UNIVERSITETET I AGDER

Statistical Modelling and Risk Analysis of Bitcoins Exchange Rate

FILIP FILIPOVIC

JOAKIM NILGARD

SUPERVISOR

Jochen Jungeilges

University of Agder, 2018

School of Business and Law

Department of Economics and Finance

Preface

As this thesis marks the end of five rewarding years of studies for us, we would like to express our thanks and gratitude towards all those who have helped us through the challenges to achieve the MSc in Business and Administration and to write and proudly present this paper that is a result of three months of hard work. Despite a belated start, several demanding challenges along the way, and hours of struggles for learning new software, we have completed this research that we are immensely proud of to have done.

First of all, we wish to give our sincere thanks to our supervisor, Professor Jochen Jungeilges, that has not only given the proposition to this incredibly interesting topic we have studied, but also given us invaluable support and contributions through his comprehensive knowledge about the topic and most issues we have had to deal with.

We want to acknowledge and express our appreciation and heartfelt gratitude towards our families, friends, fellow students and close ones for your understanding, contribution and love throughout this time. Even if not necessarily expressed directly, your support has been immeasurable and essential for us to have come this far. Thank you.

Kristiansand, 31st May 2018

Joakim Nilgard

Kristiansand, 31st May 2018

Filip Filipovic

Abstract

Bitcoin is a phenomenon that is new and there is little information on how and why it behaves as volatile as it does. This thesis uses existing data on Bitcoin's exchange rate to estimate a model that describes the pattern and use it in a financial risk analysis. We also aim to contribute as a foundation for further studies in this field.

The statistical properties of the log-return of the exchange rate are analysed and it is deemed to be *iid*. From the eleven distributional candidates we study is the fitted skew generalised *t* distribution proven to represent the data best after evaluation by criteria and statistics. The estimated VaR and ES show that the rate is volatile and that the risk from investments is still high.

The findings show that it is necessary to describe the exchange rate with complex and flexible distributions, and even if the data shows more stability today than earlier is it important to show caution in interpretations and evaluations on the topic.

Keywords: Bitcoin, cryptocurrencies, statistical distributions, statistical analysis, exchange rate, modelling

Table of Contents

List of figures, tables, and pictures

1. Introduction

Bitcoin is a currency, an ideology, and an unpredictable phenomenon that has been getting much attention over the last few years. Even with the enormous amount of financial methods and techniques that exist today, few of them have been able to explain or describe the development and volatility of Bitcoins exchange rate well. There have been a handful of articles and papers that have attempted to do just this. If some results have been obtained, there is certainly room for improvement in many aspects of the topic.

The purpose of this thesis is that we want to get a better understanding of Bitcoin and its properties and dynamics. This will be attained through the process of obtaining two financial risk measurements, value at risk and expected shortfall, abbreviated VaR and ES. If these techniques are proven to inhibit strong volatility, it is certain that the exchange rate of Bitcoin will inhibit the same characteristics. Even though it is very clear that Bitcoin is volatile, it is desirable to have some numbers to relate to than a vague concept that is not proven. VaR and ES are often assessed in risk management when one needs relatively simple, but well-informing, indicators on the risk involved in potential investments or financial decisions. John Hull (2015, pp. 258-259) defines the concepts of VaR and ES very simply, with VaR as an estimate on “how bad can things get?”, while ES asks what the expected loss is in case things go bad.

In order to estimate VaR and ES is it necessary to have a fitted model, or distribution, of the data we want to study. VaR is a model-dependent measurement, and ES, which is sometimes (misleadingly) called conditional value at risk, is again dependent on VaR, so a well-fitted model is required. A challenge in general is to choose which distribution to use, and this becomes especially difficult when studying a new phenomenon that behaves very differently than others, more well-known currencies or assets. Luckily, there has been some research on this topic before, where several distributions have been attempted fitted to the exchange rate of Bitcoin. By learning from these, some frequently distributions used in finance can be included. Admittedly, one should not take for granted that these are the ones that should be used when there could exist some distribution that is able to explain a phenomenon better, especially when it is a more complex and diverse phenomenon than normally studied.

There are different prerequisites the data must fulfil in order for us to be able to estimate and fit a distribution to the observed data. We want a model that has estimated parameters that gives the highest likelihood of representing the data, and that is why we will perform the model estimation through the maximum likelihood estimation (MLE) procedure. The observations of the sample data are required to be independent and identically distributed, *iid*, if the MLE procedure is to be exercised. This means that before the distribution can be estimated, the observations must be checked to be random, non-correlated with each other, and have the same variance over time. This will be our starting point in the analytical part of the thesis after some basic statistical properties of the data have been calculated.

Because Bitcoin is such a new concept, and thus in order to have an understanding of why Bitcoin behaves as it does and the need for such comprehensive studies, we consider it necessary to give the reader an insight into what this concept is all about. So, before the statistical aspects are approached, there will be an introduction to Bitcoin where we look at what makes a cryptocurrency different from other currencies, the technology behind Bitcoin called the blockchain, how mining, hash, forks and security are important contributors to explain Bitcoin's behaviour, and, lastly, some historical and exchange events. The information we provide on Bitcoin here is only the tip of the iceberg, but we believe that this will give enough to acknowledge the necessity of deeper and thorough research on the properties and characteristics of Bitcoin.

