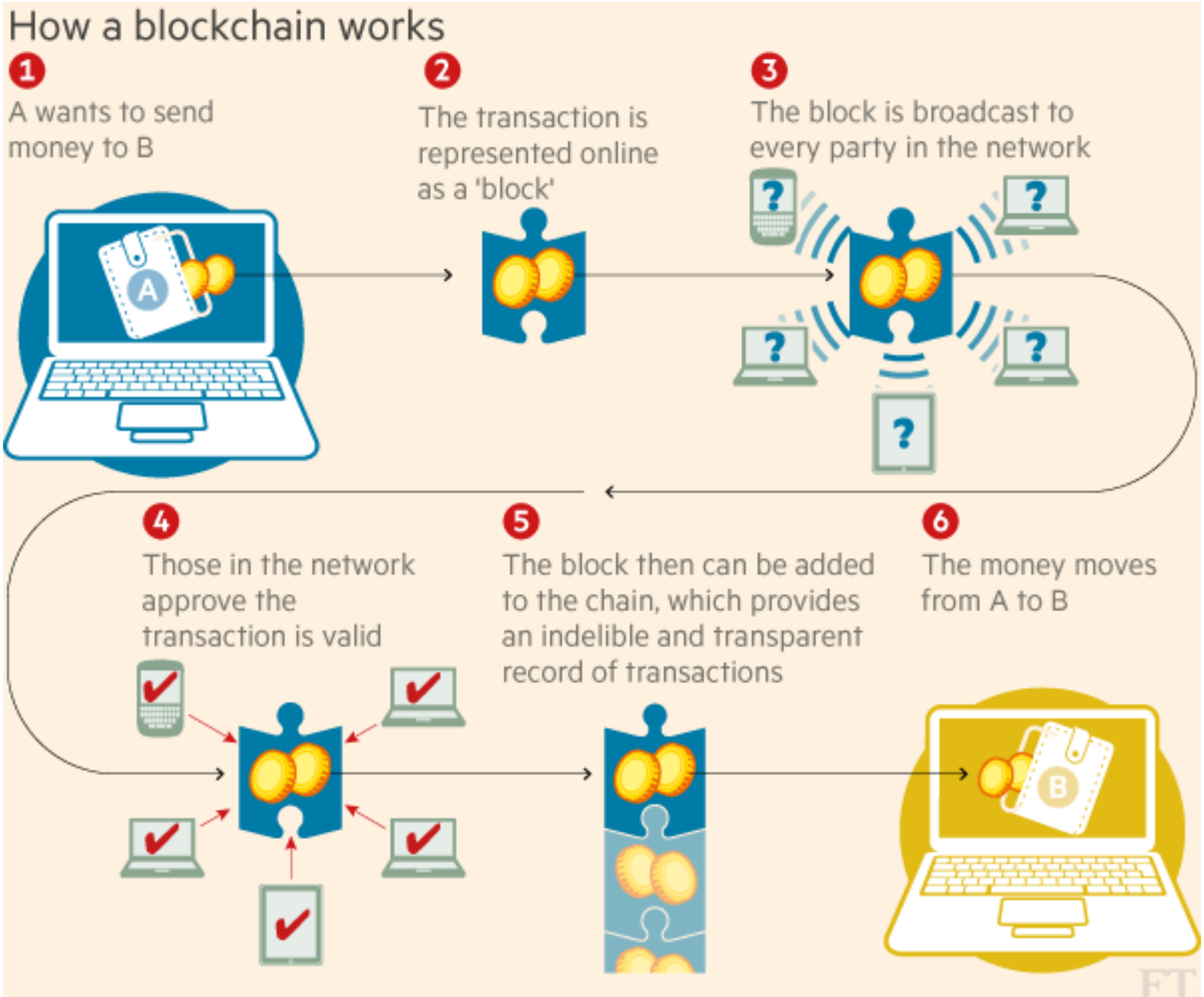


***Figure 2-7** Overview of Bitcoin Transaction Process (Crosby et al., 2016)*

If successful, the transaction is added into the unalterable public ledger, that is, the blockchain that thus guarantees security and the preservation of the transaction. The updated ledger is then distributed across all users, constituting a transparent and consistent system (Crosby et al., 2016).

2.3 Value of cryptocurrencies

Cryptocurrencies were conceptualized as a decentralized anonymous payment system, which enables the transfer of money without the need of an intermediary (Nakamoto, 2008). Nevertheless, when researchers investigated the Bitcoin transaction graph to answer questions about the behavior of users, they found that between 55% and 73% of Bitcoins were kept in dormant accounts, i.e. accounts that were used to acquire Bitcoin but with no subsequent outgoing transaction (Ron & Shamir, 2013). This discovery led Cheah and Fry (2015) to postulate that Bitcoin is more used like an a speculative asset than a currency.

Building on this hypothesis, the researchers investigated the fundamental value of Bitcoin and concluded that the value inherent to the cryptocurrency is zero. Additionally, Cheah and Fry attribute the price formation of Bitcoin to speculative bubbles. Such bubbles can either be irrational, i.e. when investors are driven by psychological factors which are not related to the asset's value or rational, i.e. when investors expect the sale of an overvalued asset for an even higher price. According to the researchers, both scenarios are reasonable explanation for the formation of bubbles in the Bitcoin economy.

The ensuing question thus became, what drivers lead users, from here on investors, to acquire Bitcoin even though it does not have a fundamental value.

2.3.1 Macro-Financial Factors

One hypothesis was that cryptocurrencies are linked to macro-financial factors, such as the gold price, the price level, the oil price and the stock market, arguing that these economic indicators influence the investors' buying and selling decisions.

Ciaian, Rajcaniova, and Kancs (2016) have investigated whether the impact of changes in the exchange rate of the US Dollar and the Euro, the oil price and the Dow Jones Index (DJI) exert an influence on the Bitcoin price but concluded that while there might be significant short-term implications for the Bitcoin price, the long-term effects of these indicators do not have a significant impact on Bitcoin.

Kristoufek (2015) analyzed whether Bitcoin is considered by investors as a safe haven in times of uncertainty and financial stress. The argument was that if Bitcoin truly is a safe haven to which investors turn in distressed times, then the Bitcoin price would be correlated with other established safe havens such as gold or the Swiss Franc. In their analysis they could not find such interconnections, leading them to believe that Bitcoin cannot be considered a safe haven. Thus, both investigations indicate that investors are not influenced by macro-financial variables in their decision to buy or sell Bitcoin.

2.3.2 Interest-driven Value

Another hypothesis proved more successful: investors' buying and selling decisions are driven by their interest, which is triggered by the attractiveness of assets. Thus, the more attractive cryptocurrencies are to investors, the more interested they are and the more likely they are to buy cryptocurrencies. In this cycle, the value is driven by interest.

Researchers have shown that due to the abundance of different stocks and the resulting scarcity of time for information search on all stocks, those that grab the investors' attention experience high abnormal trading volume and returns. Simply put, of all opportunities available to investors, those opportunities that attract their attention are more likely to be considered and thus more likely to be chosen, while less attractive options are more likely to be ignored (Barber & Odean, 2007).

Barber and Odean highlight the separation of the market in retail investors and institutional investors in the stock market. Institutional investors are likely to have access to professional tools and have generally more resources for information search unlike retail investors, whose resources are limited. The researchers found that retail investors are even more prone to making decisions based on attractiveness than their institutional counterparts (Barber & Odean, 2007).

Furthermore, Barber and Odean argue that these dynamics affect the acquisition of stocks more heavily than the sale of stocks: retail investors are more likely to sell the stocks that they already own, i.e. the search for information on which stock to sell is not relevant.

The findings of Barber and Odean (2007) and the implications of Kristoufek (2015) and Ciaian et al. (2016) that macro-financial variables are not useful in explaining trading activities, opened the gates for researchers who analyzed the connection between various indicators of investors' interest and financial assets.

Generally, the proxies are differentiated in their purpose to investors:

a) Investors search the internet to acquire information about assets.
b) Investors have already gathered information and decide to share it over the internet.

2.3.2.1 Information search

Outside of finance, Google search volume has proven a reliable source to detect patterns in users' search behavior. A famous example is the emergence of influenza epidemics in the United States (Ginsberg et al., 2009). Through Google search queries, the researchers were able to correctly estimate the level of influenza activity in different regions in the United States in a timelier fashion than the traditional surveillance systems, who typically rely on one to two weeks delayed data.

The idea that the search behavior of users has predictive power was taken up by the finance sector, starting with the stock market. Joseph, Wintoki, and Zhang (2011) use the search intensity for the tickers of all S&P 500 stocks and regress them against excess returns and abnormal trading volume of the stocks. They find a positive relationship that is even more pronounced if the stock is hard to arbitrage, as arbitrageurs find it harder to push the price to a fair price.

In 2013, Preis, Moat, and Stanley analyzed the DJI over a period of seven years and used 98 search terms related to trading. They then built a portfolio based on increases and decreases of the previous week's search volume of these terms and compared this portfolio to a random strategy portfolio. The portfolio based on search volume for specific trading terms achieved significantly higher returns than the random portfolio, signaling positive support for the search volume approach.

Similar implications were drawn from the analysis of Bijl, Kringhaug, Molnár, and Sandvik (2016), where they analyzed the broader S&P 500 index and used lagged instead of contemporaneous search volumes. Additionally, the researchers concluded that search volume better predicts the directions of excess returns rather than their magnitude.

Kristoufek (2013b), went one step further and used Google search volume to try and optimize portfolio diversification, arguing that stocks with higher search intensity are riskier and should thus be weighted less in the portfolio. This search-intensity weighted portfolio dominated both the benchmark index and a uniformly weighted portfolio over the entire observation period.

As most analyses were focused on the US stock market, Takeda and Wakao (2014) investigated whether the results would be the same on other stock markets. The results for 189 stocks from the Japanese Nikkei 225 index were in accordance with the results from the US stock market, confirming the search intensity approach outside the United States. Additionally, they found that the small firms are impacted more strongly by changes in search intensity than large firms and that trading volume tends to be greater during volatile times.

Moat et al. (2013) decided to take a closer look at Wikipedia data and distinguished between page views and page edits. They followed the approach of Preis et al. (2013) and build a portfolio based on whether page views/ edits were above or below the previous week's numbers and compared the portfolio to a random portfolio. They found that only page views lead to portfolio returns that are significantly different than the random portfolio's returns. Page edits do not generate significantly different returns than the random strategy. They do, however, state that the insignificance of page edits may be attributed to the scarcity of data for page edits as pages are edited notably less than they are viewed.

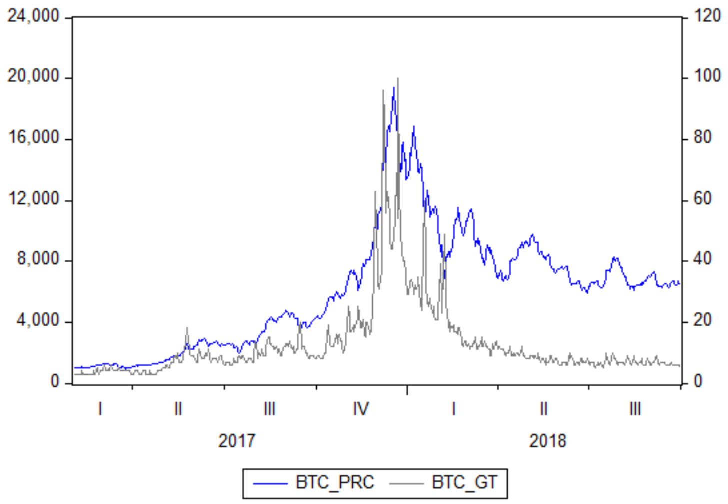


Figure 2-8 *Development of Bitcoin price and Google searches (own creation)*

Based on the findings from the stock market, Kristoufek (2013a) and Garcia, Tessone, Mavrodiev, and Perony (2014) borrowed these ideas and applied them to the Bitcoin market. They used Google search volume and, for robustness, include Wikipedia page views, arguing that Wikipedia serves as another source for investors to inform and help them in their buying decisions.

With impulse-response functions, Kristoufek found impulses from Bitcoin prices to Search queries and from Search queries to Bitcoin prices. Furthermore, the analysis showed that Bitcoin price movements are likely to be persistent as increasing prices cause increased interest, which in turn again drives prices up – the same spiraling effect is found for downwards tendencies.

2.3.2.2 Information sharing

In 2011, Mao, Counts, and Bollen (2011) investigated more proxies for investors‘ interest. They compared survey data, news from financial outlets and Social Media data to Google search volume in the pursue to explain movements of the DJI, Volatility Index (VIX) and gold prices. More specifically, they analyzed correlations and Granger causality dynamics between the different variables and found that Google search volume leads the DJI. The VIX, on the other hand, does not lead gold prices.

The researchers argued that search volume may be particularly useful for prediction in times of high volatility. Twitter data also proved to be a viable proxy for investors' interest as it holds predictive power over DJI returns while news data significantly correlates with the VIX.

Garcia et al. (2014) apply the information sharing component to Bitcoin and add Tweets from Twitter and Facebook post re-shares as proxies for information sharing in addition to search volumes. The researchers found that search volume increases with price, information sharing (Twitter and Facebook) increases with search volume and price again increases with positive changes in information sharing, completing a social cycle within the Bitcoin economy.

In 2015, Garcia and Schweitzer refined their 2014 investigations and analyzed Tweets for opinion polarization and emotional valence using sentimental analysis. According to the authors, social polarization is a precursor of political and economic phenomena. Additionally, they expanded the subject field and included trading and transaction volume of Bitcoin.

In their vector autoregressive analysis, they found that opinion polarization and trading volume have a positive impact on returns, especially on the first day after shocks. The trend, then, decreases rapidly thereafter. Emotional valence, which aims at identifying positive or negative sentiment towards Bitcoin, also has a significant effect on the trading volume of Bitcoin.

In the same year, Matta, Lunesu, and Marchesi (2015a, 2015b) contribute to the groundwork of Kristoufek and Garcia et al. with another investigation into Google search volume and Twitter data using sentiment analysis. The results of both studies were in accordance with previous findings.

2.3.2.3 Forums

Forums are used for both information search and information sharing. The creation of posts and comments serves the sharing of information while reading posts and comments serves the information search purpose (Bickart & Schindler, 2001; Osatuyi, 2013).

Kim, Y. Bin et al. (2015) used forums to develop a prediction system for value fluctuations within virtual worlds. They used comment and view counts of virtual communities of multiplayer online games to try and predict the fluctuations of the respective virtual currency using sentiment analysis. After success in the virtual gaming world, they later apply that prediction system on Bitcoin fluctuation and could correctly predict fluctuation with an average precision of up to 76% (Table 12, p.14).

Ciaian et al. (2016) based their model on Wikipedia views, the number of new members and posts on the popular Bitcoin forum Bitcointalk.org. In the short run, they found that of the three attractiveness proxies only new forum posts exert a continued significant impact on the Bitcoin economy over the entire observation period.

The impact of Wikipedia and the number of new members decreases over time leading the researchers to believe that Wikipedia may only be helpful in the acquisition of initial general information of Bitcoin. It does not, however, serve as a source for information on buying and selling decisions thereafter.

New forum posts, on the other hand, may give better insight into the current Bitcoin economy and help investors in their decisions after having gathered general information on Wikipedia.

2.3.3 Value of Altcoins

Ether and Ripple were the first followers of Bitcoin that started to attract researchers. Some techniques that proved successful for Bitcoin were also applied on Ether and Ripple.

Kim, Y. Bin et al. (2016) followed the approach of Ciaian et al. and crawled the respective forum of Bitcoin, Ether and Ripple for comments and replies, analyzed their texts for sentiment and tried to link them to the cryptocurrencies' prices and transaction volumes. The authors found significant correlations across all cryptocurrencies and their forum data confirming the results of Ciaian et al. (2016).

In 2016, researcher from the University of Korea took a closer look at the linkage between user comments and their replies and fluctuations of the three major cryptocurrencies and were able to predict fluctuations with an accuracy of up to 70%.

In 2018, Sovbetov analyzed five cryptocurrencies, including Bitcoin and Ether. He tested several drivers for the price: a crypto market beta, trading volume, volatility, S&P 500 and Google search queries. The results suggest that attractiveness of cryptocurrencies, as proxied by Google search, does not have significant impact in the short-run, but the coefficient become significant in the long-run. Sovbetov hypothesizes that this is because attractiveness builds over time and recognition builds slowly. Furthermore, the author finds that macro-financial variables are statistically insignificant in explaining cryptocurrency prices, similar to what Kristoufek (2015) and Ciaian et al. (2016) found in their Bitcoin focused examinations.

The results from studies on altcoins suggest that the finding from the Bitcoin analyses are also applicable to other cryptocurrencies.

2.4 Research Question

To summarize, this chapter has shown that cryptocurrencies exhibit unique characteristics (anonymity and decentralization), which were the key drivers in the establishment of cryptocurrencies and are still the main reasons for users to trade with crypto-coins.

Furthermore, the market is dominated by Bitcoin, the best-known cryptocurrency, followed by Ether and Ripple, who together own 75% of the market and gathered a market capitalization of more than 150 billion US Dollar as of October 2018.

In terms of trading, Chinese investors were the biggest force for the initial success of cryptocurrencies, but regulation curbed the Chinese hype and shifted the trading activity to countries with more lenient jurisdictions, namely the USA, Europe and Japan. Outside of China, regulators found it difficult to find proper legislation on the novice that are cryptocurrencies and have yet to come up with definitive answer on regulations.

An analysis into the usage of cryptocurrencies revealed two key insights: firstly, many users utilize the anonymous aspect of cryptocurrencies to engage in illegal trading activity and secondly, many users do not use them as a payment but as an investment instrument. Conclusively, this suggests that Bitcoin should be considered as an investment opportunity for speculative investors rather than a payment method.

A more in-depth look at Bitcoin offered additional information on the Blockchain technology, how this technology facilitates transaction without third parties and how Bitcoin is ultimately traded.

Finally, the question of value of cryptocurrencies was raised. Price formation relies on the basic interplay of supply and demand as the price builds the intersection between the two curves. Since the supply side is based on a known algorithm that is largely dependent on the energy consumption needed for the production of Bitcoin, the answer lies in the demand side, i.e. the investors that decide to buy Bitcoin. In a quest to gain a better understanding of the demand-side dynamics of cryptocurrencies, researchers have investigated different source that could influence the demand curve.

Different researchers have shown that the demand side of Bitcoin is not influenced by macro-financial variables, but by the interest that investors show for cryptocurrencies – an idea that has proven successful on the stock market. To proxy this interest, different sources have been considered: Google search queries constitute a viable resource for information search while Twitter and Forums are platforms that enable investors to share information. At last, research showed that altcoins exhibit similar dynamics as Bitcoin and are similarly influenced by proxies for investors' interest.

However, there has not been a thorough investigation into the three major cryptocurrencies and the linkage to the most important information-search and information-sharing proxies for investor interest. Most proxies have been linked to Bitcoin, but not Ether and Ripple while the dynamics between the three cryptocurrencies has found little attention. It is reasonable to hypothesize that the developments of Bitcoin have significant impact on the development of other cryptocurrencies as Bitcoin is perceived as the forerunner and in that way represents the overall crypto-market.

Accordingly, it is the aim of this thesis to conduct an extensive and inclusive analysis into Bitcoin, Ether and Ripple in respect to established sources for the approximation of investor interest, i.e. Google search, Twitter and Reddit. Another important aspect of the analysis will be what is widely known as the chicken or the egg conundrum: Are crypto-prices moved by an increase in interest or does the interest increase, because prices have risen?

The implication of this question is an endogeneity problem, which complicates standard analysis methods like a simple regression model, because it is dependent on a definitive answer on explanatory and explaining variables. A vector autoregressive (VAR) model will be applied, which assumes all variables endogen and shows how each variable reacts to a shock of another variable visualized in impulse-response functions.

Accordingly, the research questions of this thesis are:

1. Does the interest of investors for Bitcoin impact the Bitcoin economy?
2. Does that interest also impact other cryptocurrencies?
3. Are Ether and Ripple impacted by the development of Bitcoin?

These questions then lead to the following hypotheses:

1. H_0: Investor interest impacts Bitcoin's metrics.
2. H_0: Investor interest impacts Ripple's and Ether's metrics.
3. H_0: Bitcoin's metrics impact Ether's and Ripple's metrics.

To answer the research question and test its hypotheses, the chapter Methodology will explain what methods will be used for the empirical analysis. Additionally, the next chapter will offer more insight into the data that is needed for the later analysis.

3 Data

In the previous chapter, the pertinent research for the empirical analysis has been summarized. Furthermore, the research questions were raised, and its corresponding hypotheses were formulated. Now, the relevant variables for the analysis will be defined, the data for the variables gathered and described. At first, the observation period will be defined and the question whether all variables satisfy the global representation requirement is answered. Later, the data retrieval process is explained followed by an initial descriptive statistics analysis of the data set.

3.1 Observation period & Location

The decision on which period to analyze in the empirical analysis must be consistent with the theoretical background and the availability of data for all variables. From the theoretical perspective, the most interesting period starts in the first quarter of 2017. As mentioned in 2.1.3.2, the introduction of fees on major Chinese exchanges and the investigations into exchanges in China led to sharp decreases in trading volume. Thus, the trading volumes after the event are more representative of the real trading activity as fake trades are not included in the data after January 2017 (Parker, 2017).

In accordance with trading volume, the steep increase in Bitcoin price occurs in the first months of 2017 and thus defines the most interesting period to investigate:

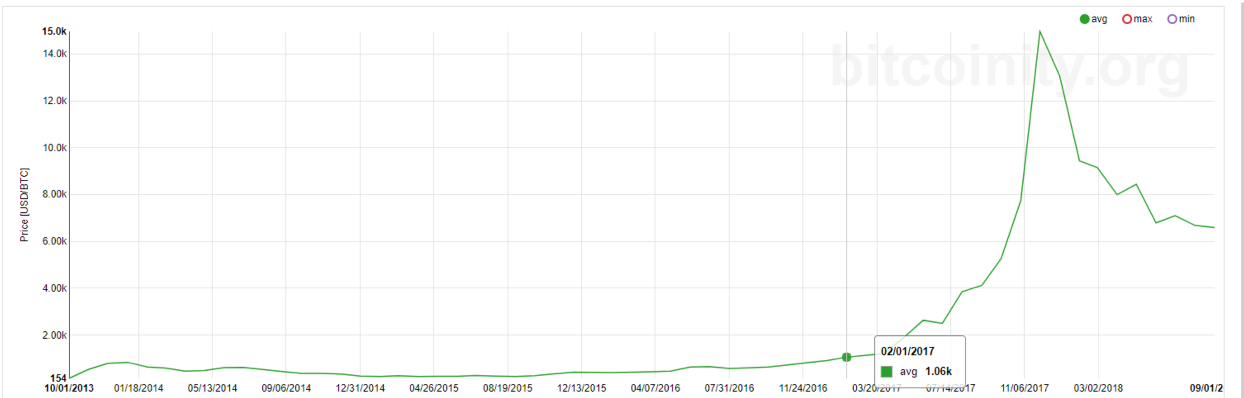


Figure 3-1 *Development of the Bitcoin Price (from bitcoinity.org)*

Furthermore, the data retrieval for the empirical analysis was done in October 2018, putting a cap on the end of the observation period in September 2018. Thus, the observation period will be between February 1st, 2017 and September 30th, 2018, totaling in 20 months or 607 days.

Secondly, the decision of the location for the analysis is decisive as it is connected to the currency, the language and the exchanges. Prices and their volatility will be retrieved in USD but are representative globally, because prices are harmonized across currencies (see the subsequent analysis in 3.2.1.1). The trading volume, as defined later in 3.2.1.3, consists of the number of Bitcoins traded, which are unrestricted from national boundaries. To estimate the monetary trading volume, this number is then multiplied with the average price on the major exchanges. Thus, both trading volumes are not bound to any specific country and represent the worldwide trading volumes of Bitcoin.

Tweets from Twitter are not restricted to any country or language either as the hashtag #BTC is used in all languages. On Google Trends, the user can opt to retrieve worldwide data and use the search term ‘Bitcoin (Cash)’, which includes all translations of Bitcoin. Thus, Google search data can be used without country limits. Wikipedia data is available on a daily basis for all translation of the ‘Bitcoin’ page and Reddit posts in the Subreddit r/Bitcoin can be posted in any language and accessed worldwide.

Accordingly, the empirical analysis is not bound by national jurisdiction and is representative of the worldwide Bitcoin economy.

3.2 Data sources

Here, the sources for the data, which will later be used in the empirical analysis will be listed and explained. Furthermore, the data retrieval process for each source will be declared. The section is divided into the three cryptocurrencies, the interest proxies and the robustness variables.

3.2.1 Cryptocurrencies

Data for Bitcoin, Ether and Ripple is split up into the price, the price volatility and the trading volumes of each currency. Price volatility refers to the intraday volatility and not the return volatility. The return volatility will be calculated directly from the price data; thus, no additional data is required for this variable. Furthermore, two measures of trading volumes are defined: the daily number of coins traded and the monetary trading volume, which is the daily number of coins multiplied with an approximation of the price paid for them.

3.2.1.1 Price

The price of cryptocurrencies is the amount of money an investor must pay to receive a certain amount of crypto-coins. Prices are always denoted in the amount of currency per one unit of cryptocurrency, i.e. one coin. For example, the price of Bitcoin on October 24th, 2018 was 6,507.67 USD/BTC (US Dollar per Bitcoin) (CoinMarketCap, 2018e). It is, however, possible to only buy in increments of one coin, e.g. one hundredth of a Bitcoin would cost 65.08 USD.

Since investors all over the world buy cryptocurrencies and exchange various currencies for crypto-money, there are several different prices, i.e. exchange rates. In that sense, there is more than one price for cryptocurrencies such as Bitcoin. Through arbitrage, however, these prices are harmonized as arbitrageurs try to exploit exchange anomalies and ensure that prices are consistent across different currencies. Short-term deviations from the concept of purchasing power parity as discovered by Dornbusch (1976) and others can be neglected as cryptocurrencies are rarely used for the purchase of real goods (Yermack, 2015).

Nevertheless, prices of cryptocurrencies differ across exchanges depending mostly on the liquidity of the exchange (Pisani, 2017). Therefore, an average of the biggest exchanges represents the global cryptocurrency price best.

According to 2.1.3.2, the most exchanged currencies are USD, EUR and JPY, thus these three currencies will be used to check whether the development of the different currencies is virtually equal. The correlation between the currencies was, as expected, almost perfect (EUR-USD: 0.999124, EUR-JPY: 0.998319, EUR-USD: 0.997644).

Accordingly, from here on, prices will be defined as the mean average price of the different USD prices from all major exchanges (h):

$$BTC\_PRC_t = \frac{1}{h}\sum_{i=1}^{h} PRC\_USD_t$$

Equation 3-1 *Bitcoin Price*

All price data was retrieved from Bitcoinity.org (https://data.bitcoinity.org/markets/price).

For Ether and Ripple, the price information was retrieved from CoinMarketCap (CoinMarketCap, 2018b, 2018c).

3.2.1.2 Price Volatility

Volatility is closely linked to the price in that it describes the magnitude of its movements, i.e. how much the prices deviate around a mean. Typically, volatility is calculated as the standard deviation of all data points in a given time range. The standard deviation is superior to the variance of a sample, because the unit of the standard deviation is the same as the underlying metric.

For cryptocurrencies, the volatility could be defined as the deviation of the prices within an hour, a day, a week or even a month. However, in accordance with all other data used for this empirical analysis, volatility will be denoted in 24-hour ranges.

As seen in 3.2.1.1, the price metric only allows for one price data point per day, which makes the calculation of a daily standard deviation impossible as more price data points are needed. Ideally, the volatility would be calculated as the standard deviation from all prices of all transaction that were executed within one day. Accordingly, price data would have to be retrieved from all exchanges across all currencies. Unfortunately, the magnitude of data would go beyond the scope of this thesis.

Thus, the prices from all transactions of one major exchange will be used, which will then be assumed representative for all other exchanges and currencies. The decision for USD was made in accordance with the decision to use USD as currency unit of prices.

For Bitcoin, the data set was retrieved from BitcoinCharts.com's API (http://api.bitcoincharts.com/v1/csv/), where .csv-files of all major exchanges were available for download. The files were already split up in the different currencies. As all transactions from the Coinbase exchange would have totaled in a file of more than six gigabytes of data, the decision fell in favor of BitStamp. The Europe-based exchange recorded 28,187,539 transactions between September 13, 2011 and October 25, 2018 and is therefore adequate for the empirical analysis.

To calculate the daily volatility, the price information was extracted for each 24-hour window and then the standard deviation of these price data points was calculated using a Python script (appendix A.3) and finally saved to a new file. The standard deviation was calculated using the formula:

$$BTC\_VOLA_t = \sqrt{\frac{\sum_{i=1}^{n}(PRC_{i_t} - \overline{PRC}_t)^2}{N-1}}$$

Equation 3-2 *Sample Standard Deviation*

, where PRC_i represents all prices within the 24-hour window, $\overline{PRC}$ is the average of all these observations and N is the number of observations.

Alternatively, https://data.bitcoinity.org offered volatility charts for Bitcoin. However, as noted by the operators of the website, their price volatility for longer periods is calculated as the average of hourly volatilities (Bitcoinity.org, 2018a). Nevertheless, daily data was retrieved from the website for the same period, which offered the intra-day price volatilities of each exchange. Since some exchanges' extreme volatilities caused a skewed distribution, the median across all exchanges was used instead of the mean average. This median was then compared to the volatility calculated using only BitStamp transaction. As expected, the correlation between the two data sets was very high (0.910345431), therefore the more accurate intra-day volatility from BitStamp transactions was used.

One other way to approximate the daily volatility of prices is to calculate the standard deviation of the highest and the lowest price on a specific day. As this method only uses two data points, it is less accurate then the first two alternatives. However, as these two price points are readily available and the small data set results in a fast and resource-efficient computation, it is a valid alternative to the more data-heavy computations mentioned before. The correlation between the volatility calculated from transactions and the volatility estimated from the high and low price was calculated as well and exhibited an even stronger correlation than the previous data sets (0.970255892).

Given this near-perfect correlation and the advantage of easy computation, Ether's and Ripple's volatility was calculated from the highest and the lowest price. Consistent with price data, the initial data came from CoinMarketCap (CoinMarketCap, 2018b, 2018c).

3.2.1.3 Trading Volumes

Trading volume represents the number of crypto-coins traded within a certain time frame. Usually this time frame is one day – 24 hours to be more specific. For securities, exchanges are typically open for public trading during business hours (approximately 9am – 5pm) and on weekdays depending on the time zone of the location of the exchange. Cryptocurrencies, on the other hand, are tradable 24 hours on any given day of the week. This raises the question of what time data providers use to distinguish between one day and another.

Unfortunately, the providers used to retrieve trading volume data for this analysis do not share information about which time zone is used to differentiate between days. The providers could either define a day according to the time zone of the location they are operating in, or the headquarters of the exchange. Alternatively, they could adjust the trading volume according to the location of the user. For the sake of this thesis, it shall however be sufficient to consider the trading volume as the number of traded Bitcoins within a 24-hour time frame.

Another definitory question is how trading volume is denoted. The simplest denotation is the mere number of coins traded, which, unfortunately, does not give an insight into the actual money moved. For these insights, the trading volume would have to be multiplied by the price (in the respective currency) at the time of fulfillment of the trade order. This information is tedious to retrieve and causes another problem that was already experienced within the price metric, namely which price should be used to align the different volumes.

As an approximation, this analysis will define two trading volumes of Bitcoin: The first will be the number of Bitcoins traded (BTC_TRADE_BTC) and the second will be the monetary trading volume (BTC_TRADE_USD).

BTC_TRADE_BTC will simply be:

$$BTC\_TRADE\_BTC_t = \sum_{i=1}^{24} BTC$$

***Equation 3-3** Daily Trading Volume Bitcoin*

, where i represents hours and the sum represents the number of Bitcoins traded within 24 hours.

BTC_TRADE_USD, however, will be calculated using the previously defined variables BTC_TRADE_BTC and BTC_PRC:

$$BTC\_TRADE\_USD_t = BTC\_TRADE\_BTC_t * BTC\_PRC_t$$

***Equation 3-4** Daily Monetary Trading Volume*

The number of Bitcoins traded is available on Bitcoinity.org (https://data.bitcoinity.org/markets/volume) and is then calculated as the sum of all major exchanges. As defined in **Equation 3-4**, the number of Bitcoins traded was then multiplied with the average price.

Alternatively, CoinMarketCap offers the trading volume in monetary value. To crosscheck the trading volume as defined here, another trading volume data set was downloaded from CoinMarketCap and compared to the first data set (CoinMarketCap, 2018a). The correlation between the two trading volumes was high (0.852180143), but not perfect, which may be due to different definitions of trading volume. However, since it is more consistent to use data from the same source, the trading volume calculated with data from Bitcoinity.org will be used.

For Ether and Ripple, the monetary trading volume from CoinMarketCap was used as price and volatility data comes from this source, too (CoinMarketCap, 2018b, 2018c). Since the mere number of Ether and Ripple coins was not available at CoinMarketCap, it had to be estimated through a conversion from **Equation 3-4**:

$$ETH\_TRADE\_ETH = \frac{\mathrm{ETH\_TRADE\_USD}}{\frac{1}{2} * (ETH\_PRC_H + ETH\_PRC_L)}$$

***Equation 3-5** Daily Trading Volume Ether*

$$XRP\_TRADE\_XRP = \frac{XRP\_TRADE\_USD}{\frac{1}{2} * (XRP\_PRC_H + XRP\_PRC_L)}$$

***Equation 3-6** Daily Trading Volume Ripple*

, where the average price as shown in the denominator was simply assumed to be average between the highest and the lowest price of that day.

3.2.2 Proxies for Investors' Interest

In 2.3.2, researchers have shown that the interest of investors is linked to assets. To approximate this interest, previous research has used different sources. Here, the relevant proxies for investors' interest, which will be used in the empirical analysis of this thesis, will be listed and explained in more detail. Additionally, the data retrieval process for the three proxies will be explained.

The Google search engine represents a tool for investors to search for information, while Twitter is used as a platform to share and distribute information across peers. Reddit, on the other hand, is used for both information-search and information-sharing. The relevant information is the activity level on each proxy, which will be approximated by the number of search queries on Google, the number of Tweets on Twitter and the number of posts on Reddit respectively.

3.2.2.1 Google search

As already mentioned in 2.3.2, Google search queries have established themselves as a viable source to analyze users' search behavior. Google offers users the possibility to search for information and thus constitutes a proxy for the information-search component of investors' interest.

The Google search engine clearly dominates the market: Between September 2017 and September 2018, 92% of worldwide online searchers used Google in their quest to find information (StatCounter, 2018). According to Statista (2018b), in all major countries except for Russia and China, Google was the number one search engine. 38.96% of Russian users use Google while in China only 5.72% use Google as the search engine was shut down in 2010 due to restrictions imposed by the Chinese government (Arrington, 2010).

Google offers its own tool to access information about the search engine. Via Google Trends (https://trends.google.com/), users can look up and retrieve information on certain search terms. One can either enter the plain search term or choose the related search term from Google. The difference between the two is that the plain search term only returns results for exactly that search term while the suggested search term includes translation into other languages and typographical errors.

After the user has chosen a search term, a graph is shown, which depicts the interest level for that specific topic over a certain period. Location, time range, category and search type are options within this screen. By default, data from Google Trends that dates back longer than eight months is only available on a weekly basis. Since four search terms can be compared at most, the maximum time range for daily data is 32 months. The graph for the search term 'Bitcoin (Cash)' for the period from October 2016 to September 2018[3] can be found in the appendix (A.2).

One thing to note about the Google Trends data is that Google does not make available the real number of searches but instead offers normalized search data. The day with the highest search volume is set as 100 and all other data points are set in respect to the maximum number of searches. For example, if the maximum number of searches was 100,000 on any day, a volume of 30,000 on another day results in a data point of 30 for that day. Thus, there is no possibility for the user to retrieve the actual number of searches – only the relative numbers are available. The downloaded data is available in 'comma-separated values'-files (.csv) and was retrieved on October 29, 2018

[3] Downloaded via the query: http://bit.ly/2qa26AR

using the query https://trends.google.com/trends/explore?date=2017-02-01%202017-06-30,2017-07-01%202017-11-30,2017-12-01%202018-04-30,2018-05-01%202018-09-30&geo=,,,&q=%2Fm%2F05p0rrx,%2Fm%2F05p0rrx,%2Fm%2F05p0rrx,%2Fm%2F05p0rrx for the observation period from February 2017 to September 2018.

The hypothesis for the later analysis is that changes in the normalized number of searches for the term 'Bitcoin' significantly impact the different metrics of Bitcoin, i.e. the price, its volatility and the trading volume.

3.2.2.2 Twitter

Twitter is an online social network, where users can post messages to the network, which are used to communicate among the users. These messages are called tweets and are meant to be quick and short pieces of information, thus tweets are limited to 280 characters (Rosen, 2017). Users can retweet, i.e. repost a tweet, comment on a tweet, like tweets and directly message each other. The concept of following enables users to see the tweets of their favorite users. Twitter is also used by artists, politicians and starts to communicate to their followers, most famously the president of the United States of America, Donald J. Trump (@realDonaldTrump).

To categorize tweets, Twitter introduced the option to include hashtags in tweets, which are then grouped together according to the characters following the hashtag, e.g. #BTC. In 2012, Twitter additionally introduced 'cash tags', which follow the concept of hashtags, but instead use the dollar sign and are used for ticket symbol of stocks, e.g. $AAPL (Kim, E., 2012). In 2017, an average of 125 million hashtags were tweeted per day (Twitter Inc., 2018)

Unfortunately, Twitter does not offer a tool like Google Trends, which lets users easily download information about the network. Instead Twitter offers application programming interfaces (APIs), through which data requests are facilitated. Different tiers of APIs currently exist: a standard, a premium and an enterprise endpoint. They differentiate themselves in their capabilities and their price.

The hypothesis is that changes in the number of tweets per day relating to Bitcoin (through #BTC) significantly impact Bitcoin's metrics. Thus, the needed information from Twitter is the number of tweets per day for tweets, which contain the term 'BTC'. The request for count data is, unfortunately, handled through Twitter's premium API[4], which is priced at 99$ per month (Twitter, 2018).

To obtain access to Twitter's APIs, a developer account must be created and a subscription to the relevant API is needed. Every developer receives credentials, which are later used to authenticate with the API. Requests can be sent via different tools, one of which is Postman (https://www.getpostman.com). Using URLs, Postman can send so-called GET-requests to the respective API. The URL includes certain parameters, such as the from- and to-date and the query. The application, if successful, then delivers the expected response in JSON[5] format as seen in **Figure 3-2**:

[4] For reference see: https://developer.twitter.com/en/docs/tweets/search/api-reference/premium-search.html#CountsEndpoint

[5] JavaScript Object Notation

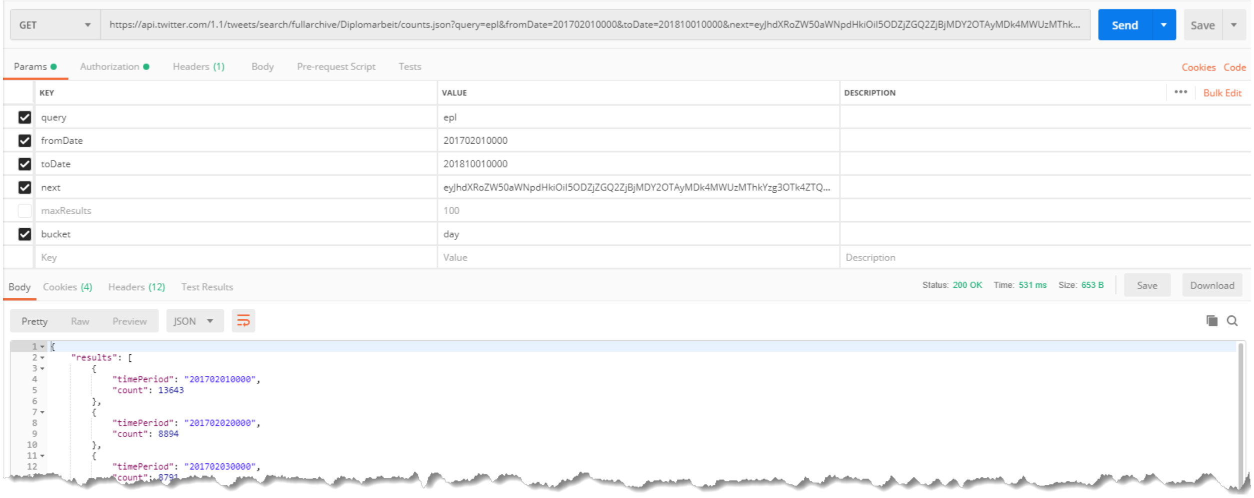


Figure 3-2 Example of a GET-request to Twitter's API using Postman (own creation)

3.2.2.3 Reddit

In 2005, Reddit (https://reddit.com) was founded by Alexis Ohanian and Steve Huffman. It is an online platform, where users can engage in discussion about topics of their interest (Crunchbase, 2018). Users can post content to the network, such as links, text posts and images. At the same time, users can also vote on existing content by voting it up or down.

To categorize content by subject, subreddits are used. Subreddits are created by the users themselves and follow the URL convention https://reddit.com/r/[subreddit], where [subreddit] can be any topic of the users' choice, e.g. Bitcoin-related topics are posted to the subreddit https://reddit.com/r/Bitcoin.

In 2017, the Reddit community created close to 1.2 million subreddits (Richter, 2017) and the website reported more than 1.6 billion monthly visits in July of 2018 (Statista, 2018a). In October 2018, Reddit was the fifth most visited website in the US and took place 17 globally (Alexa Internet Inc., 2018b, 2018a).

Reddit offers Bitcoin investors a forum-like platform, where they can post and exchange information, discuss latest developments around Bitcoin and thus serves as an excellent proxy for investors' interest. In comparison to forums like https://bitcointalk.org, Reddit is superior in that the platform is not Bitcoin-specific, has a larger user base and is globally better known than any Bitcoin-specific forum. Thus, Reddit is more likely to be the first contact point for investors who are interested in Bitcoin.

Accordingly, the number of posts per day are supposed to give an insight in the activity level of Bitcoin investors. The hypothesis is that changes in the number of posts, like the number of tweets and search queries on Google, have a significant impact on the Bitcoin economy.

Unfortunately, the retrieval of the number of posts within a certain subreddit is currently not possible via Reddit's public API (Reddit Inc, 2018). Therefore, a custom API was used, namely https://api.pushshift.io/reddit. This API allowed for the search within a certain subreddit and the specification of a time range.

For example, a request for the subreddit r/Bitcoin from 09/30/2018 to 9/29/2018 is sent with the following URL:

https://api.pushshift.io/reddit/search/submission/?subreddit=bitcoin&filter=id,created_utc&sort=desc&size=10000&before=1538265600&after=1538179200

(Note that dates must be converted to computer time.)

The request then returns the ID and the creation time of each post that satisfies the search criteria (see **Figure 3-3**). The ID can then be used to verify the post by using https://reddit.com/r/Bitcoin/comments/[ID], e.g. https://www.reddit.com/r/Bitcoin/comments/9k1qdb/.


```
{
    "data": [
        {
            "created_utc": 1538265373,
            "id": "9k1qdb"
        },
        {
            "created_utc": 1538264618,
            "id": "9k1ng1"
        },
        {
            "created_utc": 1538264571,
            "id": "9k1n8s"
        },
        {
            "created_utc": 1538263944,
```


***Figure 3-3** Example of API response*

Unfortunately, the API's response limit is set to 1000, thus the time range needed to be limited to three hours as on certain days the number of posts was as high as 4191.

To automatize the adjustment of the request URL to fit the required search criteria, a Python script was written, which can be found in the appendix (A.4). First, the URL is looped so that the requests retrieve posts in three-hour intervals within a certain time range as defined in line four and eight respectively. The response of each request was counted for the number of items and the count was saved to a dictionary (countdict). At last, the count dictionary was sent to an output file (line 27f.).

Since the file reflected the number of posts within three-hour intervals, the data file was later modified to represent the daily number of posts.

3.2.3 Robustness variables

As an alternative resource for investors to gather information (information search), Wikipedia is an established proxy for investors' interest. Wikipedia is a free encyclopedia, which was founded in 2000 by Jimmy Wales and Larry Sanger and is now hosted under the parent company Wikimedia Foundation. The website offers articles about topics, which are written and edited by the network of users. According to Wikipedia, approximately 72,000 contributors are working on more than 48 million articles in more than 300 languages. Of all articles, close to 6 million (~ one eight) were written in English. (Wikimedia Foundation, 2018)

Similar to Google, users can search for specific topics and find information on that topic. For example, Bitcoin's English Wikipedia page contains information about the history of Bitcoin, its architecture, the ideology behind Bitcoin, its economics and legal status and criticism of the cryptocurrency (Wikipedia, 2018). As such, it offers investors an abundance of information and helps them in their decision to buy Bitcoin.

The number of page views of a specific Wikipedia page can be retrieved using a tool that was created through the Toolforge environment from the Wikimedia Foundation. Toolforge offers developers access to develop services based on Wikipedia data. More specifically, the page view analysis tool (https://tools.wmflabs.org/pageviews) offers users the option to view the page views of a specific term in all languages. Using the URL https://tools.wmflabs.org/langviews/?project=en.wikipedia.org&platform=all-access&agent=user&start=2017-02-01&end=2018-09-

30&sort=views&direction=1&view=list&page=Bitcoin, the number of page views for the Bitcoin Wikipedia page was downloaded in .csv format. The data set included 94 unique languages. The page view count for the observation period ranged from around 21 million (English) to only 80 (Egyptian Arabic). To approximate the worldwide page count, all languages were later summed up.

Since Wikipedia serves as a substitute and/ or complement for information search for investors who seek information, the number of page views are expected to have a similar impact on the cryptocurrency economy than Google searches.

To check for the robustness of Bitcoin-related Reddit posts, the post count retrieval as described in 3.2.2.3 was repeated for a subreddit that is unrelated to the cryptocurrency topic. Taylor Swift was assumed adequate for this purpose, which was supported by the weak correlation between the two posts counts (0.303453665).

The same approach was followed for Twitter data, where an unrelated hashtag will be used, in this case #EPL, representing the English Premier League. According to the data, there is no pronounced relationship between the #EPL and #BTC series as the correlation is -0.04713970.

The random Reddit post and Tweet counts should not have a significant impact on cryptocurrencies as no interdependencies are expected between either Taylor Swift or the English Premier League and any cryptocurrency.

3.3 Descriptive Statistics

After all data was retrieved from the respective sources, the first step before it will be used in the final analysis is to take a closer look at the data set, describe the different variables, analyze the correlation between the variables and check the data for inconsistencies. The variable names are depicted in **Table 3-1**:

#	Variable	Description (defined in Chapter)
1	BTC_PRC	Bitcoin price in USD (3.2.1.1)
2	BTC_TRADE_BTC	Number of Bitcoins traded (3.2.1.3)
3	BTC_TRADE_USD	Monetary Trading Volume Bitcoin (3.2.1.3)
4	BTC_VOLA	Price Volatility Bitcoin (3.2.1.2)
5	BTC_GT	Google Trends data for the search term 'Bitcoin (Cash)' (3.2.2.1)
6	BTC_TWITTER	Number of Tweets containing the hashtag #BTC (3.2.2.2)
7	BTC_REDDIT	Number of Reddit posts in the subreddit r/Bitcoin (3.2.2.3)
8	ETH_PRC	Ether price in USD (3.2.1.1)
9	ETH_TRADE_ETH	Number of Ether coins traded (3.2.1.3)
10	ETH_TRADE_USD	Monetary Trading Volume Ether (3.2.1.3)
11	ETH_VOLA	Price Volatility Ether (3.2.1.2)
12	XRP_PRC	Ripple price in USD (3.2.1.1)
13	XRP_TRADE_XRP	Number of Ripple coins traded (3.2.1.3)
14	XRP_TRADE_USD	Monetary Trading Volume Ripple (3.2.1.3)
15	XRP_VOLA	Price Volatility Ripple (3.2.1.2)
16	BTC_WIKI	Number of page views of Bitcoin's Wikipedia page (3.2.3)
17	EPL_TWITTER	Number of Tweets containing the hashtag #EPL (3.2.3)
18	TS_REDDIT	Number of Reddit posts in the subreddit r/TaylorSwift (3.2.3)

Table 3-1 *Overview of Variable Names*

At this point, the data set consists of 18 different variables, excluding the 'DATE' variable, which sorts the data chronologically and the two variables 'ETH_AVG' and 'XRP_AVG', which were used to calculate the trading volume as mentioned in 3.2.1.3. Each variable has 607 observations comprising daily data points for the observation period from February 1st, 2017 to September 30th, 2018. The sample does not include missing values; thus, no entries must be omitted.

The set is divided in Bitcoin metrics (1-4) [see 3.2.1], the proxies for investor interest (5-7) [see 3.2.2], the Altcoin metrics (ETH: 8-11, XRP: 12-15) and the robustness variables (16-18) [see 3.2.3]. Since tables and figures would be two crowded if the entire data set was analyzed at once, descriptive statistics will be provided for each block.

3.3.1 Cryptocurrencies

	BTC_PRC	BTC_TRADE_BTC	BTC_TRADE_USD	BTC_VOLA
Mean	6098.103268385492	147591.7759437038	922429938.5599616	112.5663509308583
Median	6383.031184297204	132332.9582266065	610902463.2572371	62.2738158532145
Maximum	19389.01279310084	620987.0576831199	7554416528.251469	832.737146404337
Minimum	935.4322232578262	28973.11486680858	45993016.16099853	2.33833719202624
Std. Dev.	3915.067536933681	79607.28952953411	920981250.0216176	133.3419031796938
Skewness	0.7318299573765224	1.847659925091814	2.521431136405733	2.368139678057084
Kurtosis	3.339715208924283	8.408877244383401	12.25798308052945	9.248236500092885
Observations	607	607	607	607

Table 3-2 *Descriptive Statistics of Bitcoin*[6]

As seen in **Table 3-2** and **Figure 3-4**, the median of the Bitcoin price is higher than the mean price, indicating a slightly left-skewed distribution. The highest price of $ 19,389.01 was reached on December 17th, 2017 (observation 320) while the lowest price was reported on March 25th, 2017 (observation 53). Prices deviated extremely in the period with a standard deviation of close to two thirds of the average price.

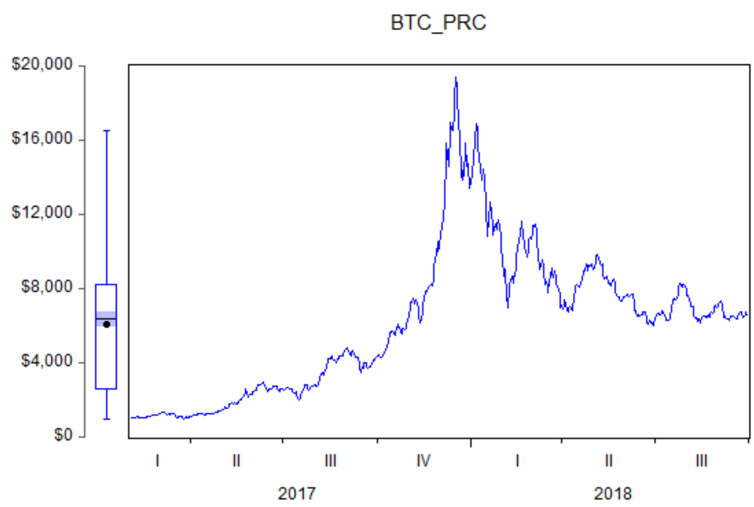


Figure 3-4 *Bitcoin price evolution with Boxplot*

In contrast to prices, both trading volumes are skewed to the right indicating that a few large trades pull the mean higher than the median. Interestingly, the largest trade of Bitcoin occurred on February 6th, 2018, where the average price was $6,899.39 while the most money was traded for Bitcoin on December 22nd, 2017 where the price was close to its all-time high. Furthermore, the deviation of the monetary trading volume is significantly higher (~100 % of average) than the deviation of the number of Bitcoins traded (~50 % of average), which is likely in part due to the inheritance of volatility from the Bitcoin price itself.

Intra-day price volatility is also strongly skewed to the right and the maximum data point was reached on the same that the monetary trading volume was highest (12/22/17). Interestingly, the 48 lowest intra-day volatilities were recorded in the first three months (February through April) of the sample, suggesting that price volatility has taken off in February and never resettled.

[6] Column 2,4,5 in USD; column 3 in BTC

	ETH_PRC	ETH_TRADE_ETH	ETH_TRADE_USD	ETH_VOLA
Mean	400.2206425041187	3662468.72236889	1524985522.135087	24.49056764655954
Median	322.11	3292211.095791478	1373219968	15.26643540581758
Maximum	1396.42	12069963.40114052	9214950400	294.927167365097
Minimum	10.73	499162.138475022	5695440	0.0707106781186545
Std. Dev.	284.8405033463748	2087765.890965961	1385618179.215334	30.36284568067386
Skewness	0.8765627242214715	1.334454289575641	1.981757138467727	3.50968524047709
Kurtosis	3.48163082649035	5.070614794390958	9.069705991485019	21.49134514994305
Observations	607	607	607	607

Table 3-3 *Descriptive Statistics of Ether*[7]

Table 3-3 shows the descriptive statistics for the second cryptocurrency under investigation – Ether while **Figure 3-5** visually depicts the price development over time. Compared to Bitcoin, the median of this series is smaller than the average, resulting in a right-skewed distribution that is caused by many small price points compared to a few large ones that drive the two averages apart. The maximum value of the Ether price series was reached almost a month later than Bitcoin's on January 13th, 2018 (observation 347). On the other hand, the lowest price was reported on the first day of the observation period, indicating that the price has been increasing ever since then. The standard deviation of the Ether price compared to the mean is slightly higher than Bitcoin's ratio (~ 71% vs. 64%) suggesting that Ether prices were even more volatile.

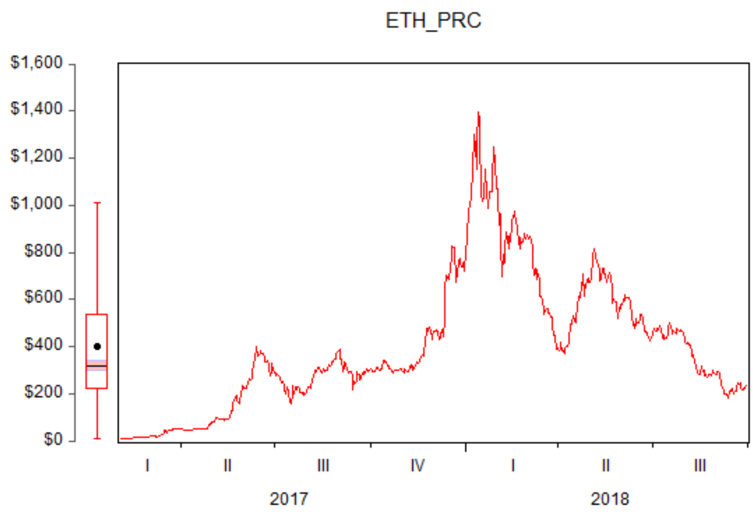


Figure 3-5 *Ether price evolution with Boxplot*

The most Ether coins were traded on September 21st, 2018 (observation 598), while the most money was traded on January 10th, 2018 at an Ether price of $1,255.82 - three days before the highest price was reached. Similar to the lowest Ether price, the lowest trade of Ether coins occurred almost in the beginning of the observation period (observation six), which coincides with occurrence of the lowest monetary trading volume. Both trading volumes have higher means than medians, signaling that a few large trades push the mean relatively higher as has already been witnessed with Bitcoin.

Ether's intra-day volatility is strongly impacted by a few days with extreme price deviations, which reveals itself in the significantly higher mean compared to the median of the series. The 29 most volatile days lie within one month before and after the peak of the price series, marking the most defining period for Ether.

[7] Column 2,4,5 in USD; column 3 in ETH

	XRP_PRC	XRP_TRADE_USD	XRP_TRADE_XRP	XRP_VOLA
Mean	0.4591376227347613	573108602.215814	970529307.8049782	0.04095666620530734
Median	0.2846	230418000	664735732.7350539	0.01621890824007583
Maximum	3.38	9110439936	9360196945.832242	0.9121677477306458
Minimum	0.005408	230479	36488403.38795219	0.000049497475
Std. Dev.	0.4795303322295606	1124543158.798536	1035301151.044416	0.08467419600857801
Skewness	2.622950355474678	4.269910136807612	3.454901186659171	5.150253011079335
Kurtosis	12.67720802170572	24.38490497341811	19.04128197983273	36.61423946082648
Observations	607	607	607	607

Table 3-4 *Descriptive Statistics of Ripple*[8]

Ripple, the third and last cryptocurrency under investigation exhibits a price development (see **Figure 3-6**) that is significantly different from Bitcoin and Ether. The single peak at the end of 2017 is more pronounced and not followed by a second peak as seen in Bitcoin's and Ether's price evolution. The peak does also not announce itself in the preceding months as witnessed in Q2 and Q3 of 2017 of the other two cryptocurrencies. In that way, Ripple's price surge is more immediate and explosive.

Ripple's mean-median proportionality is the same as Ether's (see **Table 3-4**), thus the price distribution also exhibits the same right-skewness. The standard deviation is extremely high (168% times the mean) compared to the other two cryptocurrencies, which is likely due to the overall low level of prices. The maximum price of \$3.38 was reached on January 7th, 2018 (observation 341) – six days prior Ether's peak and 21 days after Bitcoin's. The lowest price was recorded on March 1st, 2017 (observation 29).

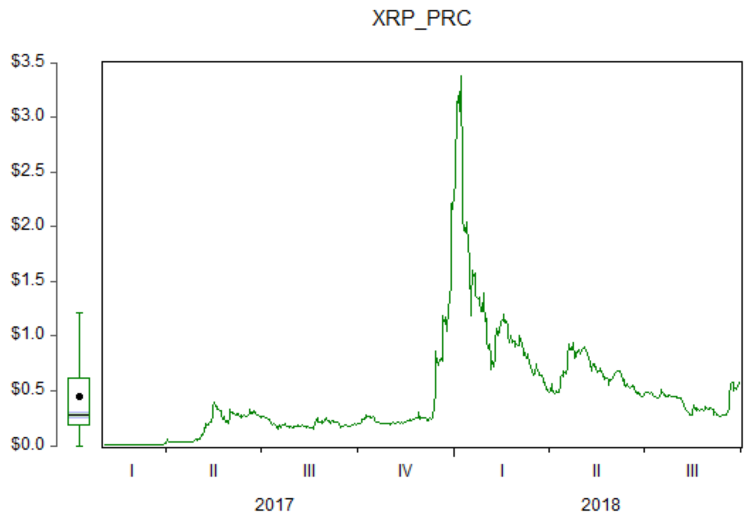


Figure 3-6 *Ripple price evolution with Boxplot*

The fewest coins were traded on February 12th, 2017, but interestingly the period between October 26th and October 31st of 2017 marks a period of very low coin trading, too. Similar characteristics appear between September 20th and September 24th of 2017. These occurrences of low trading late in the observation period are not in accordance with Ether and Ripple where trading saw a relatively continuous increase over the entire observation period. The highest number of coins traded occurred on December 14th, 2017 at an average trading price of \$0.67933. In contrast, the most trading denoted in USD happened on January 18th, 2018 shortly after Ripple's price peak. As also observed for Ether, Ripple's highest monetary trading volumes revolve one month around the highest price: the 35 highest trading volumes occur between December 2017 and February 2018

[8] Column 2,4,5 in USD; column 3 in XRP

with two exception on September 20th and 21st, 2017. The most coins were traded on December 14th, 2017 at an average price of $0.67933 – three days before the highest Bitcoin prices were recorded.

In **Figure 3-7**, the three price series of the cryptocurrencies are normalized and overlapped in the same graph. It appears as though Bitcoin is the leader of them. With some lag, Ether closely mimics the behavior of Bitcoin prices, but reacts slightly later then Ripple. Ripple on the other hand shows a slightly decoupled development in 2017, but couples with Ether and Bitcoin after the respective peaks between December 2017 and January 2018.

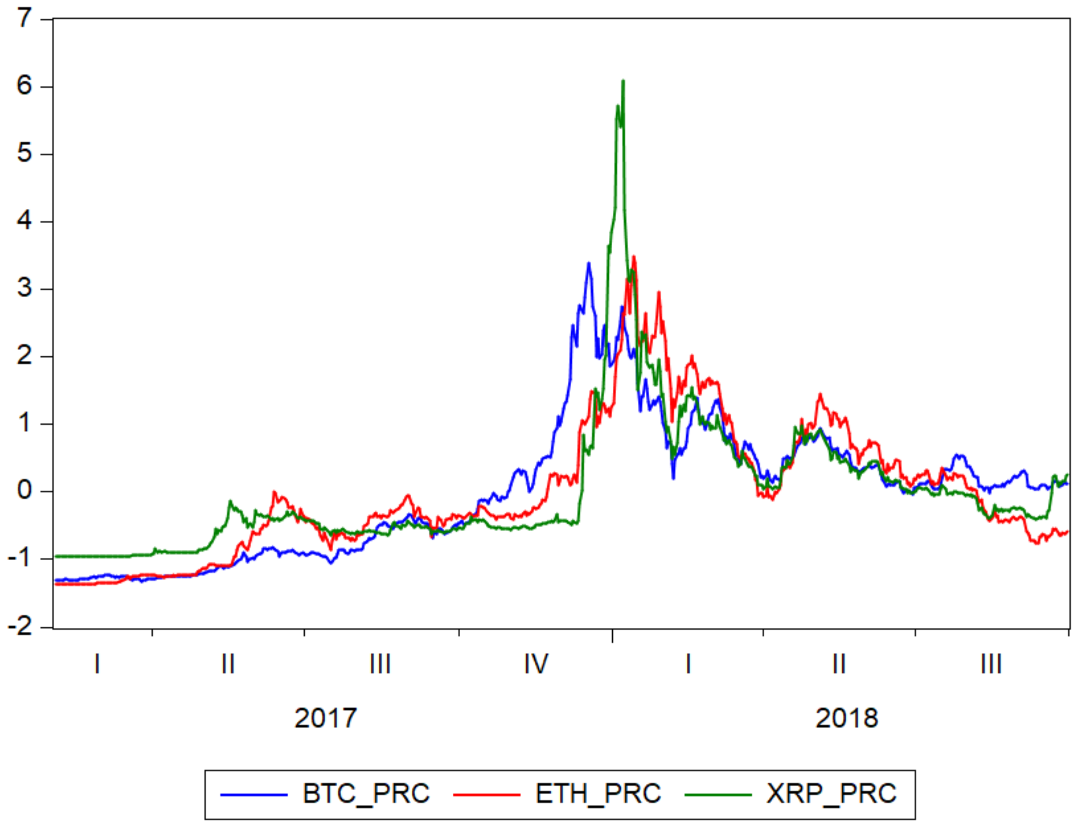


***Figure 3-7** Comparison of price development (normalized scale)*

However, the price series share common characteristic: all have their peaks within one month, which coincides with the highest trading numbers and the highest volatility. This synchronicity is also reflected in the high correlation between the various series. The only metric that is only weakly correlated with the other metrics is the number of coins traded, which is expected as it is decoupled from the price unlike the monetary trading volume, the volatility and the price itself. The detailed overview of the different correlations can be found in appendix A.5.

Similar among all cryptocurrencies, prices, trading volumes and volatilities see a continuous increase from the start of the observation period that leads to the peak season. Furthermore, they leave an initial level to which they never return afterwards.

3.3.2 Proxies for Investors' Interest

	BTC_GT	BTC_TWITTER	BTC_REDDIT
Mean	12.64579901153212	47868.61120263601	517.0296540362435
Median	9	53693	383
Maximum	100	143124	4191
Minimum	3	7519	125
Std. Dev.	12.35322052766326	28969.20793492015	431.8473928789532
Skewness	3.128277431172214	0.2096893928158363	3.873073274574524
Kurtosis	15.57923742822456	2.065762880440557	24.19766505416616
Observations	607	607	607

***Table 3-5** Descriptive Statistics for Interest Proxies*

Table 3-5 shows that the descriptive statistics of the three interest proxies differ significantly. However, like the different cryptocurrency prices, the three series exhibit a very similar development over time on a normalized scale.

As already mentioned in 3.2.2.1, the data downloaded from Google Trends does not correspond with the actual number of Google search queries, but a normalized scale of them. Thus, the maximum value is, by definition, 100. The minimum of the series, however, does not have to be 0 as the series does not have to contain a data point, which exhibits 100 times fewer searches. In the BTC_GT series this is not the case, which is why the minimum is 3. Note also that the series only contains integers, thus for example a value of 9.22 is not possible. Google Trends would round 9.22 to the next integer, i.e. 9.

The highest number of searches occurred on December 22nd, 2017, which is in accordance with the peak of Bitcoin prices five days earlier. The second pronounced peak (see **Figure 3-8**) was recorded 15 days earlier on December 7th at a value of 96. The date of the lowest number of searches cannot be determined as 39 dates have a Google Trends value of 3, ranging from February 1st, 2017 to April 24th, 2017. Presumably, the actual lowest number of search queries is likely to occur in February 2017 as similar data series show a continuously increasing trend and Google is likely not to be any different. The mean average of the BTC_GT series for the observation period was 40.5% higher (12.64) than the median (9) suggesting the few extreme values pushed the mean higher.

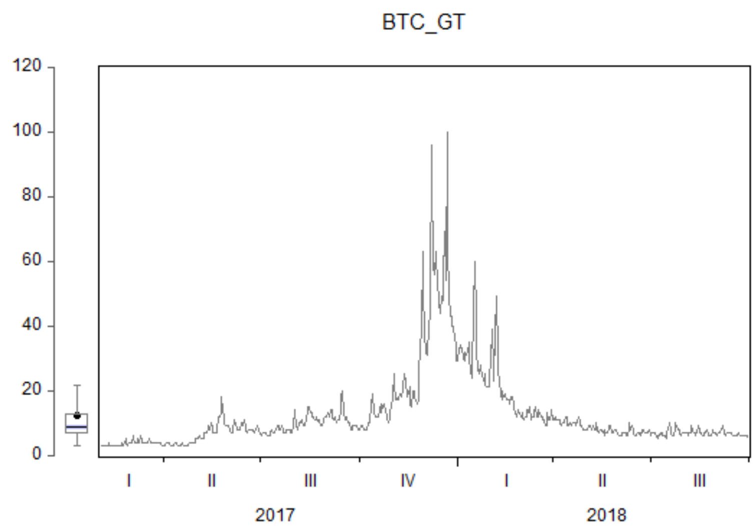


***Figure 3-8** Google search evolution with Boxplot*

At the same time, the standard deviation was roughly 100% of the mean, which does not mean that the actual searches deviated that heavily, but the normalized series. Since the normalized series is rounded to integers, the underlying series could be even more volatile, because small changes are not reflected in the normalized data set.

The number of tweets on Twitter related to Bitcoin, on the other hand, truly represent the number of tweets on any given day within the observation period. As such, an average of 47,868.61 tweets containing the term 'BTC' were recorded compared to a higher median of 53,693. Unlike BTC_GT, the distribution is left-skewed. Interestingly, the highest number occurred on April 26th, 2018 at 143124 tweets, more than four months after the highest Bitcoin price was reported. Only the fourth and fifth highest tweet volume occurred one month within the peak season of Bitcoin while only the tenth and eleventh highest volume was registered closely around the peak (12/7/17 and 12/20/17). In accordance with other data series, Twitter users tweeted the least about Bitcoin in February and March 2017. The number of tweets also deviated less than other series, exhibiting only approximately 60% deviation around the mean.

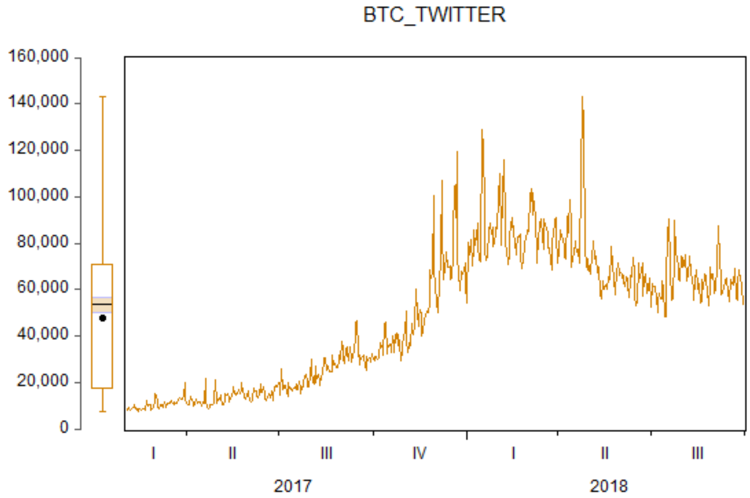


***Figure 3-9** Tweet volume evolution with Boxplot*

Referring to **Figure 3-9**, it is apparent that the number of tweets is differently distributed over time than all other series. The series shares the steep increase in Q4 of 2017 that concludes in a peak, but unlike BTC_GT or the cryptocurrency prices, there are several other peaks afterwards. Especially the main peak in April is unique to BTC_TWITTER. Furthermore, the overall level of tweets does not drop as heavily as other series. Presumably, Twitter users kept their interest level long after the peak in prices in December 2017 and January 2018 and in that way prolonged the cryptocurrency hype on Social Media.

The distribution of Reddit posts in r/Bitcoin exhibit a typical distribution with a steep increase before and a rapid decrease after the peak. This peak of BTC_REDDIT occurred on November 29th, 2017 at 4191 posts, more than two weeks before the Bitcoin price peaked. The second significant peak was recorded on December 7th followed by the third peak around December 20th. Over the 20 months, Reddit users posted an average of 517.03 posts while the median was

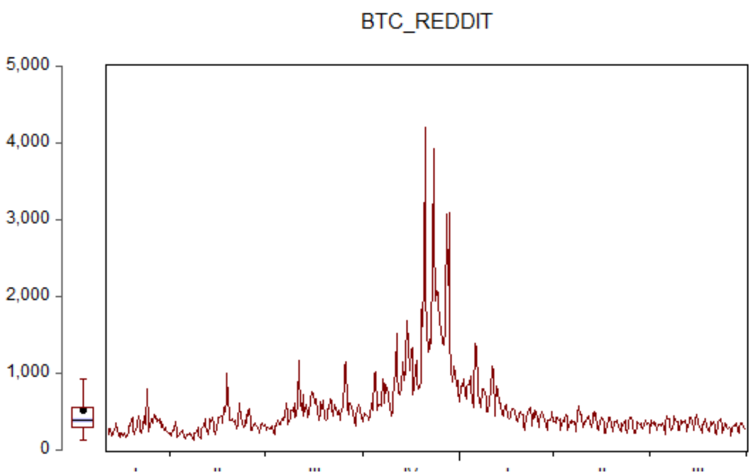


***Figure 3-10** Reddit posts evolution with Boxplot*

significantly lower at 383 posts constituting a right-skewed distribution. The number of posts was rather volatile with a standard deviation of 431.85, approximately 84% of the mean.

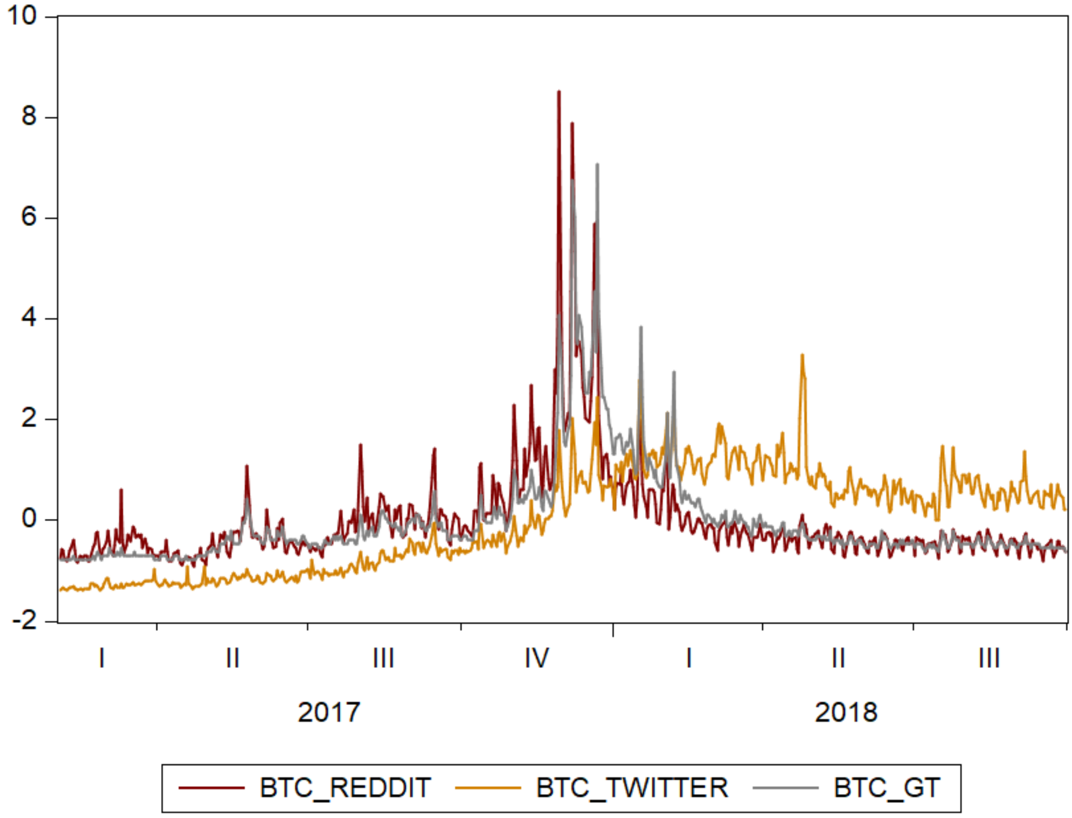


***Figure 3-11** Comparison of interest proxies' development (normalized scale)*

The comparison of the three series, as depicted in **Figure 3-11** confirms what has been already mentioned earlier: BTC_REDDIT and BTC_GT exhibit a trend that is extremely similar, while BTC_TWITTR differs significantly. All three series share a peak at the end of Q4 2017, but BTC_TWITTER is the only series that records abnormal volumes in Q2 2018. Furthermore, while Google and Reddit users seems to lose their interest almost immediately after a peak, Twitter users seem to build up an interest level more slowly but keep that interest more continuously.

The synchronicity between Reddit and Google is confirmed in the high correlation coefficient (0.89), while the significant difference of Twitter trickles down to a weak to medium correlation of 0.32 (with BTC_REDDIT) and 0.48 (with BTC_GT) respectively.

Summarizing, it seems that the behavior of Twitter and Reddit users differ drastically: Reddit users appear to be very fast in the processing of new information. A topic becomes obsolete very quickly (as seen in the many narrow peaks), while Twitter users slowly take up a topic and increase interest in it relatively slowly over time as it diffuses across the network. Accordingly, the pass-through time of information seems significantly higher in the Reddit network than in the Twitter network, which makes sense if the large, diverse Twitter network is considered and compared to the presumably Bitcoin-specific user network within the r/Bitcoin subreddit.

Alternatively, the different behaviors could also be explained by the construction of the networks: Within one post in a subreddit, there can be many comments to that post. That is important, because it is unlikely that users would create a new post on a current topic but rather comment on the existing post. This way the comments within the old post would keep the activity level of users high, without the need to create a new post. On the other hand, it is more likely that Twitter users create new tweets with the same topic, simply because they might not have an overview of the latest topics. Accordingly, the difference could be explained in what data is retrieved from the respective networks, i.e. the comment counts on Reddit might give a better insight into the network.

At last, the correlation between the different interest proxies and the cryptocurrency markets were analyzed and the results are depicted in **Table 3-6**. The three highest correlation coefficients of each interest proxy are highlighted in bold. Interestingly, seven out of the nine occur between the proxy series and the Bitcoin economy, which was to be expected since the proxy variables were constructed in a way that they convey the interest towards Bitcoin and not Ether or Ripple. All the more surprising that the Tweet volume is more highly correlated with the Ether price and monetary trading volume than other Bitcoin metrics. This could be due to the close relationship between the Bitcoin and the Ether price as seen in the previous chapter. Also noteworthy is that the correlation coefficients between the number of reddit posts and the two altcoins are rather weak with the highest reaching a value of 0.379717. The same trend is visible in the BTC_GT series, but the coefficients average at a moderate correlation level. Furthermore, the relationship between coin trading volumes and the proxies is notably weak: only two out of nine exhibit moderate correlation compared to six out of nine which are correlated moderately or higher with the monetary trading volume.

	BTC_TWITTER	BTC_REDDIT	BTC_GT
BTC_PRC	**0.833462**	0.527951	**0.714084**
BTC_TRADE_BTC	0.101529	**0.546743**	0.507
BTC_TRADE_USD	0.657771	**0.793193**	**0.908294**
BTC_VOLA	0.609549	**0.674753**	**0.839431**
ETH_PRC	**0.78537**	0.277596	0.512104
ETH_TRADE_ETH	0.290387	0.106474	0.157694
ETH_TRADE_USD	**0.729504**	0.311037	0.545968
ETH_VOLA	0.56854	0.379717	0.611368
XRP_PRC	0.646588	0.169461	0.442604
XRP_TRADE_USD	0.456947	0.287711	0.528089
XRP_TRADE_XRP	0.262634	0.318691	0.422615
XRP_VOLA	0.409025	0.271775	0.500122

Table 3-6 *Correlation between Interest Proxies and Crypto-Markets*

Between Bitcoin and the proxies, BTC_GT shows three strong correlations, while BTC_REDDIT and BTC_TWITTER only show one. On the other hand, the Reddit series is the most correlated proxy with all four highest correlation occurring between Reddit and Bitcoin. Twitter is the only proxy that shows a quasi-strong correlation coefficient with the Ripple economy (XRP_PRC).

Overall, 20 out of 36 pairs exhibit a moderate (13) or strong (seven) correlation, which strengthens the suspicion that the interest of investor for Bitcoin does, indeed, have an impact on the cryptocurrency markets or vice versa. However, correlations do not give answer on questions such as the directionality or more importantly causality. Thus, a correlation analysis is a valid starting point, but a more in-depth analysis is needed, to gather deeper insight in the dynamics between the interest of investors and the cryptocurrency economy.

3.3.3 Robustness variables

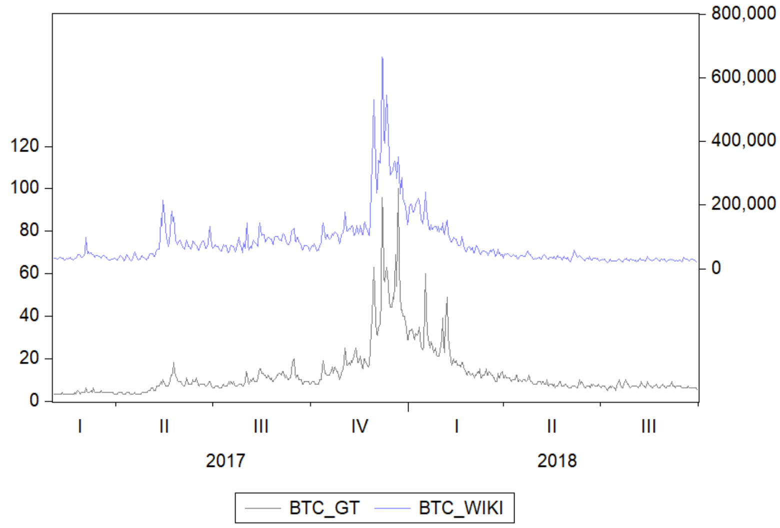


Figure 3-12 *Google Search volume vs. Wikipedia page views*

The number of page views of the Wikipedia article of Bitcoin in all languages is viewed on average 81067.76 times a day over the entire observation period. The median is significantly lower at 55377.00 views resulting in a right-skewed distribution. The most views of 665, 488 were recorded on December 7th, 2017 while the article was viewed only 18,929 times on July 7th, 2018.

In reference to **Figure 3-12**, the two series are almost identical with smaller differences in the mid of Q2 2017 and during the two smaller peaks in Q1 2018. The biggest difference occurs within the peaks: While BTC_GT exhibit two more pronounced peaks, the BTC_WIKI series appears to be more unimodal. Nevertheless, the visual simultaneousness between the two series is confirmed with a very strong correlation of 0.911568. Thus, the number of page views on Wikipedia constitute a valid alternative as information search proxy.

For robustness of the Bitcoin tweet volume, the tweet volume relating to the English Premier League is used. Compared to BTC_TWITTER the overall activity level of users is significantly lower: Over the same period only an average of 8051.48 tweets containing 'EPL' were recorded.

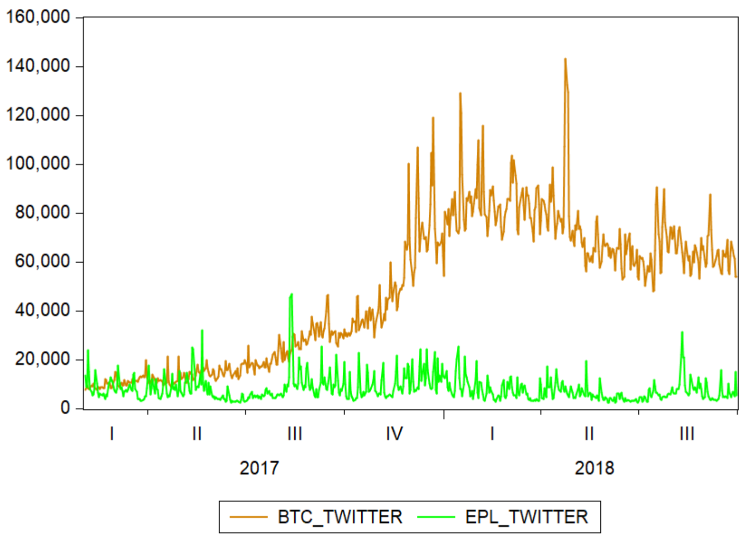


Figure 3-13 *Number of tweets containing 'BTC' vs. 'EPL'*

The median was even lower at 6453, resulting in a right-skewed distribution unlike BTC_TWITTER. The maximum number of tweets on a day (47033) is smaller than the average of the Bitcoin series and occurred on August 12th, 2017, which was the date of the start of the 2017/18 Premier League season. The ten days with the least activity were reported during the summer breaks of each season. However, the EPL_TWITTER series deviated slightly more than the BTC_TWITTER (64.84% vs. 60.51%).

As seen in **Figure 3-13**, the two series share no common characteristics. The most pronounced periods of EPL_TWITTER are the two summer breaks between May and August of each year, surrounded by peaks before and after the break. The peak before signals the heighted attention due the season finale, while the peaks after show the enthusiasm of fans at the beginning of each new season. Interestingly, the interest level at the end of season 2016/17 is significantly higher than in season 2017/18. This is likely to the dominance of Manchester City in the later season as the title race was decided early on (Oberstone, 2017) compared to the suspenseful season finale, which saw Chelsea take the title only in May 2017 (Basu, 2017). During the season, the tweet volume is relatively stable, with spikes on the respective game days, which are mostly from Friday to Monday. As already seen in 3.2.3, there is a nearly non-existent relationship between the two Twitter series, which was expected as the two series share no contextual commonalities. As such, the EPL series is expected to have no significant impact on the cryptocurrency economy.

The number of reddit posts in the subreddit r/TaylorSwift averaged at 28.97 posts per day while the median in the observation period was lower at 23. The maximum of 370 posts was reached on August 25th, 2017 while only one post was recorded on July 19th, 2017.

As seen in **Figure 3-14**, the series (TS_REDDIT) shows three major peaks, which coincide with the announcement of the new album "Reputation' (8/23/2017), the release of the album (11/10/2017) and the start of the Reputation tour (5/8/2018) (Swift, 2018).

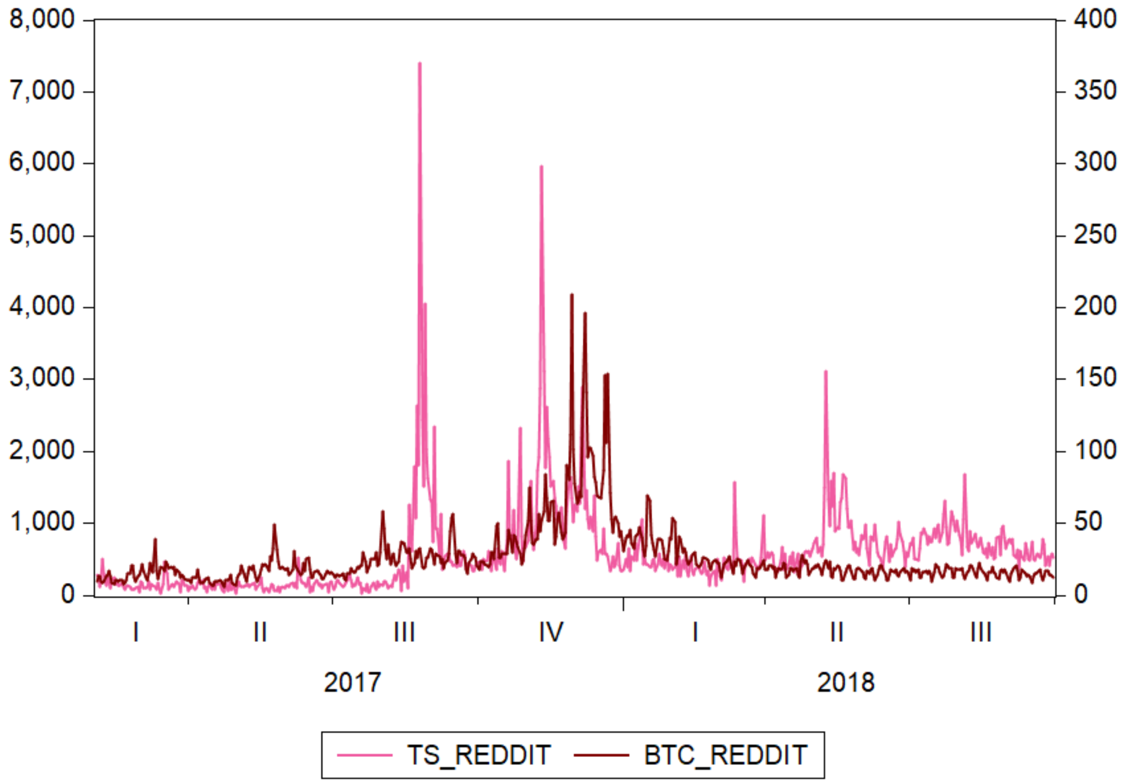


Figure 3-14 *Reddit posts: r/TaylorSwift vs. r/Bitcoin*

Compared to the number of reddit posts in the subreddit r/Bitcoin, there are no apparent synchronicities aside from the slight overlap between the release of the album (11/2017) and the peak of all cryptocurrencies at the end of 2017, beginning of 2018, which is likely to cause the weak correlation discovered in 3.2.2.3. However, Taylor Swift's musical endeavors are not expected to have an impact on the cryptocurrency markets. This contextual non-relation between the two series thus justifies the inclusion of TS_REDDIT as robustness variable.

To summarize, in the beginning of this chapter the observation period was defined, and all 18 variables were checked for their appropriateness to be representative globally. Then, each variable's source and data retrieval process was shown, followed by the descriptive statistics analysis of the data set. Most interestingly, this showed that the Ripple price series differs significantly if compared to Bitcoin's and Ether's price. However, the three series still share many characteristics, most importantly the peak around December 2017. In regard to the interest proxies, Twitter volume behaved similar to Ripple's price in the way that is also differed significantly from Google and Reddit. Twitter is the only series that experienced another pronounced peak in Q2 2018, long after the cryptocurrencies peaked. Additionally, the activity level kept relatively stable after its peak unlike Reddit and Google, which fell sharply after their peaks.

As for the robustness variables, Wikipedia was confirmed as an adequate substitute for Google, while the number of tweets containing "EPL" expectedly did not correlate with the Bitcoin price. The coinciding release of Taylor Swift's "Reputation" album in November 2017 might be the reason for the weak correlation between the seemingly unrelated topics.

Furthermore, the distribution of each variable showed a skewness, either to the right or to the left caused by outliers. Such distributions could later be smoothed by using logarithmic series instead of the raw data.

The correlation analysis shows that the crypto-markets follow similar paths as the proxies for investor interest. Especially Reddit and Google exhibited a strong correlation with the Bitcoin metrics, while the number of tweets containing 'BTC' unexpectedly showed strong correlations with Ether. One explanation is that this correlation is inherited from the synchronicity between the Bitcoin and the Ether price, as the correlation between Twitter and Ether overlap with the correlation between Twitter and Bitcoin. However, Reddit and Google did not show the same correlation with Ether.

While this correlation analysis was the first step in the analysis of the relationship between cryptocurrencies and the interest of investors for them, it cannot be taken as the answer to the research questions raised earlier. More specifically, the question of the dynamics between the series demands further investigation. However, this investigation is complicated by the fact that none of the variables can clearly be defined as exogenous or endogenous as it cannot be stated by theory, which variable impacts the other. This endogeneity problem is an important factor in the decision for an adequate model in the next chapter. A typical regression model cannot be properly utilized as it would call for a clear definition of an exogenous variable. In such cases, vector autoregressive models have been widely used by researchers. The formulation of these models and their hypotheses will be shown on the subsequent pages.

4 Methodology

After all variables have been defined and described and an initial correlation analysis of selected variables has been conducted, a more in-depth analysis of the variables follows. Correlation analysis is an inadequate tool in the pursuit to analyze the complex dynamics inherit to this data set. Correlation could be spurious and at the same correlation cannot give insight into the influence of the analyzed variables on each other. In the case of cryptocurrencies and their interest variables, the previous chapter concluded that the data set exhibits an endogeneity problem, caused by the lack of knowledge of exogeneous and endogenous variables. Thus, an endogenous model that is the vector autoregressive model is employed to reveal these unexplored relationships.

First, as a conclusion from the previous chapter's descriptive statistics overview, all variables will be transformed logarithmically. Secondly, new variables representing the return volatility of the cryptocurrencies will be introduced. Thirdly, the 21 log variables will be analyzed for the desired stationarity property, where results will expectedly show that most series are non-stationary and must be transformed to first differences. At last, the general VAR model is introduced followed by the specific models for this thesis, which will later be the foundation for testing the previously formulated hypotheses.

4.1 Logarithmic series

As already mentioned in the previous chapter, most data series exhibit a significant skewness, which is mostly caused by outliers. The approach by most researchers is to use the logarithmic series of the raw data to smooth the data and achieve a distribution that is closer to a normal distribution. Thus, all variables are logarithmically transformed, and the prefix 'L' is added to the variable name. On the example of BTC_PRC, the equation for the transformation is simply:

$$LBTC\_PRC_t = \log(BTC\_PRC_t)$$

Equation 4-1 *Logarithmic Transformation*

, where log () is the natural logarithm (also sometimes denoted as $\ln()$).

The descriptive statistics for the modified data set can be found in appendix A.6. As expected, mean and median have moved closer together and the data points are distributed more evenly.

4.2 Return Volatility

In finance, there are typically two measures which are used to characterize an asset: its return and the returns' volatility. Returns are better suited to describe and compare assets' profitability as returns are relative and thus independent of the respective price level. If the price increase of Bitcoin from 5000 to 6000 USD is compared to a price increase of Ether from 100 to 130 USD, it may be misleading, because the absolute difference of Bitcoin is bigger than that from Ether. However, in relative terms, Ether's price increase was higher, because it sore 30% compared to Bitcoin's 20%. Secondly, many financial time series are non-stationary, which is an undesirable property in time series analysis. The calculation of returns has the advantage that in many cases a non-stationary series is made stationary by using first differences, i.e. returns. Thus, for each cryptocurrency's price, a respective return variable will be introduced, which follows the naming convention of adding '_D' (for 'differenced') to the variable names. On the example of LBTC_PRC, the continuously compounded returns are calculated as:

$$LBTC\_PRC\_D_t = LBT\_PRC_t - LBTC\_PRC_{t-1}$$

Equation 4-2 *Return Calculation*

, where LBT_PRC_t is the logarithmic Bitcoin price of the current period and $LBTC\_PRC_{t-1}$ is the logarithmic Bitcoin price of the previous period. Note that these returns are raw returns and not percentage returns. The same is done for LETH_PRC and LXRP_PRC, so that two more variables (LETH_PRC_D, LXRP_PRC_D) are introduced.

Since returns are calculated by differencing the current and the previous price, the return series are created with one missing observation, because for the first observation on February 1st, 2017 no return can be calculated. Thus, all return series have only 606 observations.

	LBTC_PRC_D	LETH_PRC_D	LXRP_PRC_D
Mean	0.003159	0.005078	0.007406
Median	0.004472	0.000737	-0.002981
Maximum	0.177566	0.290130	1.027356
Minimum	-0.154875	-0.315469	-0.616273
Std. Dev.	0.038562	0.066181	0.100094
Skewness	-0.291982	0.315870	2.443719
Kurtosis	5.159437	6.342736	27.13291
Observations	606	606	606

Table 4-1 *Descriptive Statistics for Return series*

As seen in **Table 4-1**, both the average and the median of the three return series is smaller than one percent over the observation period. The highest daily return between February 2017 and October 2018 was 17.76%, 29.01% and 102.74% respectively, while the lowest daily return was -15.49%, -31.55% and -61.63% respectively. Most interestingly, the daily returns of Ripple are significantly more volatile (~ 10%) than both Ether's (~ 6.62%) and Bitcoin's (~ 3.86%), which becomes already apparent in the range between minimum and maximum value of the series.

A 100 USD investment in each of the three cryptocurrencies on February 1st, 2017 would have resulted in 676.21 USD (BTC), 2153.07 USD (ETH) and 8748.77 USD (XRP) on September 30th, 2018. That equals an ROI of 676.21% for Bitcoin, 2153.07% for Ether and 8748.77% for Ripple.

However, these returns come at the expense of risk that investors had to bear in that period. This risk is represented by the volatility of returns. Volatility is usually calculated as the standard deviation of returns over the entire observation period. Thus, the return volatility would be represented by one single value for the sample.

However, in this analysis, one goal is to analyze the changes of investor's interest and their impact on changes of the volatility of return. Since return volatility would be same for the entire sample, there would not be any changes in the volatility. Thus, researchers have used daily, weekly or monthly return volatility, which changes within the observation period, because it is calculated from a subperiod of the observation period.

Daily volatility would require intra-day return data, as more than one return per day must be available for the calculation of the standard deviation. Since these data points are not available for this analysis, the return volatility will be defined as the standard deviation of the previous days' logarithmic returns:

$$LBTC\_RET\_VOLA_t^j = STDEV.S \sum_{i=1}^{j} LBTC\_PRC\_D_{t-i}$$

Equation 4-3 *Return Volatility*

, where t is the time period, j denotes the number of previous days considered in the volatility calculation, $STDEV.S$ represent the sample standard deviation function and $LBTC\_PRC\_D$ are the returns of the logarithmic series $LBTC\_PRC$ as defined in **Equation 4-2**.

Exemplary for Bitcoin, the return volatility for the previous three days would thus be:

$$LBTC\_RET\_VOLA_t^3 = STDEV.S \sum_{i=1}^{3} LBTC\_BTC\_D_{t-i}$$

Equation 4-4 *Three-day Return Volatility*

For regular financial markets, the decision of how many previous days should be included typically falls in favor of five days as markets are open Monday through Friday, equaling five days and thus representing the weekly volatility of the asset. However, as already mentioned, cryptocurrencies can be traded 24/7, which complicates the decision of how many days are best suited to represent the volatility of returns.

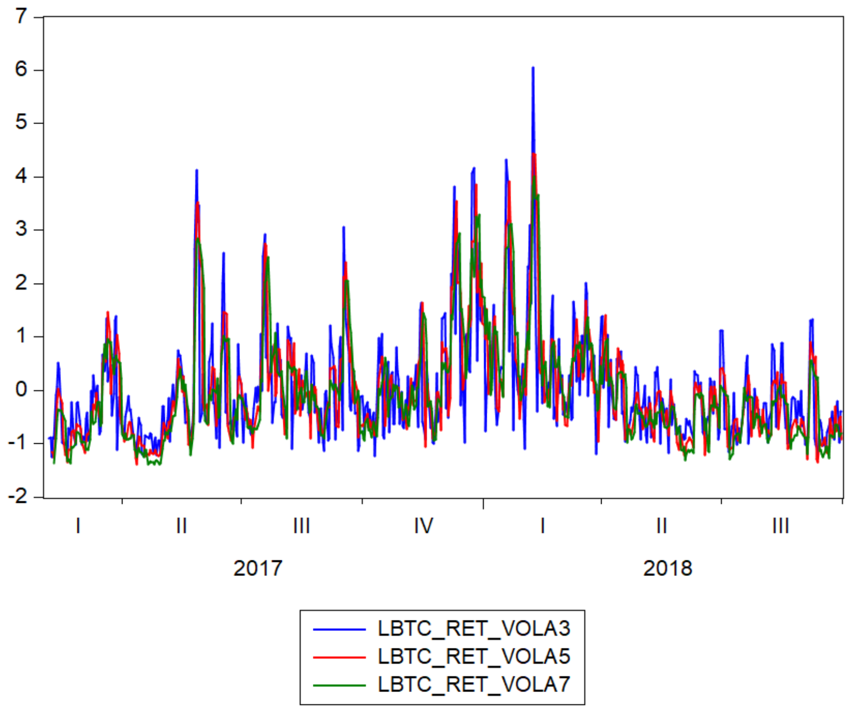


Figure 4-1 *Comparison of 3-, 5- and 7-day Return Volatility*

To find a valid answer to the question of how many days should be considered, three different return volatilities were calculated. Each was calculated as defined in **Equation 4-3**, with the only difference in j, which was set to three, five and seven respectively. A visual comparison is depicted in **Figure 4-1**.

Most notably, LBTC_RET_VOLA3 exhibits significantly more pronounced peaks than the other two volatilities. At the same time, the three-day return volatility's peaks decrease significantly faster after the peak, because peak values are discarded faster compared to the longer-perioded volatilities where a peak value influences the volatility longer. Overall, the more previous returns are considered in the calculation, the more smoothed the series is.

Since the cryptocurrency markets and the interest of investors is assumed to be very fast-paced and quick in its reaction to changes of the markets, the decision falls in favor of LBTC_RET_VOLA3 as it represents the behavior of the markets better than the more smoothed return volatilities.

In addition to LBTC_RET_VOLA3, two additional variables were added to the data set, namely LETH_RET_VOLA3 and LXRP_RET_VOLA3. Again, the return volatility series create missing values, because the first three-day return volatility can be created on February 4th, 2017.

	LBTC_RET_VOLA3	LETH_RET_VOLA3	LXRP_RET_VOLA3
Mean	0.028018	0.052016	0.064891
Median	0.022109	0.042820	0.044405
Maximum	0.155319	0.297768	0.826903
Minimum	0.001367	0.001725	0.000343
Std. Dev.	0.020987	0.039202	0.073706
Skewness	1.938401	2.183587	4.909063
Kurtosis	8.312063	10.93989	42.75083
Observations	604	604	604

***Table 4-2** Descriptive Statistics for Return Volatilities*

Thus, the three series only have 604 observation, three less than the original price series. As shown in **Table 4-2**, returns deviated on average 2.8%, 5.2% and 6.49% respectively. Obviously, the high volatility of Ripple's return is also visible in the three-day volatility of returns. Within three days, volatility was highest at 82.69% on April 4th, 2017 caused by a huge price jump between April 1st and 2nd. Bitcoin's returns were most volatile between February 5th and February 7th, 2018 while Ether's returns deviated most in March 2017.

4.3 Stationarity

As already mentioned in the discussion about returns, stationarity is an important assumption in time series analysis. More specifically, the property of stationarity is desirable for the application of a vector autoregressive model for several reason. First, VARs usually use shocks to demonstrate the influence of one variable's change on another variable. These shocks fade away over time if the series are stationary, as the effect diminishes over time, i.e. the impact on the second variable is smaller in t=5 than in t=1. If, however, the series are non-stationary, then the impact does not decrease and die away eventually but persists and might even increase over time. In these cases, the series is said to contain a 'trend' element. Secondly, statistical significance is not guaranteed in non-stationary series, thus the results of analysis with non-stationary series cannot be interpreted validly.

To test for stationarity, researchers typically conduct two separate tests: The Augmented Dickey-Fuller (ADF) test and the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test. While ADF tests for the existence of a unit root, KPSS analyses the series for stationarity. If a series is non-stationary it contains a unit root, thus the two hypotheses of the tests are contradictory. Results about the (non-)stationarity of a series can only be interpreted significantly if there are different outcomes of the tests. More specifically, the null hypothesis of ADF is that the series contains a unit root while the null hypothesis of KPSS is that the series is stationary. Thus, to label a series stationary, ADF testing must reject the null and KPSS testing must not reject the null. On the other hand, if ADF assumes the null, but KPSS rejects it, the series is non-stationary.

As mentioned before, the desired result is a stationary series. However, if the series is non-stationary, the series can be modified in a way that it is likely to become stationary. The approach for this is to use the first differences of the series, as differencing removes the trend element of the series. This transformation coincides with the calculation of returns as mentioned in 4.2.

Since the three price series have already been transformed, they will be tested for stationarity under the pretense that they already are stationary.

The 15 log variables, the three log return variables and the three return volatility variables have been tested by both ADF and KPSS and an overview of the results can be found in **Table 4-3**. Below the possible outcomes of the two tests are summarized:

(1) Reject ADF null, but do not reject KPSS null → Series is stationary.
(2) Do not reject ADF null but reject KPSS null → Series is non-stationary.
(3) Reject ADF and KPSS null → Stationarity is not clear.
(4) Do not reject ADF and KPSS null → Stationarity is not clear.

In the case of 1. and 2., the directive for the researcher is clear: 1. suggests using the data as is, i.e. levels, while in 2. the series should be differenced as the level series is non-stationary and contains a unit root.

Unfortunately, for some variables both tests reject the null hypothesis (case 3), which signals contradictory results as ADF suggests the series to be stationary while KPSS suggests non-stationarity. Thus, these variables were tested again on the first differences level to see whether they were stationary after differencing. Seven were stationary after differencing (marked with *). For the remaining five variables, no distinctive statement can be made on either levels or first differences level. However, the decision to use first differences for LETH_PRC, LETH_TRADE_ETH and LETH_VOLA was in accordance with the respective variables of Bitcoin and Ripple.

Furthermore, some of the already differenced variables (denoted with *_D) and the three return volatilities, which are calculated based on differenced series, are still not stationary according to KPSS. Differencing first differences, i.e. using the second differences, would make the interpretation of the later results meaningless, thus the variables will be kept at first-difference level despite the results of stationarity testing.

Variable	ADF	KPSS	Conclusion
LBTC_PRC_D	Reject H0	Reject H0	?
LBTC_RET_VOLA3	Reject H0	Reject H0	?
LBTC_TRADE_BTC*	Reject H0	Reject H0	(2)
LBTC_TRADE_USD	Do not reject H0	Reject H0	(2)
LBTC_VOLA	Do not reject H0	Reject H0	(2)
LBTC_GT	Do not reject H0	Reject H0	(2)
LBTC_TWITTER	Do not reject H0	Reject H0	(2)
LBTC_REDDIT	Do not reject H0	Reject H0	(2)
LETH_PRC_D	Reject H0	Reject H0	?
LETH_RET_VOLA3	Reject H0	Do not reject H0	(1)
LETH_TRADE_ETH	Reject H0	Reject H0	?
LETH_TRADE_USD*	Reject H0	Reject H0	(2)
LETH_VOLA	Reject H0	Reject H0	?
LXRP_PRC_D	Reject H0	Do not reject H0	(1)
LXRP_RET_VOLA3	Reject H0	Do not reject H0	(1)
LXRP_TRADE_XRP*	Reject H0	Reject H0	(2)
LXRP_TRADE_USD*	Reject H0	Reject H0	(2)
LXRP_VOLA*	Reject H0	Reject H0	(2)
LBTC_WIKI	Do not reject H0	Reject H0	(2)
LEPL_TWITTER*	Reject H0	Reject H0	(2)
LTS_REDDIT*	Reject H0	Reject H0	(2)

Table 4-3 *Results of Stationarity Analysis*

As a conclusion, new variables were created, which included the additional suffix _D and were calculated the same way as the price return of Bitcoin were calculated in **Equation 4-2**. On the example of LBTC_TRADE_BTC, the equation was formulated as:

$$LBTC\_TRADE\_BTC\_D = LBTC\_TRADE\_BTC_t - LBTC\_TRADE\_BTC_{t-1}$$

Equation 4-5 *First Differences Calculation*

At this point, the data set consisted of 18 variables containing the raw data (as depicted in **Table 3-1**), 18 corresponding log variables (starting with L*) and their 18 first difference pendants (ending with *_D). Additionally, there were three more variables for the three-day return volatility of the price return series of Bitcoin, Ether and Ripple totaling in a data set of 57 variables. Only the 21 variables added last will be included in the empirical analysis as they exhibit the desired properties necessary for the VAR model.

4.4 VAR

Vector autoregressive models have more than one dependent variable compared to regular regression models like OLS. Therefore, there is one equation for each dependent variable containing lagged variables of itself and the other variables in the model. VARs are a generalized version of the univariate autoregressive model, where the model is based on solely one variable and its lags. The advantage of VAR models is that they are non-structural, as they do not require a priori knowledge about the dynamics among the introduced variables, i.e. they are a-theoretical. This solves the previously mentioned endogeneity problem, because the introduction of a variable into the model does not require the knowledge of its exo- or endogeneity beforehand.

A VAR(p) model for p lags is formulated as shown in **Model 4-1**:

$$y_{1t} = \beta_{10} + \sum_{i=1}^{p} \beta_{1i} y_{1\,t-i} + \sum_{i=1}^{p} \alpha_{1i} y_{2\,t-i} + \cdots + \sum_{i=1}^{p} \alpha_{ji} y_{k\,t-i} + \varepsilon_{1t}$$

$$y_{2t} = \beta_{20} + \sum_{i=1}^{p} \beta_{2i} y_{2\,t-i} + \sum_{i=1}^{p} \alpha_{2i} y_{1t-i} + \cdots + \sum_{i=1}^{p} \alpha_{ji} y_{k\,t-i} + \varepsilon_{2t}$$

$$\dots$$

$$y_{kt} = \beta_{k0} + \sum_{i=1}^{p} \beta_{2i} y_{k\,t-i} + \sum_{i=1}^{p} \alpha_{2i} y_{1t-i} + \cdots + \sum_{i=1}^{p} \alpha_{ji} y_{k-1\,t-i} + \varepsilon_{kt}$$

Model 4-1 *General VAR model for p lags*

, where $\beta_0$is the constant term, the first sum operator constitutes the lag term of the variable itself, sum operators 2 through k constitute the lag terms of the other variables in the model. For each variable, there is one equation, thus there are as many equations as variables, i.e. k equations for k variables.

The first model will be focused on solely the interest proxies. The goal is to analyze the dynamics among the interest variables themselves. The hypothesis is that changes in Tweet volume affect changes in the number of Reddit posts and changes in Google search volume and vice versa. Independent of which variable changes first, this change is likely to be transmitted to the other two variables, under the assumption that all three variables serve as valid interest proxies.

$$LBTC\_TWITTER\_D_t = \beta_{10} + \sum_{i=1}^{p} \beta_{1i} LBTC\_TWITTER\_D_{t-i} + \sum_{i=1}^{p} \alpha_{1i} LBTC\_GT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{1i} LBTC\_REDDIT\_D_{t-i} + \varepsilon_{1t}$$

$$LBTC\_REDDIT\_D_t = \beta_{20} + \sum_{i=1}^{p} \beta_{2i} LBTC\_REDDIT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{2i} LBTC\_GT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{2i} LBTC\_TWITTER\_D_{t-i} + \varepsilon_{2t}$$

$$LBTC\_GT\_D_t = \beta_{30} + \sum_{i=1}^{p} \beta_{3i} LBTC\_GT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{3i} LBTC\_REDDIT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{3i} LBTC\_TWITTER\_D_{t-i} + \varepsilon_{3t}$$

Model 4-2 *Model for Interest Variables*

Secondly, in reference to the first hypothesis that Bitcoin is impacted by the changes of interest proxies, **Model 4-2** will be extended to include the different Bitcoin metrics:

$$LBTC\_PRC\_D_t = \beta_{10} + \sum_{i=1}^{p} \beta_{1i} LBTC\_PRC\_D_{t-i} + \sum_{i=1}^{p} \alpha_{1i} LBTC\_GT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{1i} LBTC\_REDDIT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{1i} LBTC\_TWITTER\_D_{t-i}\, \varepsilon_{1t}$$

$$LBTC\_TWITTER\_D_t = \beta_{20} + \sum_{i=1}^{p} \beta_{2i} LBTC\_TWITTER\_D_{t-i} + \sum_{i=1}^{p} \alpha_{2i} LBTC\_GT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{2i} LBTC\_REDDIT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{2i} LBTC\_PRC\_D_{t-i}\, \varepsilon_{2t}$$

$$LBTC\_GT\_D_t = \beta_{30} + \sum_{i=1}^{p} \beta_{3i} LBTC\_GT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{3i} LBTC\_TWITTER\_D_{t-i} + \sum_{i=1}^{p} \alpha_{3i} LBTC\_REDDIT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{3i} LBTC\_PRC\_D_{t-i}\, \varepsilon_{3t}$$

$$LBTC\_REDDIT\_D_t = \beta_{40} + \sum_{i=1}^{p} \beta_{4i} LBTC\_REDDIT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{4i} LBTC\_TWITTER\_D_{t-i} + \sum_{i=1}^{p} \alpha_{4i} LBTC\_GT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{4i} LBTC\_PRC\_D_{t-i}\, \varepsilon_{4t}$$

Model 4-3.A *Model for Bitcoin and Interest Variables*

Here, the three previous interest variables are accompanied by LBTC_PRC_D, the logarithmic changes in Bitcoin price. The added variable leads to one more equation being added to the model. To test for the other four metrics describing the Bitcoin economy, four equivalent models will be analyzed for the variables LBTC_TRADE_BTC_D (4-3.B), LBTC_TRADE_USD_D (4-3.C), LBTC_VOLA_D (4-3.D) and LBTC_RET_VOLA3 (4-3.E).

In the quest to test the second hypothesis that Ether and Ripple are similarly impacted by the interest in Bitcoin, the same model is formulated for Ether and Ripple. On the example of LETH_PRD_D, the model is formulated as:

$$LETH\_PRC\_D_t = \beta_{10} + \sum_{i=1}^{p} \beta_{1i} LETH\_PRC\_D_{t-i} + \sum_{i=1}^{p} \alpha_{1i} LBTC\_GT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{1i} LBTC\_REDDIT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{1i} LBTC\_TWITTER\_D_{t-i}\, \varepsilon_{1t}$$

$$LBTC\_TWITTER\_D_t = \beta_{20} + \sum_{i=1}^{p} \beta_{2i} LBTC\_TWITTER\_D_{t-i} + \sum_{i=1}^{p} \alpha_{2i} LBTC\_GT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{2i} LBTC\_REDDIT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{2i} LETH\_PRC\_D_{t-i}\, \varepsilon_{2t}$$

$$LBTC\_GT\_D_t = \beta_{30} + \sum_{i=1}^{p} \beta_{3i} LBTC\_GT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{3i} LBTC\_TWITTER\_D_{t-i} + \sum_{i=1}^{p} \alpha_{3i} LBTC\_REDDIT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{3i} LETH\_PRC\_D_{t-i}\, \varepsilon_{3t}$$

$$LBTC\_REDDIT\_D_t = \beta_{40} + \sum_{i=1}^{p} \beta_{4i} LBTC\_REDDIT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{4i} LBTC\_TWITTER\_D_{t-i} + \sum_{i=1}^{p} \alpha_{4i} LBTC\_GT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{4i} LETH\_PRC\_D_{t-i}\, \varepsilon_{4t}$$

Model 4-4.A *Model for Ether and Interest Variables*

Note that the interest variables are the same as in **Model 4-3**, because the hypothesis is that interest in Bitcoin – not the interest in the altcoins – impacts altcoins similarly.

Again, this model is modified to test for the effects of the interest variables on the different metrics describing the Ether economy, namely LETH_TRADE_ETH_D (4-4.B), LETH_TRADE_USD_D (4-4.C), LETH_VOLA_D (4-4.D) and LETH_RET_VOLA3 (4-4.E).

Subsequently, the influence of the interest variables will be tested on Ripple. Thus, the initial model is modified for LXRP_PRC_D (4-5.A), LXRP_TRADE_XRP_D (4-5.B), LXRP_TRADE_USD_D (4-5.C), LXRP_VOLA_D (4-5.D) and LXRP_VOLA3 (4-5.E). On the example of LXRP_PRC_D the model is:

$$LXRP\_PRC\_D_t = \beta_{10} + \sum_{i=1}^{p} \beta_{1i} LXRP\_PRC\_D_{t-i} + \sum_{i=1}^{p} \alpha_{1i} LBTC\_GT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{1i} LBTC\_REDDIT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{1i} LBTC\_TWITTER\_D_{t-i}\, \varepsilon_{1t}$$

$$LBTC\_TWITTER\_D_t = \beta_{20} + \sum_{i=1}^{p} \beta_{2i} LBTC\_TWITTER\_D_{t-i} + \sum_{i=1}^{p} \alpha_{2i} LBTC\_GT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{2i} LBTC\_REDDIT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{2i} LXRP\_PRC\_D_{t-i}\, \varepsilon_{2t}$$

$$LBTC\_GT\_D_t = \beta_{30} + \sum_{i=1}^{p} \beta_{3i} LBTC\_GT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{3i} LBTC\_TWITTER\_D_{t-i} + \sum_{i=1}^{p} \alpha_{3i} LBTC\_REDDIT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{3i} LXRP\_PRC\_D_{t-i}\, \varepsilon_{3t}$$

$$LBTC\_REDDIT\_D_t = \beta_{40} + \sum_{i=1}^{p} \beta_{4i} LBTC\_REDDIT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{4i} LBTC\_TWITTER\_D_{t-i} + \sum_{i=1}^{p} \alpha_{4i} LBTC\_GT\_D_{t-i} + \sum_{i=1}^{p} \alpha_{4i} LXRP\_PRC\_D_{t-i}\, \varepsilon_{4t}$$

Model 4-5.A *Model for Ripple and Interest Variables*

To test for the last hypothesis and analyze the dynamics between the cryptocurrencies themselves, another model is formulated:

$$LBTC\_PRC\_D_t = \beta_{10} + \sum_{i=1}^{p} \beta_{1i}\, LBTC\_PRC\_D_{t-i} + \sum_{i=1}^{p} \alpha_{1i}\, LETH\_PRC\_D_{t-i} + \sum_{i=1}^{p} \alpha_{1i}\, LXRP\_PRC\_D_{t-i}$$

$$LBTC\_ETH\_D_t = \beta_{20} + \sum_{i=1}^{p} \beta_{2i}\, LETH\_PRC\_D_{t-i} + \sum_{i=1}^{p} \alpha_{2i}\, LBTC\_PRC\_D_{t-i} + \sum_{i=1}^{p} \alpha_{2i}\, LXRP\_PRC\_D_{t-i}$$

$$LXRP\_PRC\_D_t = \beta_{30} + \sum_{i=1}^{p} \beta_{3i}\, LXRP\_PRC\_D_{t-i} + \sum_{i=1}^{p} \alpha_{3i}\, LETH\_PRC\_D_{t-i} + \sum_{i=1}^{p} \alpha_{3i}\, LBTC\_PRC\_D_{t-i}$$

Model 4-6.A *Model for Cryptocurrencies*

Compared to the other models, **Model 4-6** does not include any interest variables, only lagged variables of the cryptocurrency prices. Thus, the hypothesis for this model is that the cryptocurrencies themselves exert an influence over each other. Bitcoin, as the most prominent cryptocurrency is expected to influence Ether and Ripple based on the assumption that the Bitcoin economy lead investors to other cryptocurrencies. In addition to the prices, the same analysis will be conducted for the trading volumes (LBTC_TRADE_BTC_D (4-6.B), LBTC_TRADE_USD_D (4-6.C)), the price volatility (LBTC_VOLA_D (4-6.D)) and the return volatility (LBTC_RET_VOLA3 (4-6.E)).

As seen in each of the sum operators, the lag length p must be determined before the model can be tested. To determine the adequate lag length, researchers typically rely on information criterion. Compared to the likelihood ratio test, information criteria are better suited for financial data. In short, information criteria determine the appropriate lag length by comparing the RSS value and a penalty term. RSS will decrease if more lags are added while the penalty term will increase with the addition of more lags. At the optimal trade off point, the adequate lag length is then determined. There are several different information criteria such as the Akaike's information criterion (AIC), the Schwarz information criterion (SIC), the Hannan-Quinn criterion (HQC), the final prediction error (FPE) and the Bayesian information criterion (BIC). Ideally, all information criteria would come to the same result, but since all of them are differently constructed, the results are likely to differ. Thus, a 2004 study by Liew analyzed the different criteria for a predetermined sample, where the correct lag length was known. Liew found that FPE and AIC have the highest probability of determining the correct lag length in small samples up to 120 observation, while HQC is best suited for larger samples of more than 120 observations.

The sample size in all models is larger than 600, thus the Hannan-Quinn criterion will be used to determine the appropriate lag length.

In this chapter, as a conclusion from the descriptive statistics analysis in Chapter 3, all variables were logarithmically transformed, concluding in 18 new variables with the prefix L. More importantly, the concept of returns and return volatility was explained and the respective variables were added to the data set. Here, returns were the mere first differences of the log price series of the cryptocurrencies. Based on these returns, return volatilities were calculated using the previous three days' returns.

Furthermore, the stationarity analysis of the variables showed, as expected, that most variables were non-stationary and needed to be transformed to first differences to achieve the desired stationarity property. Some of the variables could still not fulfill the stationarity requirement, but the decision to use second differences was discarded due to interpretability concerns.

At last, a general VAR(p) model was formulated, which was later adapted for the specific hypotheses raised at the end of chapter 2. A total of 21 models were introduced: one to analyze the dynamics of the interest variables themselves, five for each cryptocurrency with the interest variables and five for the analysis of the cryptocurrencies themselves. Finally, to determine the appropriate lag length for each model, literature suggests that the HQ criterion is the proper criterion for samples of the size of this data set.

In the next chapter, the tests for these models will be conducted and their results summarized. On the one hand, the results will be compared to the results of similar literature while on the other hand the outcome will be checked against the hypotheses formulated based on the research question of this thesis.

5 Results

Finally, in this chapter the main hypotheses will be tested with the models introduced in the previous chapter. The hypotheses as formulated in 2.4 are:

1. H_0: Investor interest impacts Bitcoin's metrics.
2. H_0: Investor interest impacts Ripple's and/ or Ether's metrics.
3. H_0: Bitcoin's metrics impact Ether's and Ripple's metrics.

To test the first hypothesis, the hypothesis will be checked against models **Model 4-3**.A through **Model 4-3**.E. The Impulse Response functions will show the response of each Bitcoin metric to a shock of each of the three interest variables. The second hypothesis will be split into two models, as Ether and Ripple will be tested separately in **Model 4-4**.A through **Model 4-4**.E and **Model 4-5**.A through **Model 4-5**.E respectively. Again, the responses of each altcoin metric will be the deciding factor in assuming or rejecting the null.

At last, to test hypothesis three, the responses of the altcoins will be investigated to see whether they are significantly impacted by shocks to the corresponding Bitcoin metric as defined in **Model 4-6**.A through **Model 4-6**.E.

Since one VAR model with p lags and k variables results in $p * k^2$ coefficients, the interpretation of each coefficient becomes tedious. On the example of the first model for the relationship between the interest variables themselves, appendix A.7 shows the output table with all coefficients. Each variable pair shows the coefficient, its standard error and its t-statistic. However, the results of VAR models are usually interpreted with impulse response functions, which visually depict the results in a more comfortable and better interpretable way. One graph shows the response of one variable to a shock of another variable over a selected time period.

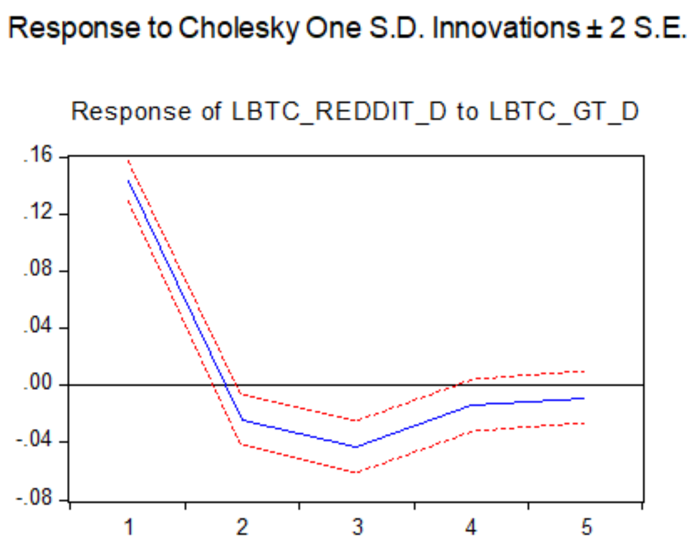


Figure 5-1 *Example of Impulse Response Function*

In reference to **Figure 5-1**, the y-axis shows the magnitude of the response which is corresponding to the coefficient in the output table. The x-axis shows the evolution of the response over time, where one to five represents the periods (days) after the shock. Here, the response of the changes of the log Reddit posts responds positively to a shock to the changes of the log Google search volume in the first period after the shock but becomes negative and then fades away in period two to period five after the shock. The blue line shows the coefficient in each period and is then linked to the successive and previous ones linearly. The red lines around the blue line build the so-called significance band, which constitutes the standard error of the coefficient. Thus, a coefficient can only be interpreted validly if not only the blue line is significant different from zero, but also the significance band. In periods one, two and three, the coefficient is significantly different from zero, while in period four and five, the upper red line crosses the zero line. This means that it cannot be said with significance that the coefficients are different from zero and thus no meaningful interpretation of the two coefficients is granted. To summarize, the y-axis shows the magnitude of a coefficient but can only be interpreted if the red band does not cross the zero line.

5.1 Interest Variables Interdependencies

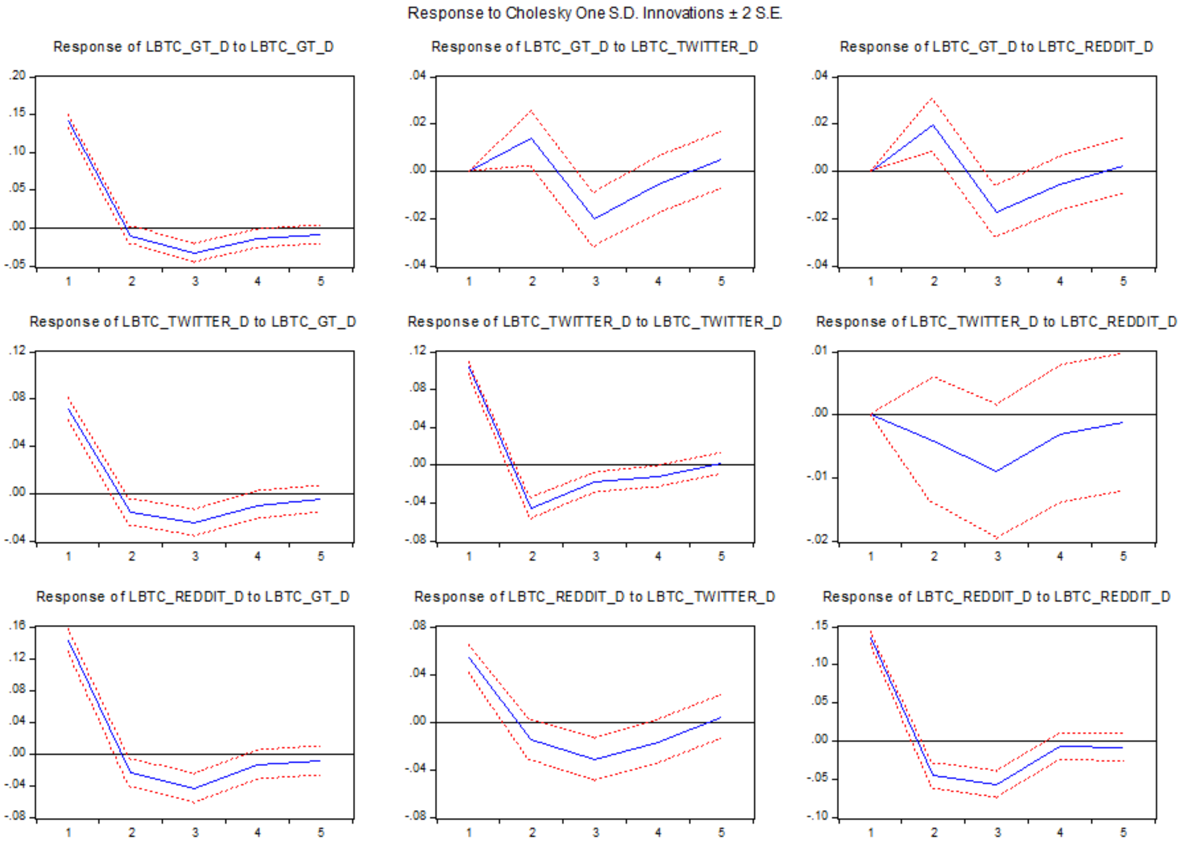


Figure 5-2** Impulse Response Functions for **Model 4-2

Figure 5-2 visually shows the nine responses of the three variables for the interest variables model (**Model 4-2**) as specified in the previous chapter. The lag length for this model was determined via HQC and the criterion suggested a lag length of six periods, i.e. six days.

The diagonal from the top left to the bottom right represents the response of the variable to a shock to the variable itself and will be discarded in future figures as it only depicts the reaction of a variable to changes of itself. These responses are typically significant but have little significance for interpretation as the scope for interpretation lies in the interaction between the different variables and not in the effects of the variable itself. Thus, the following pages' explanation will not focus on the diagonal responses.

In reference to the remaining six response functions, the strongest responses are those of Twitter and Reddit to Google search volume (middle left and bottom left). In the first period, both positively respond to changes in the normalized Google search data. Afterwards, the impact becomes insignificant or minimally negative. Vice versa, the Google variable responds only marginally significant to shock of the two other interest variables (top middle and top right). At last, the bottom middle and the middle right responses suggest that the impact of Twitter on Reddit is minimally positively significant in the first period while Reddit does not exert an influence over Twitter. As such, the relationship between Reddit and Twitter is unidirectional, i.e. shocks to Reddit do not influence Twitter, but Reddit influences the Twitter universe to some extent.

These response functions suggest that changes in the Google search volume are impacted by both the Tweet volume and the number of Reddit posts, while Twitter is not impacted by Reddit. Reddit appears to react to both Google search volume and Tweet volume and exerts only minimal positive impact on Google in the second period. This is relatively surprising as the initial analysis in 3.2.2 led to assume that Twitter would be decoupled from Reddit and Google Trends and not Reddit. However, the conjecture that Twitter and Reddit users react to shocks in Google search volume is reasonable. The suspicion that the Google search volume would change after a shock in the number of Tweets or the number of Reddit posts cannot be confirmed. As such, the relationship between the interest variables is unidirectional from Google Search to Twitter to Reddit posts.

5.2 Bitcoin – Interest Variables

As for the relationship between the Bitcoin economy and the interest variables, **Figure 5-3** through **Figure 5-5** depict the results for **Model 4-3**, where the variables of interest are Bitcoin's price changes, intra-day volatility changes, return volatility changes and trading volume changes (both BTC and USD). The lag length was uniformly determined using HQC at six lags. Originally, there were four responses for each of the four variables, but only the six, which directly show the impact on or of the respective Bitcoin metric are depicted as the others are not relevant regarding the research question.

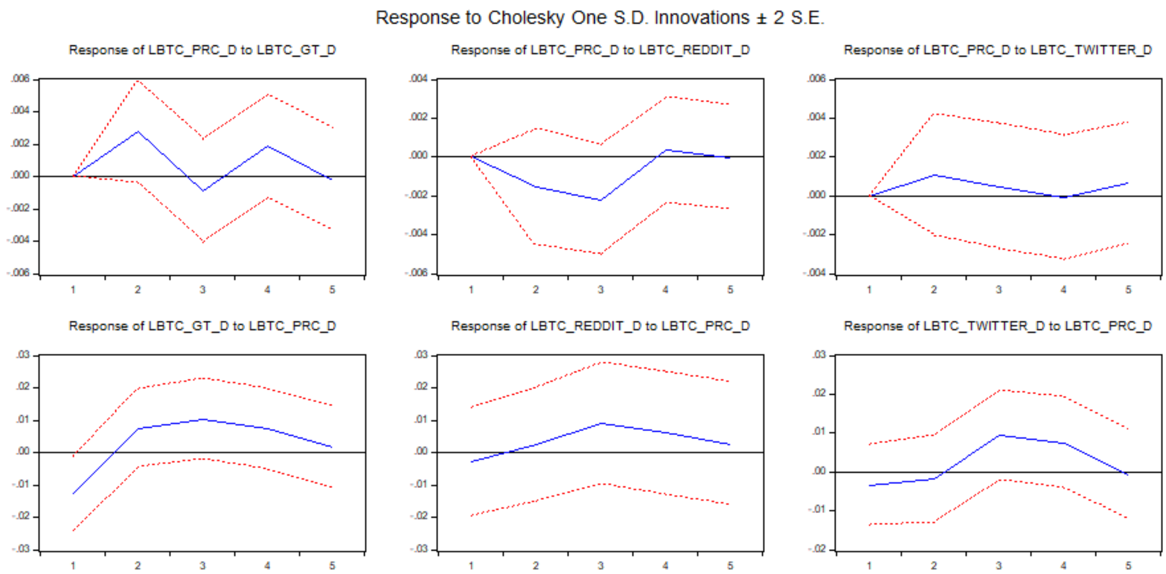


***Figure 5-3** Impulse Response Function for Bitcoin price changes (**Model 4-3**.A)*

As seen in **Figure 5-3**, changes in the Bitcoin price are neither influenced by nor exert an influence over Google, Twitter and Reddit. Except for the first period in the response in the left bottom corner all responses are not significantly different from zero and can thus not be interpreted validly. Interestingly, the one significant result suggests that Bitcoin price changes have a negative impact on the Google search volume. However, this trend seems to reverse after the first period, even though the later periods' results are insignificant again.

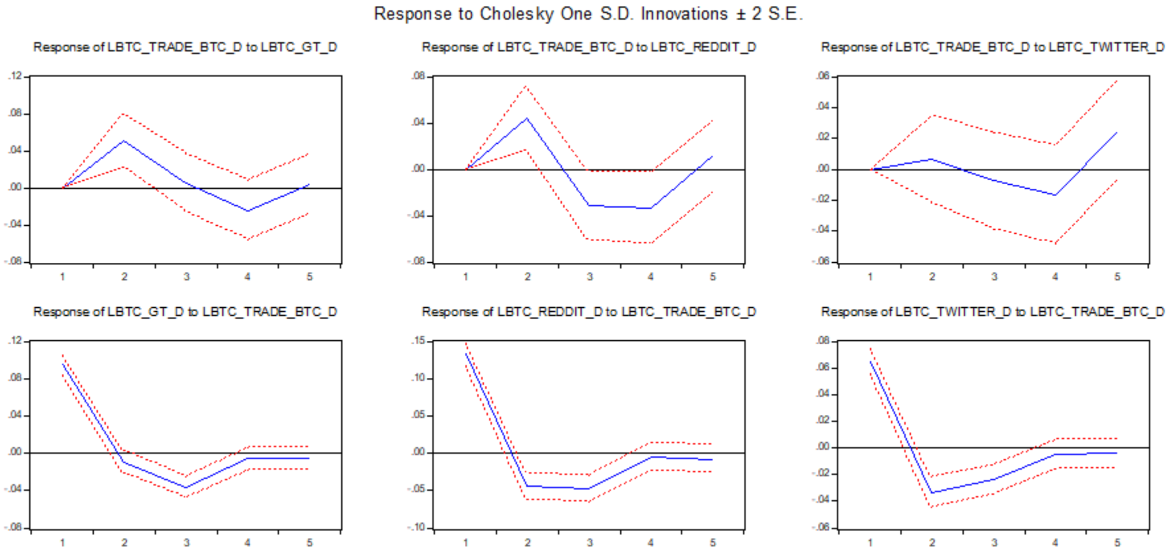


***Figure 5-4** Impulse Response Functions for Bitcoin Trading Volume (BTC) changes (**Model 4-3**.B)*

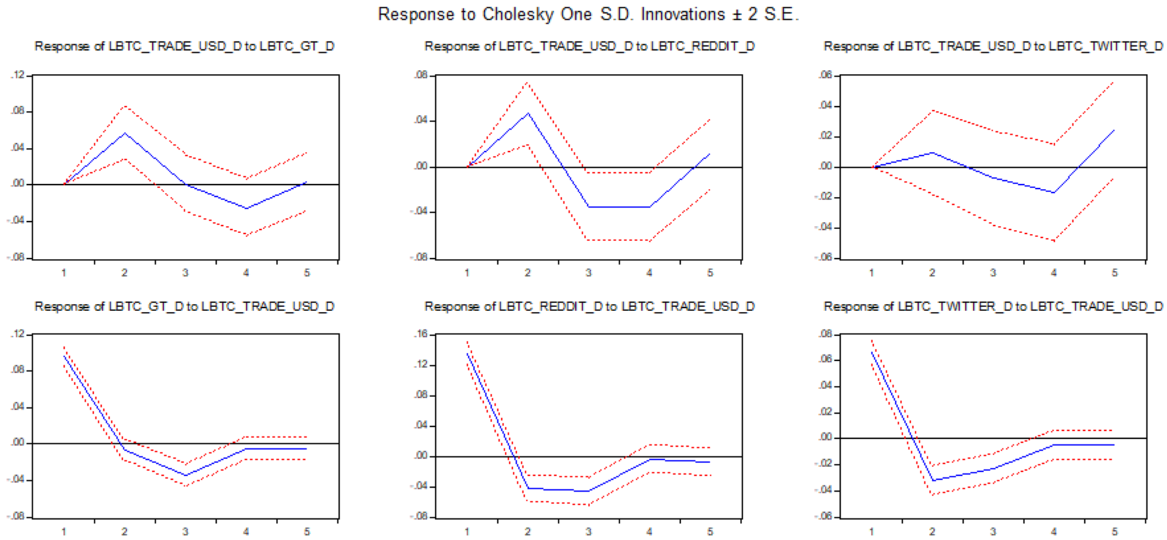


***Figure 5-5** Impulse Response Functions for Bitcoin Trading Volume (USD) changes (**Model 4-3**.C)*

Regarding changes in the Bitcoin trading volume, five out of the six responses show mostly significant results. The bottom three responses exhibit a relatively similar trend: The shock of the trading volume results in a positive response in the first period but turns negative immediately thereafter until it becomes insignificant in period four suggesting that the interest aroused by trading volume changes dies away rapidly in the respective information forums. Vice versa, the trading volume remains unimpacted in the first period after the shock to the interest variables, then responds positively in the second period until the reaction becomes insignificant or minimally negative thereafter.

Interestingly, as seen in **Figure 5-5**, the responses from and to the monetary trading volume are almost identical to the responses of the trading volume denoted in BTC. The interest variables show the same trends and so does the trading volume to shocks of the interest variables. This is surprising at first if the variations in the raw data series are considered but becomes obvious if only the logarithmic changes of the series are observed (see **Figure 5-6**). There, the two series are hard to separate thus the similar responses of the variables are reasonable.

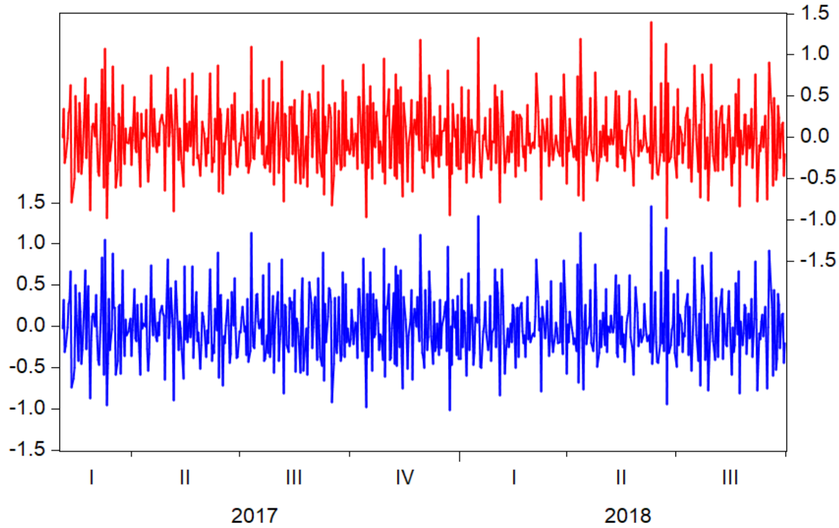


Figure 5-6 *LBTC_TRADE_BTC_D (blue) vs. LBTC_TRADE_USD_D (red)*

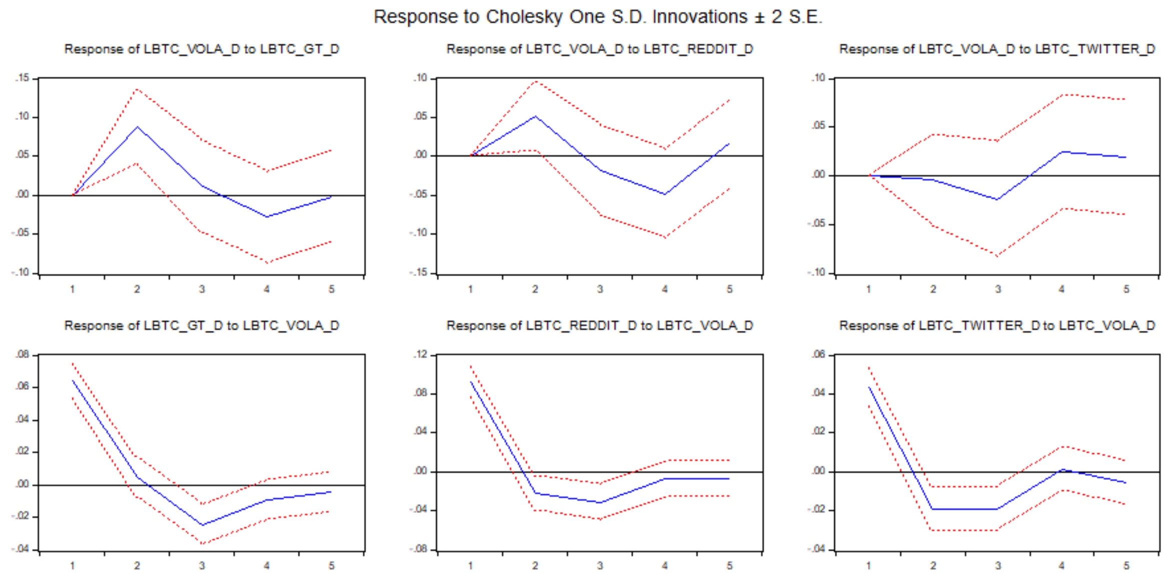


Figure 5-7 *Impulse Response Functions for Bitcoin Intra-day Volatility changes (**Model 4-3**.D)*

Even the intra-day price volatility responses share common characteristics with the two trading volumes (see **Figure 5-7**). The volatility remains robust in the first period after the shock, increases in the second period, but fades away in the following periods. Changes in the Tweet volume do not lead to significant responses of the volatility during any period after the shock. The interest variables, on the other hand, experience an initial increase in the first period, which changes direction afterwards and fade away rapidly.

At last, **Figure 5-8** depicts the responses in respect to the return volatility of Bitcoin. Here, only one graph shows significant results. The three-day return volatility increases constantly throughout period one to three after the shock of Google search volume, remains positive but the positive trend becomes weaker after period three. The other five responses are insignificant throughout all five periods after the respective shocks. To check whether this isolated occurrence is caused by the peculiarity of the Google search variable, the same analysis is conducted again, but with LBTC_WIKI_D instead of LBTC_GT_D.

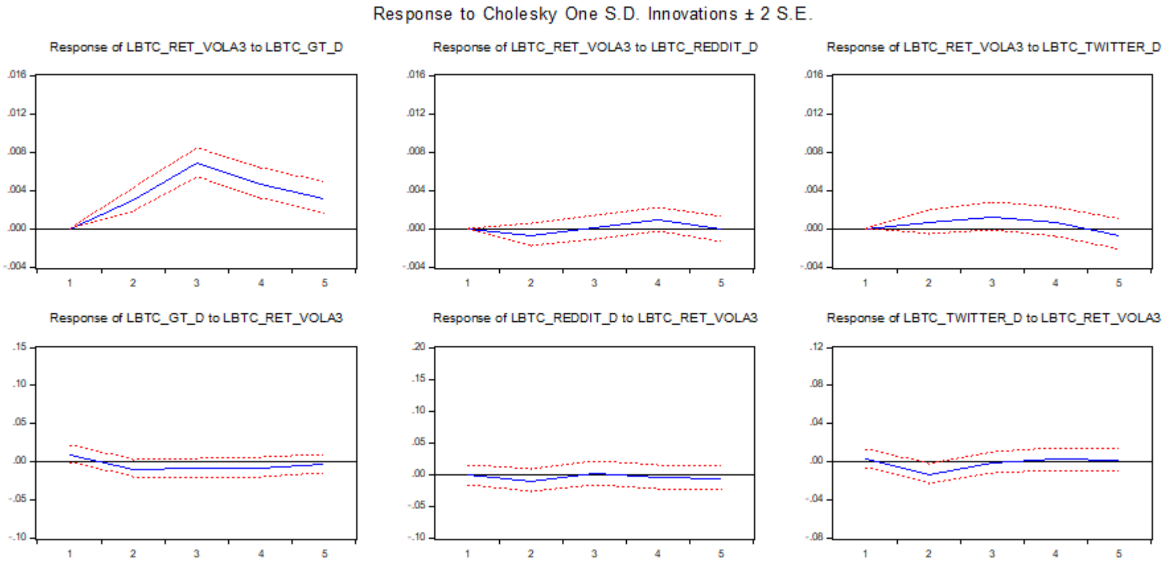


Figure 5-8 *Impulse Response Function for Bitcoin Return Volatility (**Model 4-3**.E)*

The impulse response function for the modified model can be found in appendix A.7. What is interesting is that the response of the return volatility to changes in the number of Wikipedia page views is quasi non-existent, but now the return volatility responds significantly positively to Reddit and Twitter, even though only minimally in scale. At the same time, the reaction of the interest variables to the return volatility remain insignificant. As such, all responses combined suggest that the return volatility responds to changes in the interest variables but not vice versa, signaling a unidirectional impact of the variables compared to the trading volume response where the influences of the variables were bidirectional.

In regard to the first hypothesis, the upper three responses are relevant. The null can only be assumed if the Bitcoin metric reacts significantly to shock of the interest variables. On a very small scale that is true for the price, the trading volumes and the volatilities. However, significance for most of the results is only narrowly reached. More interestingly, there is one trend that is consistent throughout all Bitcoin responses: Bitcoin does not react in the first period after the shock but shows a positive response in the second or third period, compared to the immediate response of the interest variables that fades away immediately afterwards.

Overall, the response of the interest variables to shocks in Bitcoin are more consistently significant than the responses of Bitcoin to shocks in the interest variables, which suggests a unidirectional influence from the Bitcoin economy to the interest variables. The return volatility and the price are exception, which do not influence the interest variables significantly.

5.3 Ether – Interest Variables

Like the Bitcoin price, the responses of the Ether price changes remain insignificant over all five periods (see **Figure 5-9**). Google search volume and Reddit posts only react positively on a significant level in period three and four respectively, but on a scale that is close to neglectable.

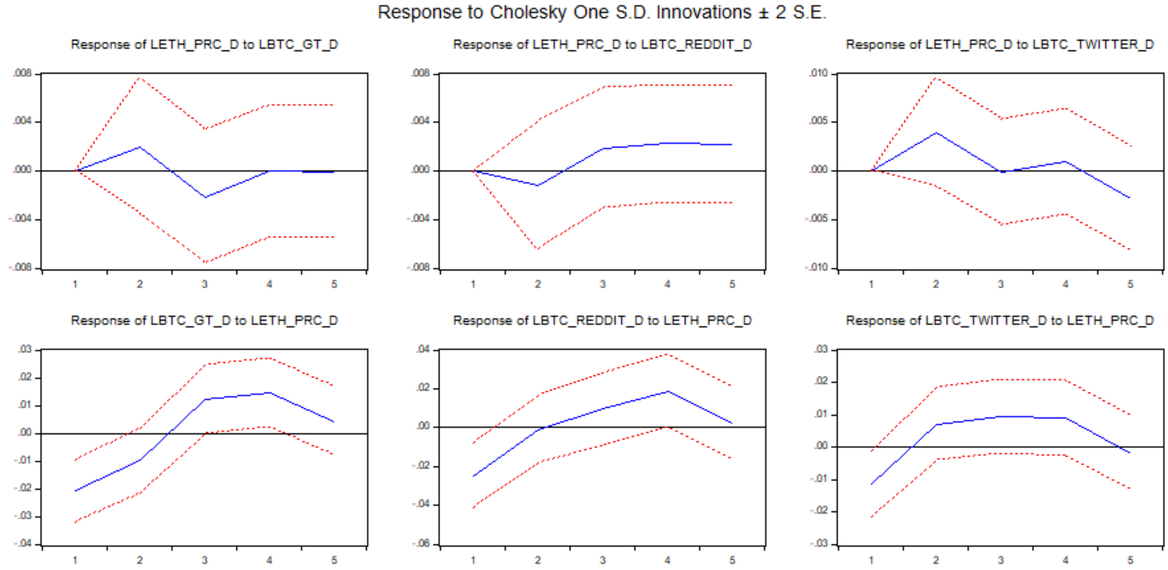


***Figure 5-9** Impulse Response Function for Ether Price changes (**Model 4-4**.A)*

Both Ether trading volumes are not significantly influenced by the interest variables. The interest variables themselves respond very similar to changes in the Ether trading volumes like they responded to the Bitcoin trading volumes. The initial positive impact in period one after the shock, turns negative in the second period and fades away afterwards:

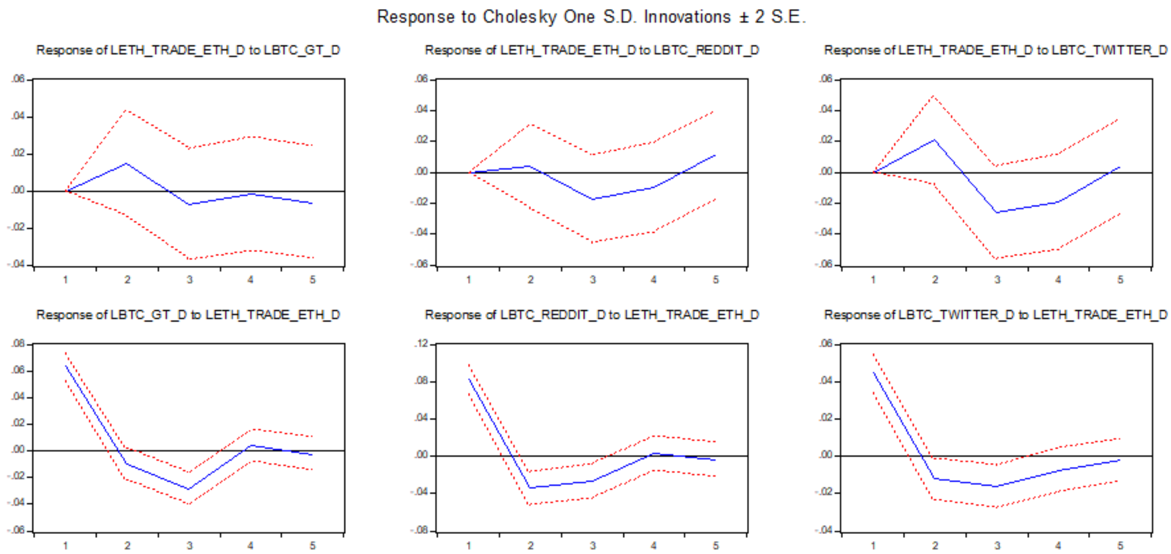


***Figure 5-10** Impulse Response Functions for Ether Trading Volume (ETH) changes (**Model 4-4**.B)*

Since the impulse responses for the monetary trading volume and the intra-day volatility almost mirror the responses as depicted in **Figure 5-10**, they are only shown in appendix A.9. The interpretation falls in line with the implications described when analyzing the Ether trading volume (ETH).

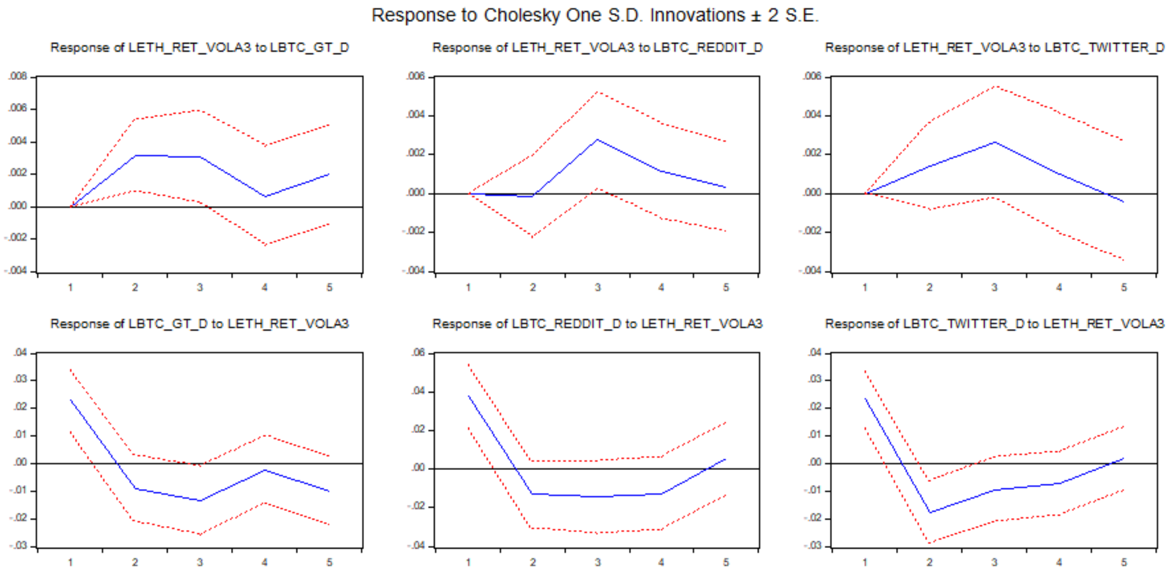


***Figure 5-11** Impulse Response Functions for Ether Return Volatility (**Model 4-4**.E)*

In contrast to the four previous models, the results for the return volatility changes of Ether suggest feedback from the volatility caused by the interest variables. Changes in Google Search volume cause the return volatility to increase over the first three days while volatility increases on the third day if Reddit posts experienced a shocks three days earlier. This trend is also visible in the third return volatility graph (top right), even though the results remain insignificant.

These results differ compared to Bitcoin's return volatility in that Reddit and Twitter did not exert any influence over Bitcoin's return volatility. At the same time, Google searches impacted Bitcoin's return volatility more meaningful. Vice versa, the responses of the interest variables are more pronounced and significant on day one after the shock compared to **Figure 5-8** where the interest variables did not respond to the shock of the Bitcoin's return volatility at all. Overall the dynamics between Ether's return volatility and the interest variables are more pronounced than those of Bitcoin's return volatility. The overall trend that the interest variables are more impacted by the cryptocurrency than vice versa remains visible also for Ether. To verify whether this trend also applies to the third cryptocurrency, the following figures show the impulse responses from Ripple's model.

5.4 Ripple – Interest Variables

Unfortunately, as seen in **Figure 5-12**, all six responses of changes in the Ripple price and the interest variables are mostly insignificant. The interest variables show a zig zag trend that starts negative in the first period, increases on day two, decreases again in period three and rises again afterwards. Most of the coefficients over that time are insignificant but the trend is clearly visible in all three interest variables. Compared to the previously seen positive responses of the interest variables after the shock, the first reaction of them to a shock in the Ripple price is negative and only turns positive in the second period.

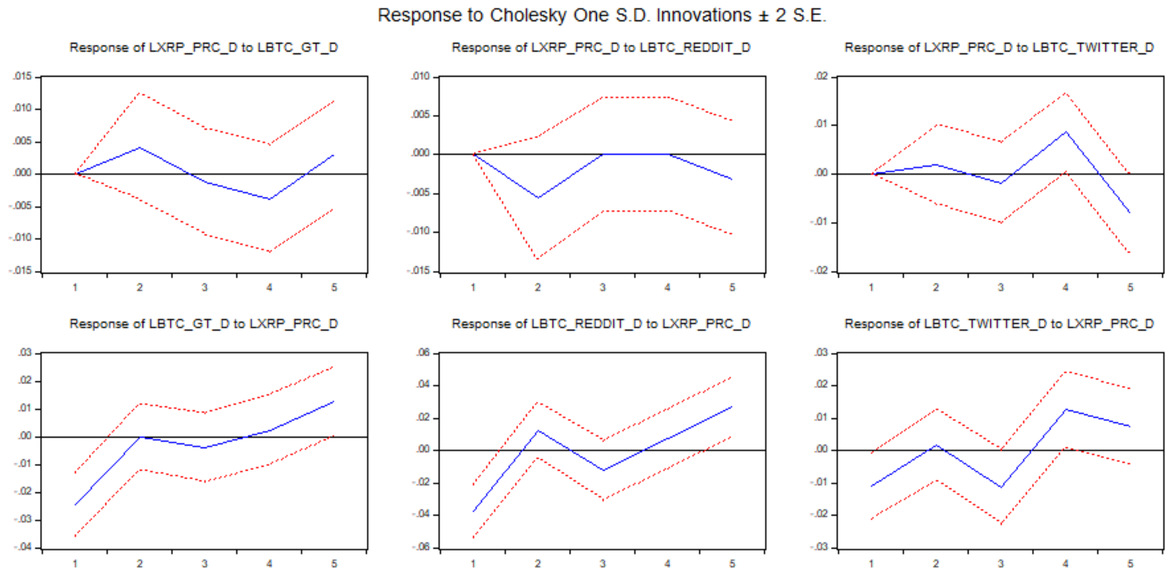


***Figure 5-12** Impulse Response Functions for Ripple price changes (**Model 4-5**.A)*

As already witnessed for Bitcoin and Ether, the two trading volumes respond very similarly, thus only the first trading volume in XRP is depicted in **Figure 5-13**. The monetary trading volume can be found in appendix A.10.

Again, the trading volumes do not significantly respond to changes of the interest variables. Instead the impact on the interest variables is positively significant in the first period after the shock, turns negative in period two and fades away thereafter strengthening the suspicion that interest variables are influenced by the crypto-markets but not vice versa.

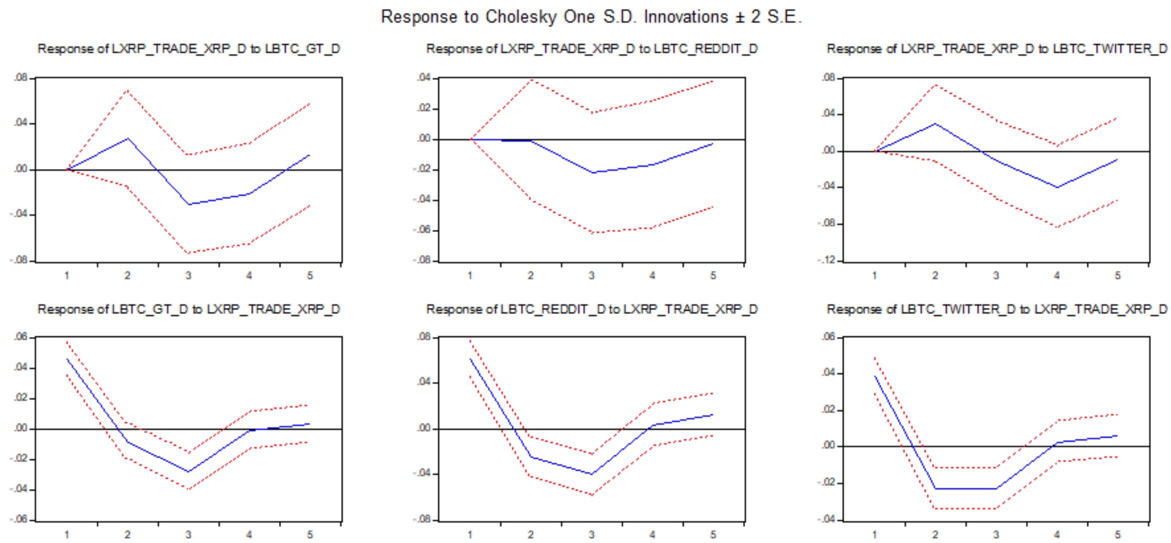


***Figure 5-13** Impulse Response Functions for Ripple Trading Volume changes (XRP) (**Model 4-5**.B)*

This finding also applies to Ripple's intra-day volatility, which can be seen in **Figure 5-14**. Like for Ether, the response of the interest variables to the trading volumes and the intra-day volatility are more or less the same (bottom three graphs in **Figure 5-13** vs. bottom three graphs in **Figure 5-14**).

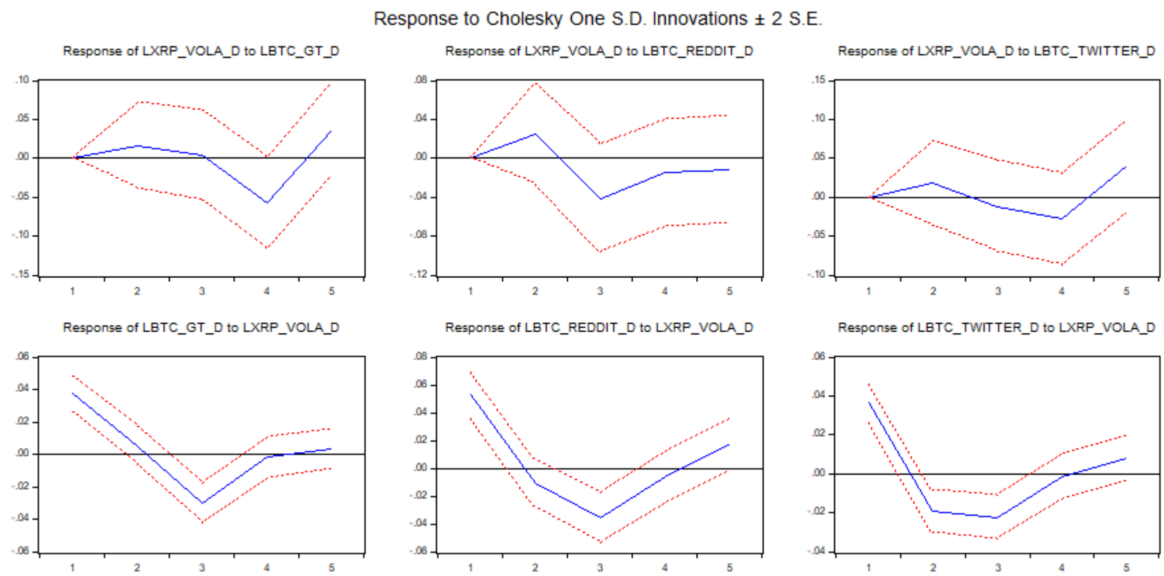


Figure 5-14 *Impulse Response Functions for Ripple Intra-Day Volatility changes (**Model 4-5**.D)*

Ripple's return volatility only responds significantly to changes in Google search volume in the first and in the second period (see **Figure 5-15**). While in the first period, the impact is zero, the second period shows a positive increase in the return volatility that fades away afterwards. However, the graphs suggest that the interaction between Ripple and Google search volume is bidirectional. Three days after the shock to the return volatility, the Google search volume decreases significantly, but the results become insignificant immediately afterwards.

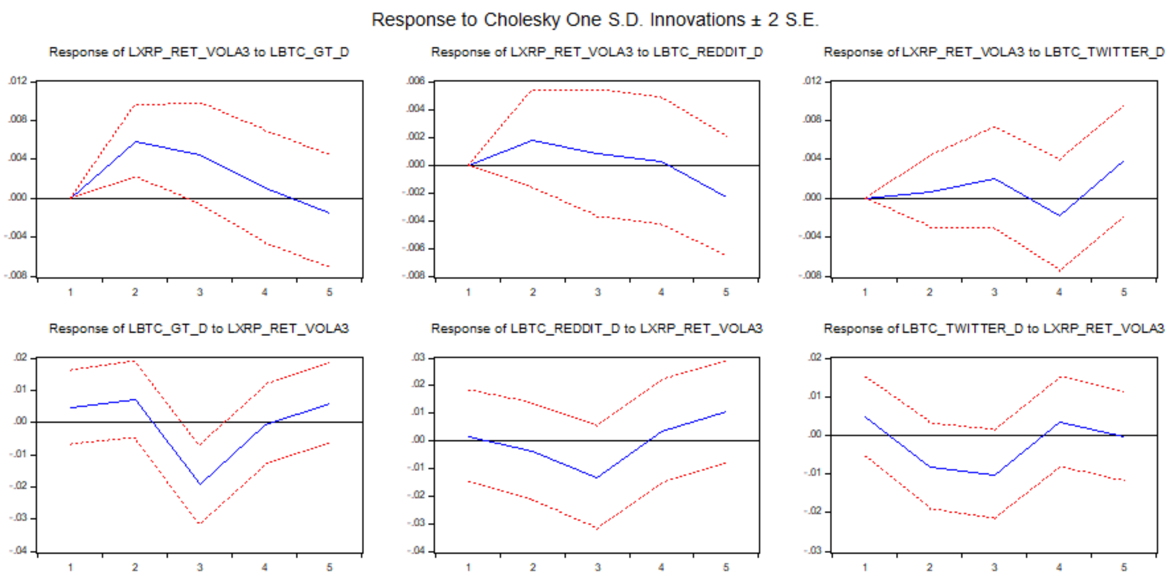


Figure 5-15 *Impulse Response Functions for Ripple Return Volatility (**Model 4-5**.E)*

The second part of the second hypothesis postulated that Ether's and Ripple's metrics are impacted by the interest variables. As seen in 5.3, for Ether this hypothesis cannot be assumed on the basis that all responses combined suggest a reversed influence. Similarly, for Ripple it cannot be said with confidence that the impact of the interest variables is consistently significant. There are too few significant responses in the 15 responses of Ripple that would lead to such conclusion.

After the analysis of all three cryptocurrencies, the most consistent result is the response of the interest variable to shocks in the trading volumes. All six responses show the same trend that is mostly significant. For the price models, the response of the crypto-prices to shocks of the interest variables is throughout insignificant discarding the notion that crypto-prices are impacted by the interest variables. Similar results are found for the return volatility, even though this metric differs across cryptocurrencies. The intra-day volatility model, however, shows a trend of the interest variables that is similar to the trading volumes for all three cryptos.

5.5 Cryptocurrency Interdependencies

Turning now to the results of the last model (**Model 4-6**), where the relationship between the three cryptocurrencies is studied. The lag determination resulted in different lag lengths for each model: HQC suggested only one lag for prices, five lags for trading volume, two lags for the monetary trading volume, three lags for the intra-day volatility and two lags for the return volatility.

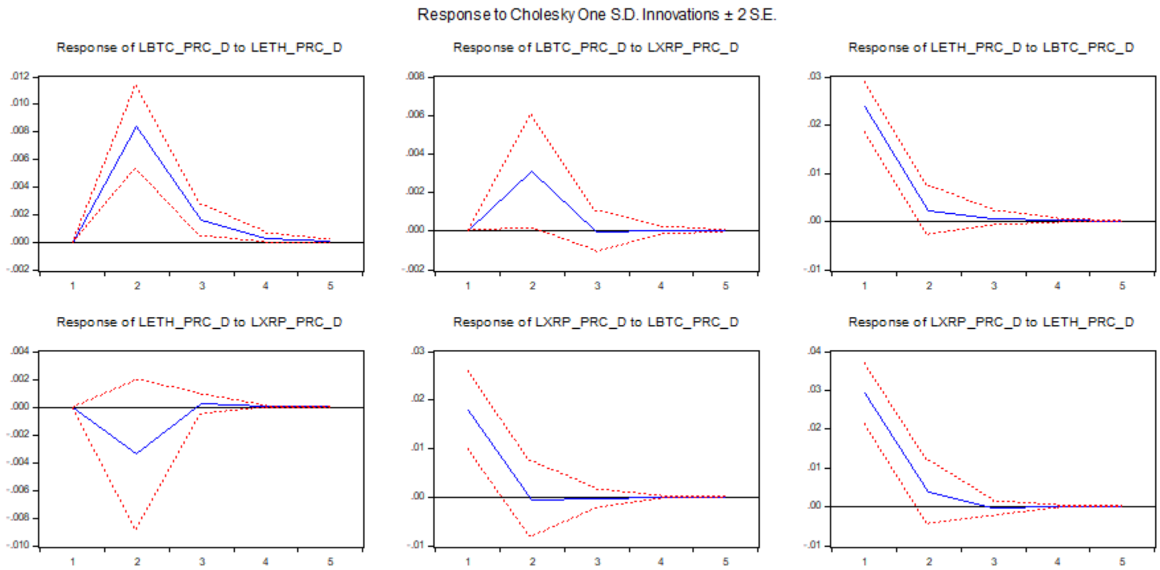


Figure 5-16 *Impulse Response Functions for Cryptocurrency Price changes (**Model 4-6**.A)*

The relationship between the Bitcoin price and the Ether price can be described as bidirectional, because both variables react positively to shocks of the other. However, the response is not equal: While Bitcoin does not change in the first period after the shock and responds only in the second period, Ether is impacted by Bitcoin's price shock immediately. Both price responses fade away in period three, but overall the magnitude of the impact is larger for Ether than for Bitcoin. Similarly, although on a smaller scale, the Bitcoin price reacts to a shock of the Ripple price. Between Ether and Ripple, only the Ripple price responds significantly to the Ether shock, while Ether's response to the Ripple shock is insignificant. At last, the response of Ripple's price to a Bitcoin price shock is similar than Ether's response completing a cycle that starts with Bitcoin and ends with Ripple.

This becomes even more apparent in **Figure 5-17** and **Figure 5-18**, where Bitcoin's trading volumes do not respond significantly to shock of both altcoins. The two altcoins, on the other hand, respond with relatively high magnitude to shocks of Bitcoin's trading volumes. Finally, Ripple's trading volumes respond positively to an abrupt increase in Ether's trading volumes.

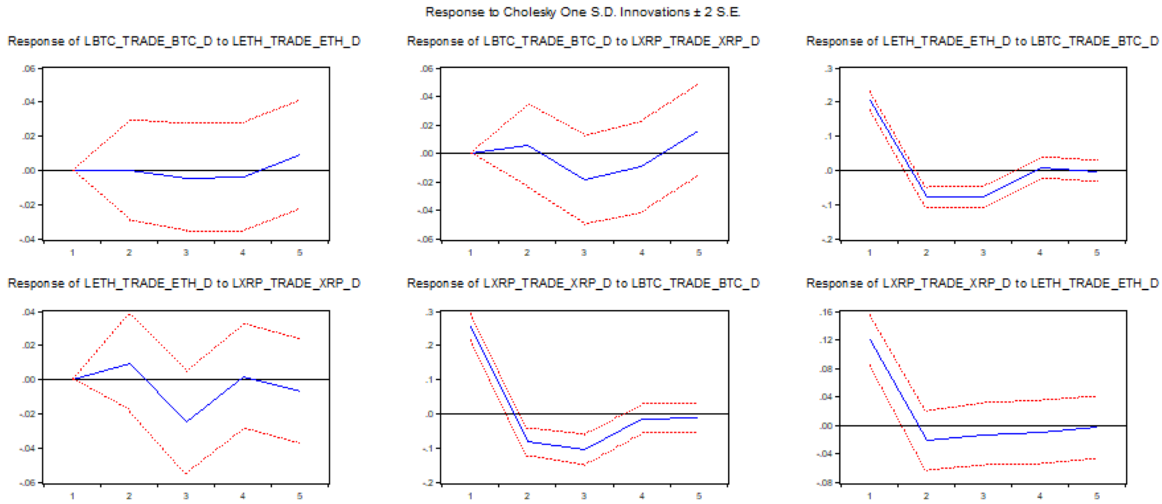


Figure 5-17 *Impulse Response Functions for Cryptocurrency Trade Volume changes (**Model 4-6**.B)*

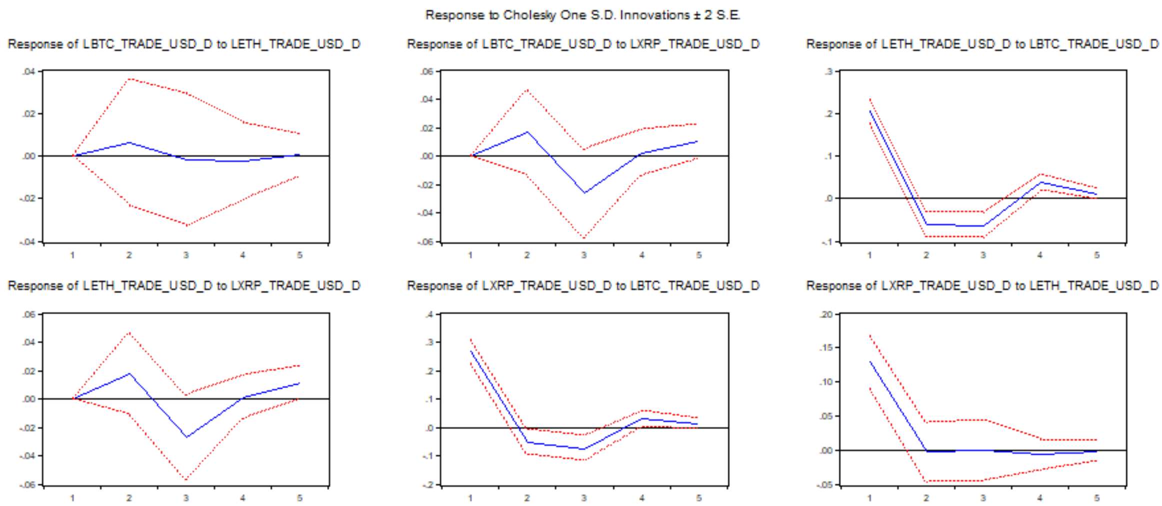


Figure 5-18 *Impulse Response Functions for Cryptocurrency Trade Volume (USD) changes (**Model 4-6**.C)*

All significant responses show the same trend, which is positive in the first period after the shock but becomes slightly negative or insignificant thereafter reminding of the trend of the interest variables to shock in the cryptocurrencies earlier. Very similarly, these trends are observable in **Figure 5-19**: The three significant graphs show the impact of the altcoins triggered by a shock of Bitcoin and Ether. Again, the magnitude is relatively high with 0.3 and 0.2 respectively.

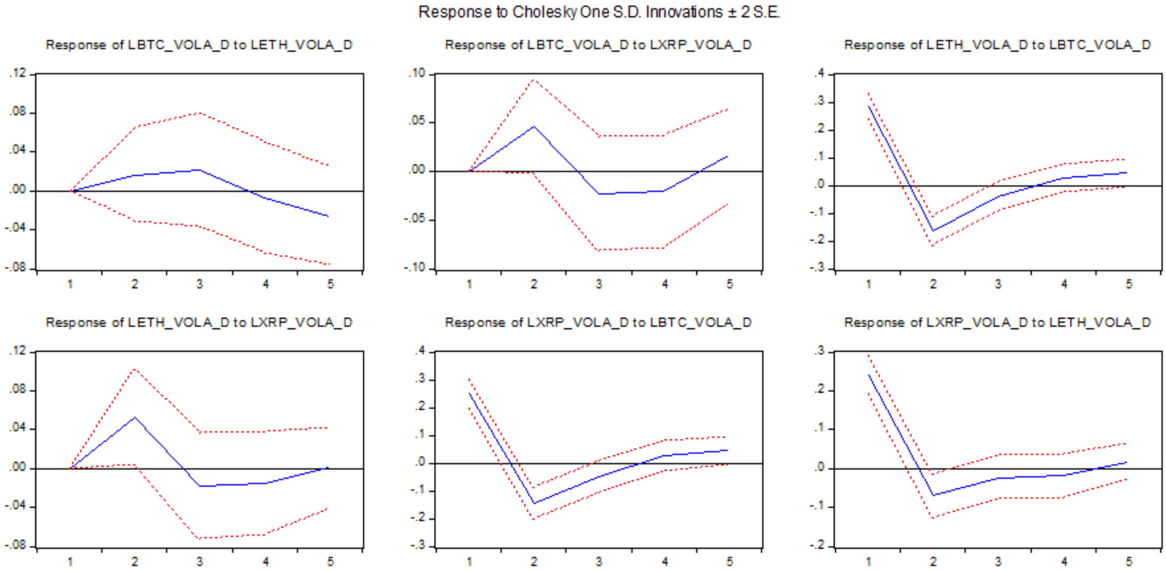


Figure 5-19 *Impulse Response Functions for Cryptocurrency Intra-Day Volatility changes (**Model 4-6**.D)*

Interestingly, this trend does not apply to the return volatilities of the three cryptocurrencies. First, the magnitude of the responses is notably smaller and secondly, Bitcoin's return volatility reacts significantly to shocks of the altcoins' return volatilities. Both Ether and Ripple do not respond significantly to shocks of Bitcoin's return volatility, but to shocks of the respective altcoin (bottom left and bottom right). However, the responses are almost insignificant.

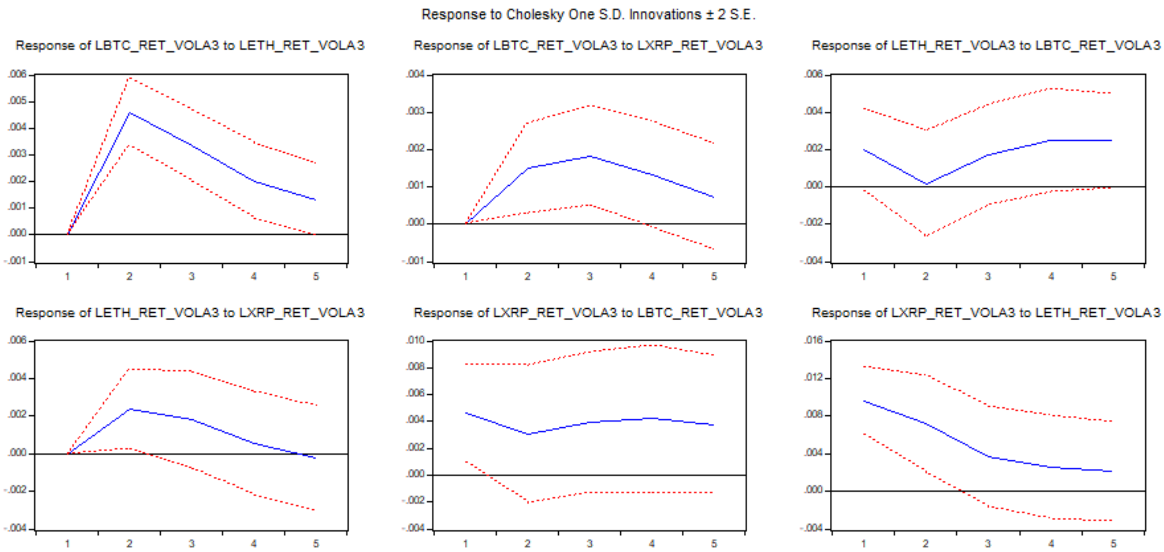


Figure 5-20 *Impulse Response Functions for Cryptocurrency Return Volatility (**Model 4-6**.E)*

All responses combined signal a consistent influence of Bitcoin over the altcoins. Ether and Ripple react positively to shocks of Bitcoin, while Ripple responds to both Bitcoin and Ripple. On a very small but significant level, the return volatility shows a reversed trend where Bitcoin is impacted by the altcoin's return volatility but not vice versa.

5.6 Robustness

In the last step, Bitcoin's price, the trading volumes and the volatilities are tested again but this time with the robustness variables as defined in 3.2.3. The results are expected to show insignificance for the number of Tweets containing 'EPL' and for the number of posts in the subreddit 'TaylorSwift', while the number of page views for the 'Bitcoin' article on Wikipedia should show a similar relationship to the Bitcoin economy as the Google search variable.

The original model's results showed that mostly the interest variables are impacted by Bitcoin, but not vice versa. This remains true also for the robustness variables: As seen in **Figure 5-21**, Wikipedia responds very similarly to what has been observed for LBTC_GT_D in **Figure 5-4**.

Initially the variable increases after the shock, but this trend is reversed afterwards and becomes insignificant in the later periods. In the right column, the response of and to the Reddit robustness variable shows that there is neither a significant response from nor to the variable. Interestingly, the trend, as witnessed in the bottom middle graph, for the Twitter robustness variable is visible in all five impulse response functions (appendix A.11). This trend suggests that the Tweet volume relating to the English Premier League increases after a shock to each of Bitcoin's metrics after three to four periods. However, only for the trading volume this trend is significant, thus the implications should not be overstated.

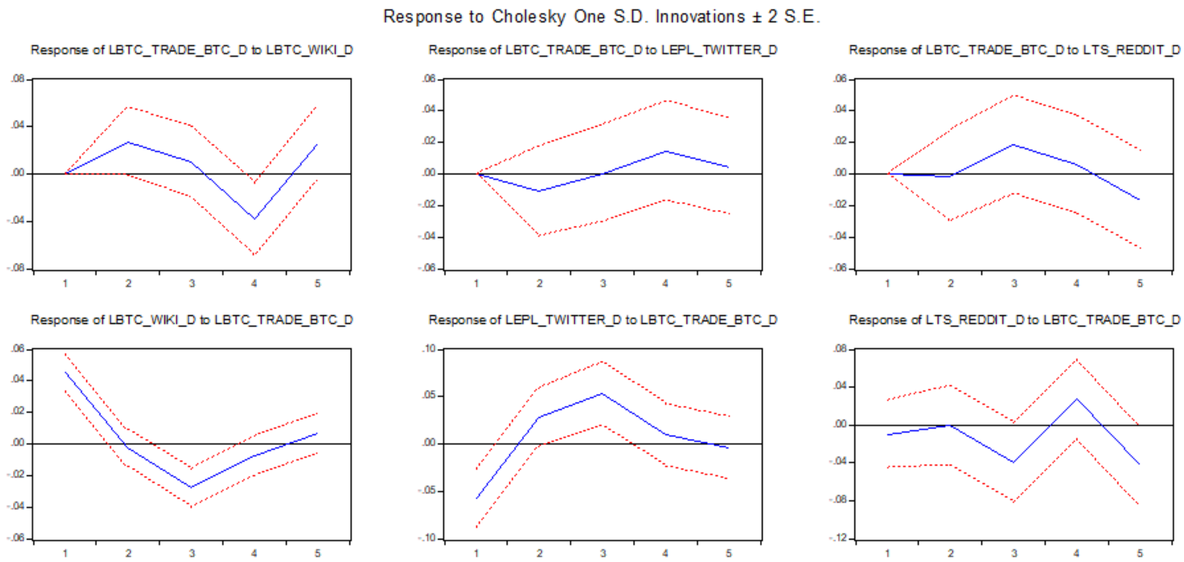


Figure 5-21 *Impulse Response Functions for Trading Volume changes*

Overall the consistent insignificance of the Reddit robustness variable strengthens the validity of the original Reddit variable as it suggests that the results are not random and are indeed related to the theoretical foundation established for the number of Reddit posts. Even if the Twitter robustness variable results were throughout significant, the response of the variable differs notably from the response of the original interest variables in that it reacts negatively at first and becomes positive afterwards. BTC_TWITTER_D and the other two interest variables consistently showed a reversed trend that starts positively and turns negative in the second period. Furthermore, the validity of the original Twitter variable remains intact due to the insignificance of the robustness variable across most of the responses.

To summarize, the first two hypotheses were constructed to test the impact of the interest variables on the cryptocurrencies. The results, however, suggest that these hypotheses must be rejected as it cannot be argued with confidence that the few significant responses of the cryptocurrencies' metric justify this statement. The only consistent trend is visible in the response of the interest variables to shocks of the trading volumes of the cryptocurrencies.

As for the third hypothesis that Bitcoin impacts the two altcoins, four of the five impulse response functions confirm that the altcoins are impacted by Bitcoin. The only exception is the return volatility where the trend is reversed, and Bitcoin does not impact but is impacted by Ether's and Ripple's return volatility. Nevertheless, the majority of responses show a significant response and suggest that the hypothesis is valid and can be assumed concluding in the first two hypotheses being rejected but the third being assumed.

In Kristoufek's 2013 research paper, the responses of the price were also insignificant and only Google and Wikipedia responded significantly. The trend of the reaction of the two interest variables is in accordance with the trends of the three interest variables introduced in this thesis: in the first period after the shock, the variables react positively, but this trend becomes insignificant or minimally negative afterwards and fades away. The author then decided to introduce a dummy variable that indicates whether Bitcoin is above or below a trend line to split the feedback in positive and negative feedback. The reasoning behind this separation is that he argues that it makes a difference whether information search occurs during upwards or downwards trends of the Bitcoin price. After the introduction of the dummy variable, prices respond significantly to positive feedback of Google and to both positive and negative feedback of Wikipedia.

The impulse response functions of Garcia et. al's VAR shows the reaction of the information search variable (Google and Wikipedia), the Word-of-Mouth variable (Twitter) and the user variable to Bitcoin price changes and vice versa. There, the interest variables show a similar trend in their reaction to Bitcoin, which is significantly positive in the first period, but becomes insignificant afterwards. However, prices react significantly to a change in the user variable, which approximates the number of new users in the network.

Most other related literature uses some form of sentiment analysis in their analysis of cryptocurrencies. This makes it difficult to compare the results of this thesis with the respective literature. However, the results that are comparable are in accordance in the way that the responses of the cryptocurrencies to the interest variables are mostly insignificant.

6 Conclusion

Bitcoin has attracted the attention of not only mainstream media, but also the researching community. The mysterious rise of cryptocurrencies has called researchers across the world to explore its characteristics. The question how an asset that does not hold an intrinsic value can achieve a price of close to 20 000 US Dollar has left many fazed, including the author of this thesis.

Economic theory suggests that a price establishes itself at the intersection of the supply and the demand side. In the case of cryptocurrencies, the supply side is known and relies on an algorithm that defines the rate by which new coins are mined. That means that buyers are confronted with a constant supply side that shifts the focus to the decision of buyers. Since these buyers cannot make a reasonable decision based on financial models, the decision must be influenced by something else. Previous research has analyzed whether these influencing factors are other macro-financial indicators, such as the overall trend of the economy, the interest rate or the gold price. The results of these investigations suggest that this is not the case as no significant results could be presented.

If Bitcoin is decoupled from economic factors, the investment decision must be based on the cryptocurrency economy itself and the attention that it attracts. On the example of stocks, researchers found that stock that attract more attention are likely to be more traded than stocks that attract less attention. Applied to the crypto-markets this would mean that the more attractive Bitcoin and Co. becomes to investors, the more likely they are to be bought. However, attractiveness alone does not lead to investment, as investors must not only be aware of the investment opportunity but also be interest to do so. Thus, the goal is to find a measure for the interest of investors for cryptocurrencies. Then, the hypothesis is that the interest of investors translates to buying decision and influences the price of the underlying. This approach has led to respectable results in the stock market, where researchers have built portfolios based on the interest of investors for certain stocks and achieved higher returns than random and equally-weighted portfolios.

In this thesis, the interest of investors was approximated by three different sources: Google, Twitter and Reddit. Google serves an investor with the opportunity to search for everything that is cryptocurrency related, while Twitter gives them the possibility to share information with other investors. Reddit should combine both information-search and information-sharing. Descriptive statistics showed that the three interest variables are strongly correlated with the crypto-markets. More specifically, the relationship between the interest variables and five metrics that describe a cryptocurrency were investigated: the price, its volatility, two different forms of trading volume and the return volatility. Due to the stationarity property necessary for a vector autoregressive model, the series were transformed to first differences, making the raw price series a return series. At the same time, the interpretation of the results changed, because all series were no to be interpreted at first differences level, resulting in shocks to the changes of the series.

The decision to use VAR models instead of OLS models was made because of an endogeneity problem inherit to the data set: A priori, it could not be stated whether the interest variables impact the cryptocurrencies or vice versa. With no clear statement of exogeneous and endogenous variables, no OLS model could be formulated but different VAR models were used instead. First, one model analyzed the dynamics among the interest variables themselves. The results showed that Twitter and Reddit respond positively to shock of the Google variable, while Google itself does not respond on a meaningful level to shock of the other two variables suggesting that Google exerts a unidirectional influence over the other two variables.

Next, five models for the Bitcoin economy were analyzed: The common trend in all impulse response functions is that each metric does only minimally react to shocks of the interest variables. On the other hand, if significant, the response of the interest variables showed a trend that is positive in the first period after the shock, which turns to no impact immediately thereafter. This trend is best observed in the response of the two trading volumes and the intra-day price volatility. The responses of returns and their volatility cannot be interpreted validly, as they are mostly insignificant.

Similar results were found in the second five models for Ether. The interest variables and the intra-day price volatility reacted almost identically to the shock of the two trading volumes. However, the responses for the fifth model, the return volatility of Ether, are different to Bitcoin's return volatility are different in that they are not insignificant over the entire period but signal some impact of the shocks to the return volatility on the interest variables. Similarly, Ripple's impulse response function for the trading volumes and the intra-day volatility exhibit the same trend as already witnessed in Bitcoin and Ether.

The total of 15 impulse response function for all three cryptocurrencies suggest that there is no consistent response of the cryptocurrencies' metrics. To the contrary, the interest variables show a consistent reaction to the cryptocurrencies' metrics leading to the conclusion that the interest variables react to the cryptocurrency markets, but not vice versa.

At last, like for the interest variables, the dynamics among the cryptocurrencies themselves were studies. There, the consistent big picture was that Bitcoin exerts a significant influence over the two altcoins. Furthermore, Ripple responds to both Bitcoin and Ether significantly. This trend is visible in each of the five responses (top right, bottom right and bottom middle responses). Vice versa, Bitcoin does only rarely respond significantly to a shock of either of the altcoins suggesting a unidirectional relationship between Bitcoin and the altcoins.

In the quest to strengthen the validity of the interest variables, robustness tests showed that arbitrary chosen interest variables did not lead to the same responses that were witnessed with the original interest variables. As such, the results from the first 20 variables can be assumed not to be random, which is also visible if the consistency of responses over all models is considered.

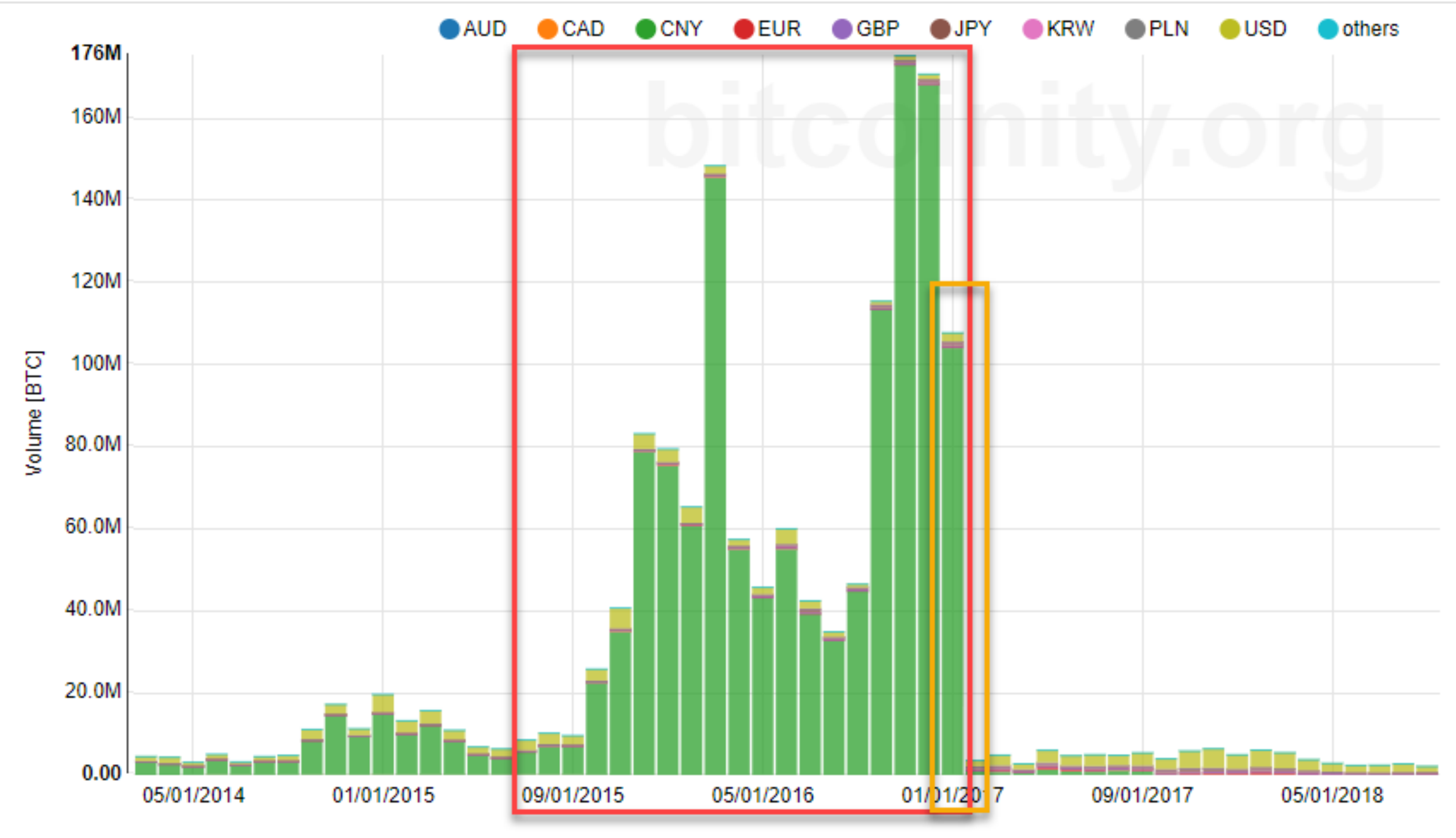
However, the consistent insignificance of the cryptocurrency metrics could be due to the frequency of the data used in the analysis. Daily data might be too infrequent in the pursuit to analyze the relationships in a fast-paced market that is the cryptocurrency markets. Thus, further investigation should consider conducting similar analyses with variables that use intra-day data as they might better capture the dynamics among the different variables. Furthermore, the Google variable constitutes a problem as the normalization of the original search query volume diminishes the substance of the original data. Since it is unlikely that Google will change their approach in how they make their data available, it might be better to use Wikipedia data instead, because it shows the true number of page views. At the same time, similar analyses were able to show more significant results, if the interest of investors was split into positive and negative feedback with the help of sentiment analysis. As such, sentiment analysis of Twitter data and Reddit posts might have benefited the outcome of the empirical analyses.

Additionally, the construction of marginal cost for the production of cryptocurrency might further help in the analysis of the impact of interest variables. The energy cost of mining coins has risen significantly with the increase in crypto trading and might be a significant driver of the prices, as miners must be reimbursed for their efforts. Otherwise, coins would not be mined.

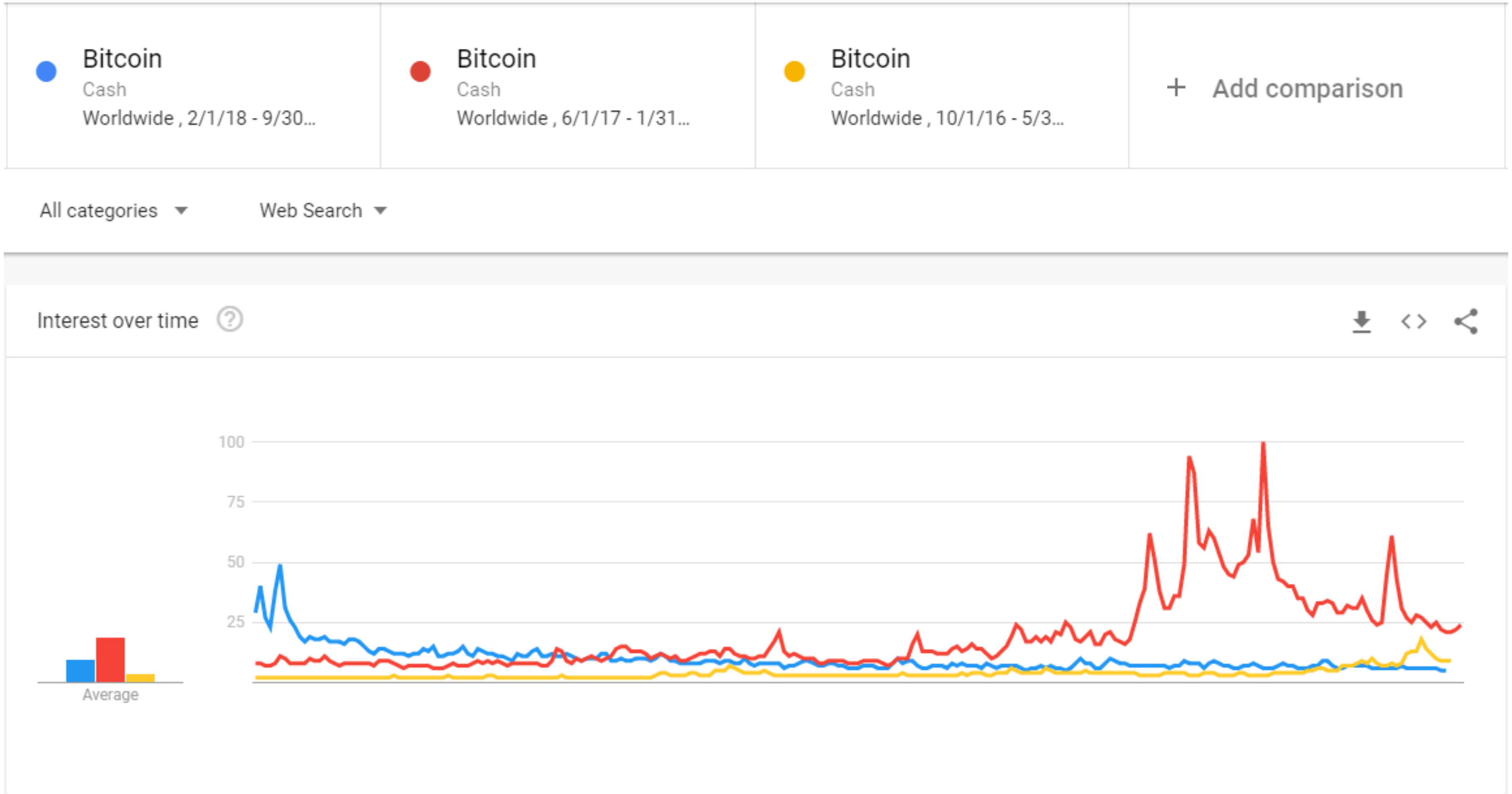
Overall, the cryptocurrency markets continue to amaze researchers and the public alike. Their recentness and disruptive characteristics pose a challenge for the researching community that has yet to be understood and explained. This analysis is no different, as a sound theoretical approach could not be translated to significant results in the subsequent empirical analysis. At the same time, recent developments of cryptocurrency prices have once again amplified the concerns of investors that cryptocurrencies might not exist forever. Their idea of decentralization and anonymity is noble and surfaced in times were banks and centralized institution faced harsh scrutiny. Nevertheless, cryptocurrencies have a long way from being considered an adequate alternative to currencies or a valid investment opportunity for that matter. Time will tell whether cryptocurrencies be able to make their way to everyday usage or if they will be an exciting experiment that will be forgotten soon.

Appendix A: Further Material

A.1 Bitcoin Trading Volume (references 2.1.3.2)



A.2 Google Trends graph (references 3.2.2.1)



A.3 Python script to calculate the daily volatility of prices

```
import csv
import json
import statistics as stat
#start and end time in computer time
start=1538352000
end=1485907200
# one day in computer time to subtract one day
oneday=60*60*24
# create dictionary for date and volatility
voladict={}
#loop for each day
while start>end:
    #difference between end of day (eod) and beginning of day (bod) set to one
day
    eod=start
    bod=start-oneday
    #create dictionary for date and price within one day
    pricedict={}
    #open the .csv file from Bitcoincharts.com's API
    with open(".bitstampUSD.csv", 'r') as inp:
        #check each row in the file for date and if true, write to price diction-
ary
        for row in csv.reader(inp):
            if int(row[0])>bod and int(row[0])<eod:
                pricedict.update({int(row[0]): float(row[1])})
    #calculate the volatility from all entries in price dictionary
    volatility=stat.stdev(pricedict.values())
    #write calculated volatility to volatility dictionary
    voladict.update({int(start): volatility})
    ###print(voladict)
    #reduce start day by one day
    start=start-oneday
# create file from complete volatility dictionary
with open('volatility_daily.txt', 'w') as outfile:
    json.dump(voladict, outfile, indent=1)
```

A.4 Python script to retrieve the number of Reddit posts

```
import requests
import json
#start date in epoch time https://www.epochconverter.com
start=1538352000
#to subtract hours (sec*min*hours)
intervall=60*60*3
#API limits responses to 1000 items, thus reduce time range of request to half
day
end=1485907200
#create dictionary
countdict={
        }
while start > end:
    eod=start
    bod=start-intervall
    url="https://api.pushshift.io/reddit/search/submission/?subred-
dit=bitcoin&filter=id,created_utc&sort=desc&size=10000&before="+str(eod)+'&af-
ter='+str(bod)
    ###print(url)
    response=requests.get(url)
    response=json.loads(response.text)
    #count number of items in response
    post_count=len(response['data'])
    # write counts into dict
    countdict.update({str(eod):post_count})
    #subtract one day
    start=start-intervall
###print(countdict)
#write complete dictionary to file
with open('count.json', 'w') as outfile:
        json.dump(countdict, outfile, indent=1)
```

A.5 Overview of Correlation between cryptocurrencies

	BTC_PRC	ETH_PRC	XRP_PRC	BTC_TRADE_BTC	BTC_TRADE_USD	ETH_TRADE_ETH	ETH_TRADE_USD	ETH_VOLA	BTC_VOLA	XRP_VOLA	XRP_TRADE_USD	XRP_TRADE_XRP
BTC_PRC	1.000000	0.850103	0.764319	0.071990	0.753192	0.188124	0.743193	0.573996	0.734799	0.534665	0.570955	0.312244
ETH_PRC	0.850103	1.000000	0.855432	0.077699	0.612132	0.115921	0.831776	0.694376	0.624492	0.558053	0.583146	0.243701
XRP_PRC	0.764319	0.855432	1.000000	0.021450	0.505857	0.181934	0.801887	0.621932	0.569446	0.764167	0.755926	0.276609
BTC_TRADE_BTC	0.071990	0.077699	0.021450	1.000000	0.619792	0.237535	0.220444	0.451551	0.524746	0.224154	0.204403	0.290927
BTC_TRADE_USD	0.753192	0.612132	0.505857	0.619792	1.000000	0.213337	0.643138	0.710492	0.913935	0.531449	0.548186	0.418440
ETH_TRADE_ETH	0.188124	0.115921	0.181934	0.237535	0.213337	1.000000	0.508756	0.373123	0.198232	0.263345	0.280091	0.358008
ETH_TRADE_USD	0.743193	0.831776	0.801887	0.220444	0.643138	0.508756	1.000000	0.859920	0.646944	0.715057	0.721394	0.423153
ETH_VOLA	0.573996	0.694376	0.621932	0.451551	0.710492	0.373123	0.859920	1.000000	0.722781	0.702875	0.656805	0.438760
BTC_VOLA	0.734799	0.624492	0.569446	0.524746	0.913935	0.198232	0.646944	0.722781	1.000000	0.584905	0.564294	0.364185
XRP_VOLA	0.534665	0.558053	0.764167	0.224154	0.531449	0.263345	0.715057	0.702875	0.584905	1.000000	0.888775	0.517191
XRP_TRADE_USD	0.570955	0.583146	0.755926	0.204403	0.548186	0.280091	0.721394	0.656805	0.564294	0.888775	1.000000	0.673866
XRP_TRADE_XRP	0.312244	0.243701	0.276609	0.290927	0.418440	0.358008	0.423153	0.438760	0.364185	0.517191	0.673866	1.000000

A.6 Descriptive Statistics for Logarithmic Series

	LBTC_GT	LBTC_PRC	LBTC_REDDIT	LBTC_TRADE_BTC	LBTC_TRADE_USD	LBTC_TWITTER	LBTC_VOLA	LBTC_WIKI	LEPL_TWITTER	LETH_PRC	LETH_TRADE_ETH	LETH_TRADE_USD	LETH_VOLA	LTS_REDDIT	LXRP_PRC	LXRP_TRADE_USD	LXRP_TRADE_XRP	LXRP_VOLA
Mean	2.258	8.455	6.059	11.779	20.233	10.525	4.136	11.004	8.840	5.593	14.955	20.544	2.531	2.970	-1.419	18.918	20.339	-4.363
Median	2.197	8.761	5.948	11.793	20.230	10.891	4.132	10.922	8.772	5.775	15.007	21.040	2.726	3.135	-1.257	19.255	20.315	-4.122
Maximum	4.605	9.872	8.341	13.339	22.745	11.871	6.725	13.408	10.759	7.242	16.306	22.944	5.687	5.914	1.218	22.933	22.960	-0.092
Minimum	1.099	6.841	4.828	10.274	17.644	8.925	0.849	9.848	7.725	2.373	13.121	15.555	-2.649	0.000	-5.220	12.348	17.413	-9.914
Std. Dev.	0.690	0.793	0.558	0.493	0.933	0.777	1.143	0.715	0.532	1.109	0.583	1.435	1.378	0.934	1.447	1.971	0.828	1.805
Skewness	0.723	-0.580	0.991	0.090	-0.137	-0.510	-0.217	0.750	0.486	-1.337	-0.404	-1.424	-1.112	-0.352	-1.301	-1.121	-0.009	-0.911
Kurtosis	3.570	2.232	4.395	3.071	2.691	1.826	2.771	3.060	3.046	4.264	3.330	4.765	4.775	3.073	4.259	4.525	3.885	4.134
Obs	607	607	607	607	607	607	607	607	607	607	607	607	607	607	607	607	607	607

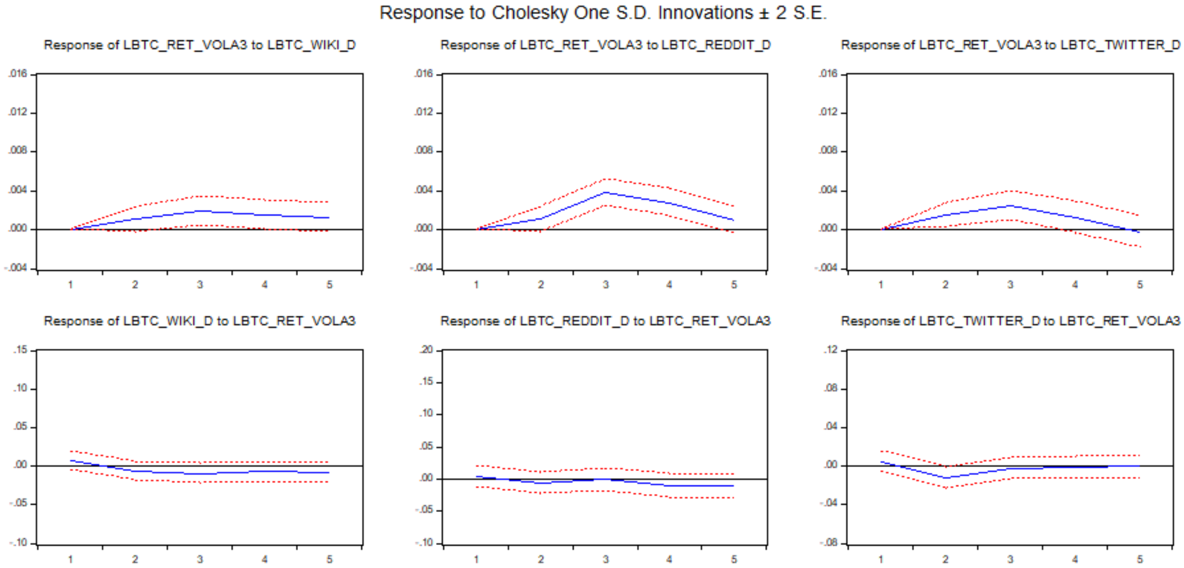
A.7 VAR Output for Model 4-2

Vector Autoregression Estimates
Date: 12/03/18 Time: 11:12
Sample (adjusted): 2/08/2017 9/30/2018
Included observations: 600 after adjustments
Standard errors in ( ) & t-statistics in [ ]

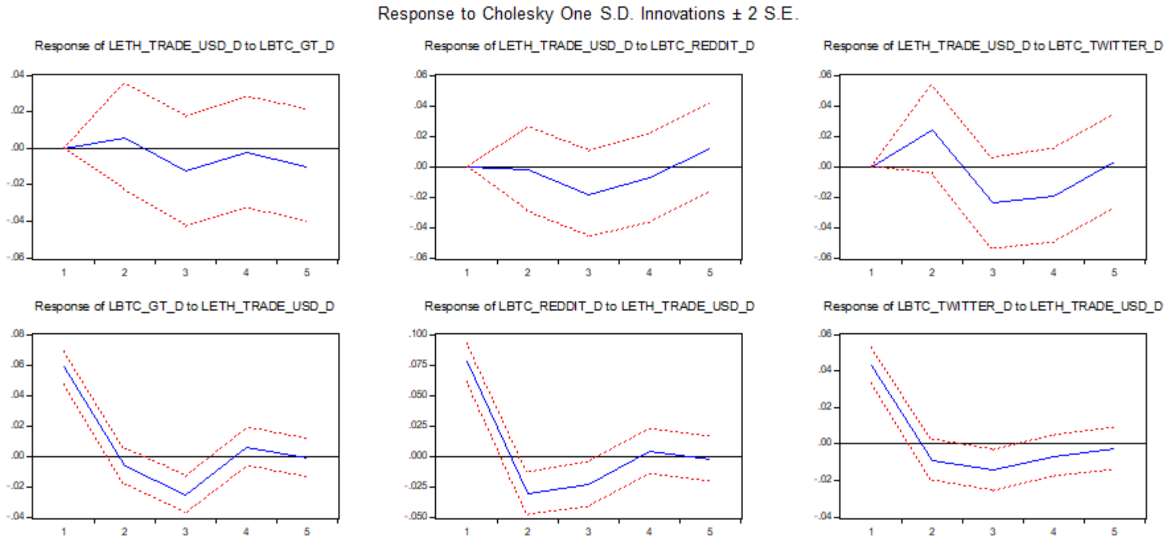
	LBTC_GT_D	LBTC_TWIT-TER_D	LBTC_RED-DIT_D
LBTC_GT_D(-1)	-0.246264	0.137501	0.156880
	(0.05841)	(0.05198)	(0.08426)
	[-4.21634]	[2.64548]	[1.86191]
LBTC_GT_D(-2)	-0.132337	0.062431	0.254557
	(0.05988)	(0.05329)	(0.08638)
	[-2.21008]	[1.17163]	[2.94692]
LBTC_GT_D(-3)	-0.091830	0.081816	0.242690
	(0.05960)	(0.05304)	(0.08598)
	[-1.54072]	[1.54256]	[2.82258]
LBTC_GT_D(-4)	-0.130604	0.059151	0.231270
	(0.05945)	(0.05291)	(0.08576)
	[-2.19680]	[1.11804]	[2.69656]
LBTC_GT_D(-5)	-0.172417	-0.007680	0.033756
	(0.05944)	(0.05289)	(0.08575)
	[-2.90071]	[-0.14519]	[0.39366]
LBTC_GT_D(-6)	-0.067551	-0.019416	0.015656
	(0.05506)	(0.04900)	(0.07943)
	[-1.22679]	[-0.39623]	[0.19709]
LBTC_TWITTER_D(-1)	0.060840	-0.427116	0.028387
	(0.06076)	(0.05407)	(0.08765)
	[1.00130]	[-7.89921]	[0.32386]
LBTC_TWITTER_D(-2)	-0.094613	-0.332831	-0.078084
	(0.06591)	(0.05865)	(0.09508)
	[-1.43556]	[-5.67488]	[-0.82127]
LBTC_TWITTER_D(-3)	-0.077532	-0.294156	-0.176294
	(0.06720)	(0.05980)	(0.09694)
	[-1.15376]	[-4.91895]	[-1.81855]
LBTC_TWITTER_D(-4)	-0.006773	-0.193522	-0.065345
	(0.06696)	(0.05959)	(0.09660)
	[-0.10114]	[-3.24759]	[-0.67645]
LBTC_TWITTER_D(-5)	0.044051	-0.072211	0.066619
	(0.06521)	(0.05803)	(0.09408)
	[0.67548]	[-1.24428]	[0.70812]
LBTC_TWITTER_D(-6)	0.052101	-0.078455	-0.008653
	(0.05983)	(0.05324)	(0.08630)
	[0.87088]	[-1.47367]	[-0.10027]
LBTC_REDDIT_D(-1)	0.144537	-0.030149	-0.340786
	(0.04154)	(0.03696)	(0.05992)
	[3.47984]	[-0.81567]	[-5.68747]

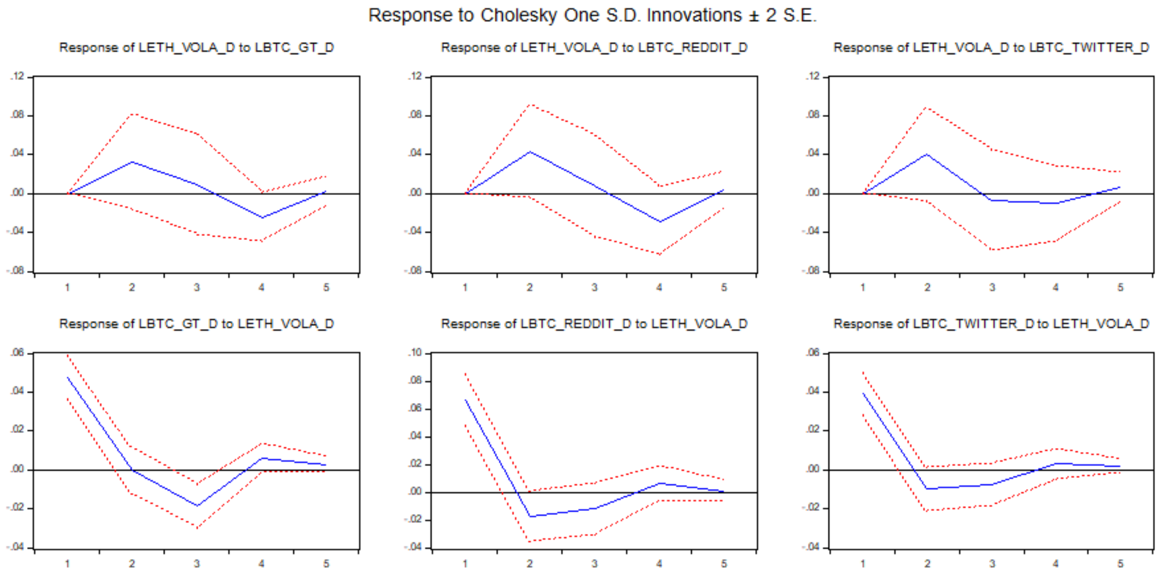
LBTC_REDDIT_D(-2)	-0.039844	-0.110131	-0.562546
	(0.04289)	(0.03817)	(0.06188)
	[-0.92894]	[-2.88530]	[-9.09143]
LBTC_REDDIT_D(-3)	-0.002706	-0.104228	-0.402571
	(0.04447)	(0.03958)	(0.06416)
	[-0.06085]	[-2.63363]	[-6.27490]
LBTC_REDDIT_D(-4)	-0.014571	-0.132555	-0.470752
	(0.04440)	(0.03951)	(0.06405)
	[-0.32818]	[-3.35488]	[-7.34966]
LBTC_REDDIT_D(-5)	-0.038658	-0.172677	-0.469248
	(0.04305)	(0.03831)	(0.06210)
	[-0.89806]	[-4.50780]	[-7.55662]
LBTC_REDDIT_D(-6)	0.007042	-0.034290	-0.242729
	(0.04222)	(0.03757)	(0.06090)
	[0.16681]	[-0.91272]	[-3.98555]
C	0.001886	0.007470	0.001085
	(0.00582)	(0.00518)	(0.00839)
	[0.32402]	[1.44245]	[0.12923]
R-squared	0.163622	0.256383	0.312059
Adj. R-squared	0.137710	0.233345	0.290746
Sum sq. resids	11.60798	9.192452	24.15710
S.E. equation	0.141348	0.125785	0.203908
F-statistic	6.314545	11.12866	14.64164
Log likelihood	332.2080	402.2010	112.3422
Akaike AIC	-1.044027	-1.277337	-0.311141
Schwarz SC	-0.904790	-1.138100	-0.171905
Mean dependent	0.000851	0.002902	-8.45E-05
S.D. dependent	0.152217	0.143657	0.242122
Determinant resid covariance (dof adj.)		3.90E-06	
Determinant resid covariance		3.54E-06	
Log likelihood		1211.241	
Akaike information criterion		-3.847471	
Schwarz criterion		-3.429762	

A.8 Impulse Response Functions with Wikipedia instead of Google

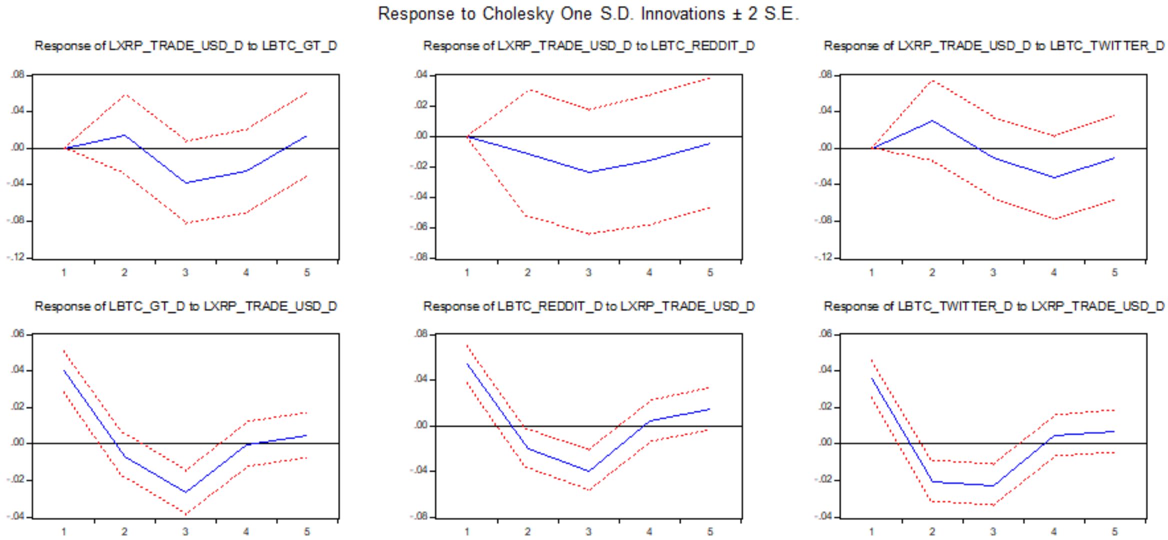


A.9 Impulse Response Functions for Ether Monetary Trading Volume

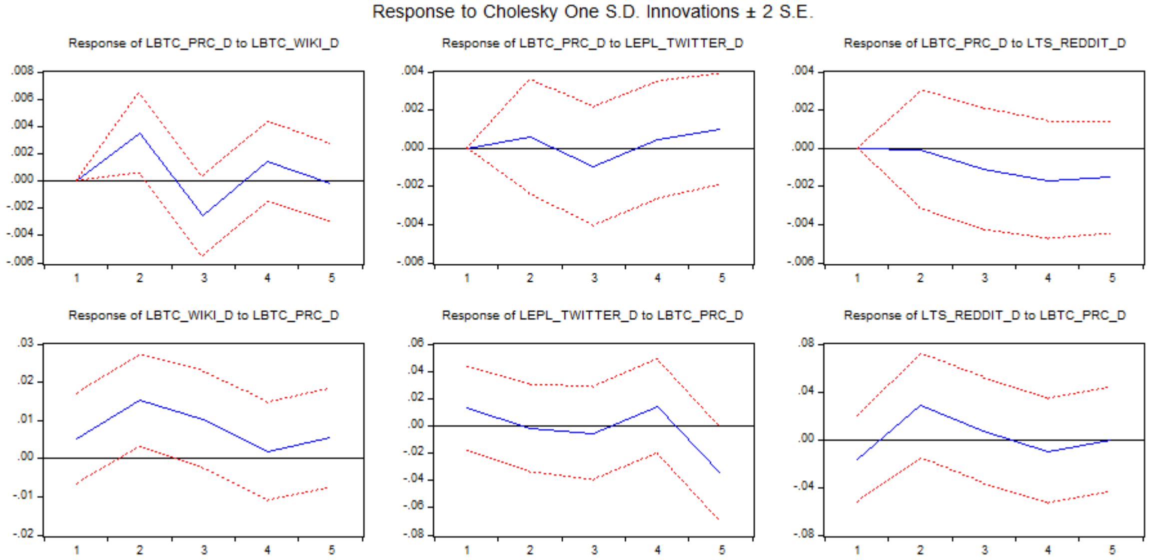


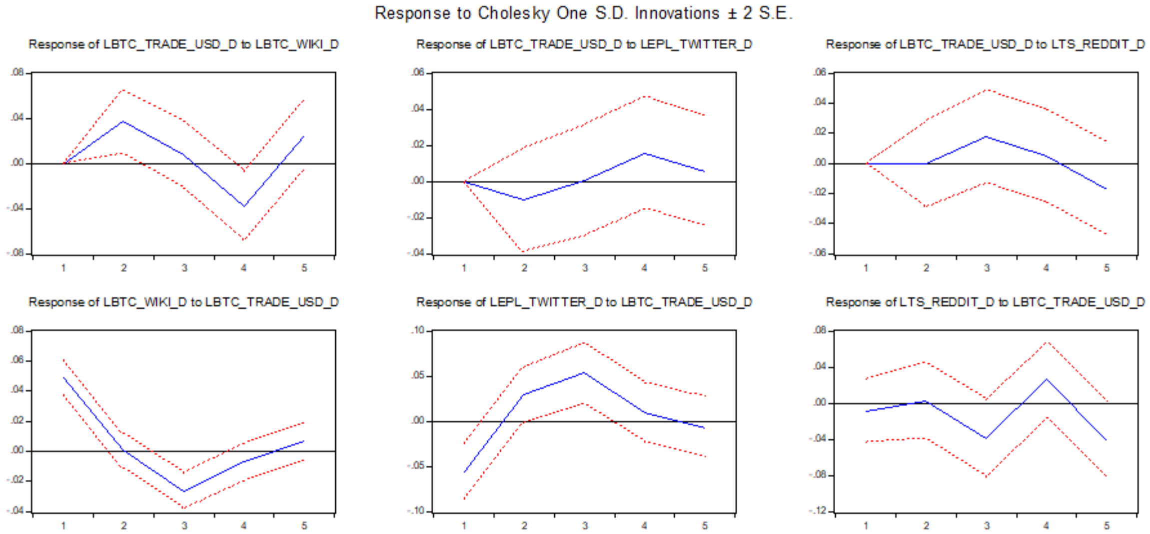


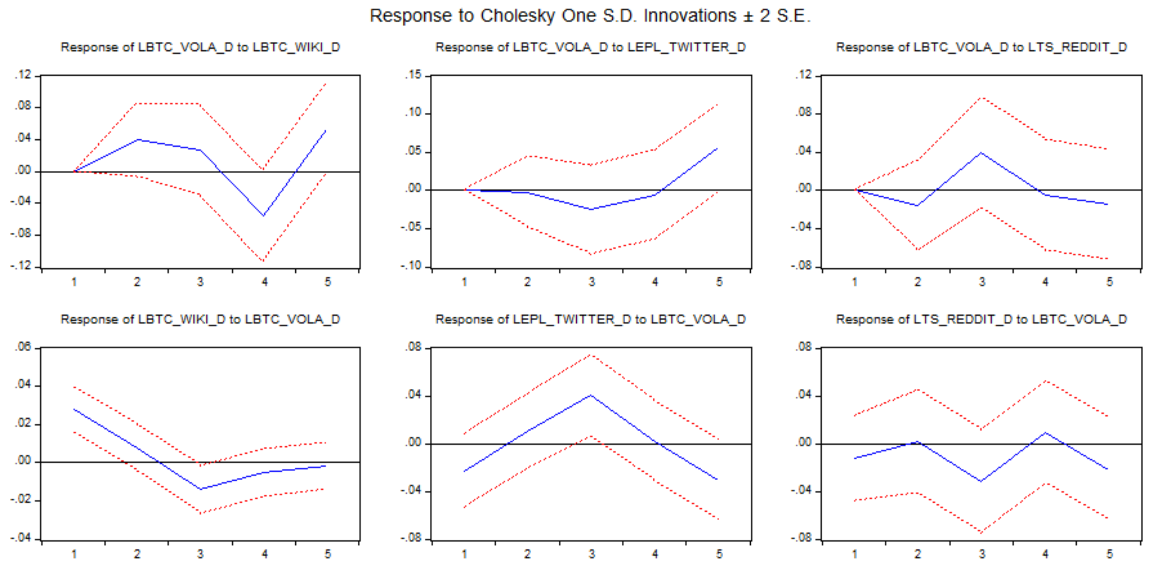
A.10 Impulse Response Functions for Ripple Monetary Trading Volume

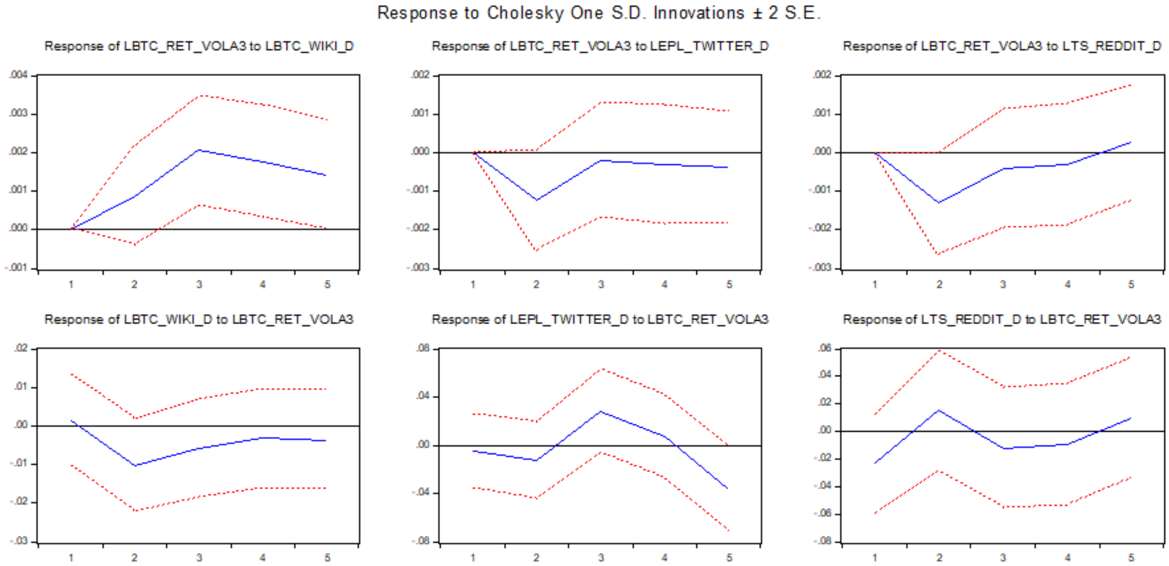


A.11 Impulse Response Functions for Robustness Variables








Response to Cholesky One S.D. Innovations ± 2 S.E.
Response of LBTC_RET_VOLA3 to LBTC_WIKI_D
Response of LBTC_RET_VOLA3 to LEPL_TWITTER_D
Response of LBTC_RET_VOLA3 to LTS_REDDIT_D
Response of LBTC_WIKI_D to LBTC_RET_VOLA3
Response of LEPL_TWITTER_D to LBTC_RET_VOLA3
Response of LTS_REDDIT_D to LBTC_RET_VOLA3

7 References

Alexa Internet Inc. (2018a). Top Sites in United States. Retrieved from https://www.alexa.com/topsites/countries/US

Alexa Internet Inc. (2018b). The top 500 sites on the web. Retrieved from https://www.alexa.com/topsites

Arrington, M. (2010, March 22). Google Shuts Down Chinese Search, Redirects To Hong Kong. *TechCrunch*. Retrieved from https://techcrunch.com/2010/03/22/google-shuts-down-chinese-search-redirects-to-hong-kong/

Barber, B. M., & Odean, T. (2007). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *The Review of Financial Studies*, *21*, 785–818.

Basu, A. (2017, May 12). March to the title: how Chelsea's season unfolded, game by game. *The Guardian*. Retrieved from https://www.theguardian.com/football/2017/may/12/title-chelsea-game-by-game

Bickart, B., & Schindler, R. M. (2001). Internet forums as influential sources of consumer information. *Journal of Interactive Marketing*, *15*, 31–40.

Bieg, M. (2018). peer-to-peer (P2P) network. Retrieved from https://en.wikipedia.org/wiki/Peer-to-peer#/media/File:P2P-network.svg

Bijl, L., Kringhaug, G., Molnár, P., & Sandvik, E. (2016). Google searches and stock returns. *International Review of Financial Analysis*, *45*, 150–156.

Bitcoin.com (2017, October 8). Bitcoin Basics Lesson 2: Essentials of Bitcoin: Medium. Retrieved from https://medium.com/einstein-exchange/bitcoin-basics-lesson-2-5727b9591a78

Bitcoinity.org. (2018a). Bitcoin price volatility. Retrieved from http://data.bitcoinity.org/markets/volatility/2y/USD?c=e&f=m10&g=15&r=day&st=log&t=l

Bitcoinity.org. (2018b). Bitcoin trading volume. Retrieved from https://data.bitcoinity.org/markets/volume/all?c=c&t=b

Bitcoinity.org. (2018c). Number of transactions. Retrieved from https://data.bitcoinity.org/bitcoin/tx_count/all?t=l

Blockchain. (2018). Largest Transactions. Retrieved from https://www.blockchain.com/btc/largest-recent-transactions

Böhme, R., Christin, N., Edelman, B., & Moore, T. (2015). Bitcoin: Economics, technology, and governance. *Journal of Economic Perspectives*, *29*, 213–238.

Bohr, J., & Bashir, M. (2014). Who uses bitcoin? an exploration of the bitcoin community. In *2014 Twelfth Annual Conference on Privacy, Security and Trust (PST)* (pp. 94–101).

Bovaird, C. (2017, January 27). Bitcoin's Volume is Continuing to Decline at Major Exchanges: coindesk. Retrieved from https://www.coindesk.com/bitcoins-volume-continuing-decline-major-exchanges/

Bovaird, C. (2018, June 8). Regulatory Concerns Dampen Bitcoin Volatility: Forbes. Retrieved from https://www.forbes.com/sites/cbovaird/2018/06/08/regulatory-concerns-place-damper-on-bitcoin-volatility/#4d0cc7b8bd8a

Cadwalladr, C., & Graham-Harrison, E. (2018, March 17). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The Guardian*. Retrieved from https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election

Chaparro, F. (2017, September 15). 2 of China's biggest cryptocurrency exchanges are shutting down trading — but bitcoin is soaring. New York: Business Insider. Retrieved from https://www.businessinsider.de/bitcoin-price-chinas-biggest-exchanges-shutting-2017-9?utm_source=marketsinsider&utm_medium=ingest&r=UK&IR=T

Chase, B., & MacBrough, E. (2018). Analysis of the XRP Ledger consensus protocol. *ArXiv Preprint ArXiv:1802.07242*.

Chaum, D. (1983). Blind signatures for untraceable payments. In *Advances in cryptology* (pp. 199–203).

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

Chohan, U. W. (2017). A History of Bitcoin.

Christidis, K., & Devetsikiotis, M. (2016). Blockchains and smart contracts for the internet of things. *Ieee Access*, *4*, 2292–2303.

Christin, N. (2013). Traveling the Silk Road: A Measurement Analysis of a Large Anonymous Online Marketplace.

Ciaian, P., Rajcaniova, M., & Kancs, d.'A. (2016). The economics of BitCoin price formation. *Applied Economics*, *48*, 1799–1815.

CoinMarketCap. (2018a). Historical data for Bitcoin. Retrieved from https://coinmarketcap.com/currencies/bitcoin/historical-data/?start=20130428&end=20181030

CoinMarketCap. (2018b). Historical data for Ether. Retrieved from https://coinmarketcap.com/currencies/ethereum/historical-data/?start=20130428&end=20181029

CoinMarketCap. (2018c). Historical data for XRP. Retrieved from https://coinmarketcap.com/currencies/ripple/historical-data/?start=20130428&end=20181029

CoinMarketCap. (2018d). Top 100 Cryptocurrencies by Market Capitalization. Retrieved from https://coinmarketcap.com

CoinMarketCap. (2018e). Top 100 Cryptocurrency Exchanges by Trade Volume. Retrieved from https://coinmarketcap.com/rankings/exchanges/

CoinMarketCap. (2018f). Total Market Capitalization. Retrieved from https://coinmarketcap.com/charts/

Crosby, M., Pattanayak, P., Verma, S., & Kalyanaraman, V. (2016). Blockchain technology: Beyond bitcoin. *Applied Innovation*, *2*, 6–10.

Crunchbase. (2018). Reddit: Overview. Retrieved from https://www.crunchbase.com/organization/reddit#section-overview

CryptoCoinCharts. (2018). Cryptocurrency Exchanges / Markets List.

Dornbusch, R. (1976). Expectations and exchange rate dynamics. *Journal of Political Economy*, *84*, 1161–1176.

Ethereum Foundation. (2018a). Ether: What is Ether? Retrieved from https://www.ethereum.org/ether

Ethereum Foundation. (2018b). Ethereum: Blockchain App Platform. Retrieved from https://www.ethereum.org

Foley, S., Karlsen, J., & Putninš, T.\=. J. (2018). Sex, drugs, and bitcoin: How much illegal activity is financed through cryptocurrencies?

Garcia, D., & Schweitzer, F. (2015). Social signals and algorithmic trading of Bitcoin. *Royal Society Open Science*, *2*, 150288.

Garcia, D., Tessone, C. J., Mavrodiev, P., & Perony, N. (2014). The digital traces of bubbles: Feedback cycles between socio-economic signals in the Bitcoin economy. *Journal of the Royal Society Interface*, *11*, 20140623.

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, *457*, 1012.

Greenberg, A. (2013). End Of The Silk Road: FBI Says It's Busted The Web's Biggest Anonymous Drug Black Market. Retrieved from https://www.forbes.com/sites/andygreenberg/2013/10/02/end-of-the-silk-road-fbi-busts-the-webs-biggest-anonymous-drug-black-market/#5caa86405b4f

Hansen, S. (2018, June 20). Guide To Top Cryptocurrency Exchanges. *Forbes*. Retrieved from https://www.forbes.com/sites/sarahhansen/2018/06/20/forbes-guide-to-cryptocurrency-exchanges/#3025adf52572

Harrington, M. (2017, January 16). Survey: People’s Trust Has Declined in Business, Media, Government, and NGOs. *Harvard Business Review*. Retrieved from https://hbr.org/2017/01/survey-peoples-trust-has-declined-in-business-media-government-and-ngos

Huang, Z. (2018, January 5). This could be the beginning of the end of China’s dominance in bitcoin mining: Quartz. Retrieved from https://qz.com/1172632/chinas-dominance-in-bitcoin-mining-under-threat-as-regulators-hit-where-it-hurts-electricity/

Joseph, K., Wintoki, M. B., & Zhang, Z. (2011). Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search. *International Journal of Forecasting*, *27*, 1116–1127.

Kelly, J. (2017, January 11). Bitcoin slides after China central bank launches investigation. London: Reuters. Retrieved from https://www.reuters.com/article/us-markets-bitcoin-pboc/bitcoin-slides-after-china-central-bank-launches-investigation-idUSKBN14V10B

Kharpal, A. (2018, August 9). Cryptocurrencies: Regulating the new economy: CNBC. Retrieved from https://www.cnbc.com/2018/08/09/cryptocurrencies--regulating-the-new--economy.html

Kim, E. (2012, July 31). Twitter unveils 'cashtags' to track stock symbols. *CNN*. Retrieved from https://money.cnn.com/2012/07/31/technology/twitter-cashtag/

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*, e0161197.

Kim, Y. B., Lee, S. H., Kang, S. J., Choi, M. J., Lee, J., & Kim, C. H. (2015). Virtual world currency value fluctuation prediction system based on user sentiment analysis. *PloS One*, *10*, e0132944.

Kristoufek, L. (2013a). BitCoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era. *Scientific Reports*, *3*, 3415.

Kristoufek, L. (2013b). Can Google Trends search queries contribute to risk diversification? *Scientific Reports*, *3*, 2713. https://doi.org/10.1038/srep02713

Kristoufek, L. (2015). What are the main drivers of the Bitcoin price? Evidence from wavelet coherence analysis. *PloS One*, *10*, e0123923. https://doi.org/10.1371/journal.pone.0123923

Liew, V. K.-S. (2004). Which lag length selection criteria should we employ?

Mao, H., Counts, S., & Bollen, J. (2011). Predicting financial markets: Comparing survey, news, twitter and search engine data. *ArXiv Preprint ArXiv:1112.1051*.

Matta, M., Lunesu, I., & Marchesi, M. (2015a). Bitcoin Spread Prediction Using Social And Web Search Media.

Matta, M., Lunesu, I., & Marchesi, M. (2015b). The Predictor Impact of Web Search Media On Bitcoin Trading Volumes.

McMillan, R. (2013, August 30). Ex-Googler gives the World a better Bitcoin: Wired. Retrieved from https://www.wired.com/2013/08/litecoin/

Moat, H. S., Curme, C., Avakian, A., Kenett, D. Y., Stanley, H. E., & Preis, T. (2013). Quantifying Wikipedia usage patterns before stock market moves. *Scientific Reports*, *3*, 1801.

Moore, S. (2018, May 26). The Costs To Trading Crypto And How To Cut Them. *Forbes*. Retrieved from https://www.forbes.com/sites/simonmoore/2018/05/26/the-costs-to-trading-crypto-and-how-to-cut-them/#470364595c93

Moskov, P. (2018, May 22). What Is Bit Gold? The Brainchild of Blockchain Pioneer Nick Szabo: CoinCentral. Retrieved from https://coincentral.com/what-is-bit-gold-the-brainchild-of-blockchain-pioneer-nick-szabo/

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

Naughton, J. (2013, April 7). Why Bitcoin scares banks and governments. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2013/apr/07/bitcoin-scares-banks-governments

Oberstone, J. (2017, November 15). The amazing stats behind Man City's dominance of the Premier League. *Planet Football*. Retrieved from https://www.planetfootball.com/in-depth/amazing-stats-behind-man-citys-dominance-premier-league/

Osatuyi, B. (2013). Information sharing on social media sites. *Computers in Human Behavior*, *29*, 2622–2631.

Parker, L. (2017, January 22). Have we just seen the end of fake Chinese bitcoin trading volumes?: Brave New Coin. Retrieved from https://bravenewcoin.com/insights/have-we-just-seen-the-end-of-fake-chinese-bitcoin-trading-volumes

Peters, G. W., & Panayi, E. (2015). Understanding Modern Banking Ledgers through Blockchain Technologies: Future of Transaction Processing and Smart Contracts on the Internet of Money. *ArXiv Preprint ArXiv:1511.05740*.

Pisani, B. (2017, December 12). Here's why bitcoin prices are different on each exchange. *CNBC*. Retrieved from https://www.cnbc.com/2017/12/12/why-bitcoin-prices-are-different-on-each-exchange.html

Pisani, B. (2018, June 14). Bitcoin and ether are not securities, but some initial coin offerings may be, SEC official says: CNBC. Retrieved from https://www.cnbc.com/2018/06/14/bitcoin-and-ethereum-are-not-securities-but-some-cryptocurrencies-may-be-sec-official-says.html

Pitta, j. (1999, November 1). Requiem for a Bright Idea: Forbes. Retrieved from https://www.forbes.com/forbes/1999/1101/6411390a.html#2742bcde715f

Popper, N. (2016, June 29). How China Took Center Stage in Bitcoin's Civil War: The New York Times. Retrieved from https://www.nytimes.com/2016/07/03/business/dealbook/bitcoin-china.html

Preis, T., Moat, H. S., & Stanley, H. E. (2013). Quantifying trading behavior in financial markets using Google Trends. *Scientific Reports*, *3*, 1684.

Prusakova, M. (2018, July 14). Cryptocurrency regulations in Russia: Medium. Retrieved from https://medium.com/@mariaprusakova/cryptocurrency-regulations-in-russia-6c1be4c8a92a

Reddit Inc. (2018). Reddit API Documentation: Overview. Retrieved from https://www.reddit.com/dev/api/

Richter, F. (2017, November 17). The Explosive Growth of Reddit's Community. *Statista*. Retrieved from https://www.statista.com/chart/11882/number-of-subreddits-on-reddit/

Ripple. (2018). XRP: The Digital Asset for Payments. Retrieved from https://ripple.com/xrp/

Ron, D., & Shamir, A. (2013). Quantitative analysis of the full bitcoin transaction graph. In *International Conference on Financial Cryptography and Data Security* (pp. 6–24).

Twitter. (2017). *Tweeting Made Easier*. Retrieved from https://blog.twitter.com/official/en_us/topics/product/2017/tweetingmadeeasier.html

Schollmeier, R. (2001). A definition of peer-to-peer networking for the classification of peer-to-peer architectures and applications. In *Peer-to-Peer Computing, 2001. Proceedings. First International Conference on* (pp. 101–102).

Sovbetov, Y. (2018). Factors influencing cryptocurrency prices: Evidence from bitcoin, ethereum, dash, litcoin, and monero.

StatCounter. (2018). Search Engine Market Share Worldwide. Retrieved from http://gs.statcounter.com/search-engine-market-share

Statista. (2018a). Combined desktop and mobile visits to Reddit.com from February 2018 to July 2018 (in millions). Retrieved from https://www.statista.com/statistics/443332/reddit-monthly-visitors/

Statista. (2018b). Share of desktop search traffic originating from Google in selected countries as of June 2018. Retrieved from https://www.statista.com/statistics/220534/googles-share-of-search-market-in-selected-countries/

Streitfeld, D. (2018). Amazon Hits $1,000,000,000,000 in Value, Following Apple. Retrieved from https://www.nytimes.com/2018/09/04/technology/amazon-stock-price-1-trillion-value.html

Swift, T. (2018). News. Retrieved from https://www.taylorswift.com/news

Takeda, F., & Wakao, T. (2014). Google search intensity and its relationship with returns and trading volume of Japanese stocks. *Pacific-Basin Finance Journal*, *27*, 1–18.

Twitter. (2018). Pricing of Twitter APIs: Search Tweets: Full Archive. Retrieved from https://developer.twitter.com/en/pricing/search-fullarchive

Twitter. (2018). *Annual Report 2018*. Retrieved from http://www.viewproxy.com/Twitter/2018/AnnualReport2017.pdf

Vaghela, V., & Tan, A. (2018, April 23). How Malta Became a Hub of the Cryptocurrency World: Bloomberg. Retrieved from https://www.bloomberg.com/news/articles/2018-04-23/how-malta-became-a-hub-of-the-cryptocurrency-world-quicktake

Webster, M. (2018). Merriam-Webster online dictionary. Retrieved from https://www.merriam-webster.com/dictionary/cryptocurrency

Wikimedia Foundation. (2018). Wikipedia:About: Wikipedia history. Retrieved from https://en.wikipedia.org/wiki/Wikipedia:About#Wikipedia_history

Wikipedia. (2018). Bitcoin. Retrieved from https://en.wikipedia.org/wiki/Bitcoin

Wolf, B., & Flitter, E. (2014, February 28). Mt Gox: The brief reign of bitcoin's top exchange. *Reuters*. Retrieved from https://www.reuters.com/article/us-bitcoin-mtgox-insight-idUSBREA1R06C20140228

Yelowitz, A., & Wilson, M. (2015). Characteristics of Bitcoin users: An analysis of Google search data. *Applied Economics Letters*, *22*, 1030–1036. https://doi.org/10.1080/13504851.2014.995359

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

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Alle Stellen, die wörtlich oder inhaltlich den angegebenen Quellen entnommen wurden, sind als solche kenntlich gemacht.

Die vorliegende Arbeit wurde bisher in gleicher oder ähnlicher Form noch nicht als Magister- /Master- /Diplomarbeit /Dissertation eingereicht.

Datum, Unterschrift