The motivation for our thesis, to study the exchange rate of Bitcoin, is to better understand its behaviour, that is very different from most other currencies and assets. The legitimacy for previous findings will be evaluated, and if they are proven not to be consistent will a new fitted distribution be proposed if it proves to describe the exchange rate better. We will perform certain risk measures with the proposed fitted distribution in order to set a starting point for later studies. With a fitted distribution will it also be easier to predict values that are out of sample.

This thesis is organised and consists of three main parts divided into seven chapters. The parts are the concept of Bitcoin, the methodology, and the results and discussion.

In chapter 2, the concept of the phenomenon Bitcoin will be established and given some general explanation on how it works, both from a historical, political, and technical way. The underlying technological breakthrough of blockchain will also be acknowledged and addressed, as it is vital to know about this in order to have some idea of why Bitcoin behaves as it does and why it has gotten the attention from the public it has done over the nine years it has existed. After the concept has been defined will we, in chapter 3, present the sheer amount of literature that have previously studied the topic we are about to venture. In chapter 4 will the data we have collected be presented and analysed accordingly. Non-parametrical tests and explanatory plots are assessed to define and assign the required properties to the data in order to be able to describe it by statistical distributions. Some strong assumptions about the data will be made, but they are described and explained when needed. Then, in chapter 5, the distributions that are to be fitted to the data will be presented together with the criteria they will be evaluated against each other by. Two measures of financial risk will also be applied, to see if the data proves to be as volatile in studies as it seems to be empirically. In chapter 6 and 7 will our estimated results for *iid*, distributions, criteria, and financial risks be presented and commented promptly and objectively. A summary and concluding remarks will be given in chapter 8, where we will draw lines from our findings to relevant sources and studies, and also propose some ideas for further research.

2. The concept of Bitcoin

2.1 Cryptocurrencies as another type of money

For as long as humans have existed we have had the need to be able to trade. A good or service for another, at a fair rate depending on needs, scarcity or availability. The relevant goods are traded at their own value in itself, which is the purpose of a commodity. Commodities have been the most used type of money through times because everything that is traded at its own value is a commodity. Types of commodities include valuables as gold, silver, copper, platinum, but also more common mediums as rocks, coffee, sugar, tobacco, cocoa, barley etc.

Commodities may be traded at its own value, but because of limits of capacity, possibilities of transportation, and other difficulties, it is useful to have a medium that may represent the value of the commodity by not having it physically present. This is what we call representative money. The medium's intrinsic value may be low or close to none, but there has to be something valuable the face value of the representative money represents. Types of this form of money includes i.e. certificates, bank notes, claims.

Because of the limitations of commodities, fiat money was introduced. Fiat comes from Latin and translates to *let it be done*. Most of the world's currencies today are fiat currencies, as the face value of the money is often greater than the value of the material it is made from. It is correct that fiat money can be perceived as representative money, but the idea is that the money is not a direct representation of something else like a commodity but rather just a value enforced by the issuing government which is agreed on by all exchanging parties. The intrinsic value of the fiat money is basically close to non-existent in principle. In the US it was so that, until president Nixon completely lifted the gold standard in 1971, most circulating money was backed by gold owned by the government.

All these forms of money are physical and tangible, and the applied value is fairly easy to comprehend. What if we were to apply value to something un-tangible, digital, and possibly unregulated? The concept of virtual money, or currencies, is fundamentally different from the other types, first and mainly because of the aspect of non-physicality, but also in other regards that will be explained later in this chapter. There are several different types of virtual currencies such as coupons, centralized systems as PayPal or eCash, fictional currencies in online games, decentralized formats as Ripple, and cryptocurrencies. Common for all these formats is that they all are "*a digital representation of value, not issued by a central bank, credit institution or* *e-money institution, which, in some circumstances, can be used as an alternative to money"* (European Central Bank, 2015). It is also called a *virtual currency scheme* because it is not considered money or a currency from a legal perspective, with regard to governments or central banks and are often controlled by the issuers themselves. Bitcoin is one of the exceptions to this, as it is not controlled by anyone in particular but rather by the network, or community, as a whole.

A cryptocurrency is fundamentally different from other virtual currencies because, as the name suggests, it is based on cryptography. This means that the currency is built upon intriguing algorithms and complex mathematical compositions which are solved by computers within the blockchain. Where does the money come from? To take Bitcoin as an example: a bitcoin is given as a reward for solving a problem, or 'block', that is the algorithm, and is transferred to the virtual wallet of the solver. This procedure is called mining – to have a computer run software to solve riddles. The difficulty of this procedure is increasing by the amount of people mining for bitcoins and by how many bitcoins that are already in circulation. The amount of mined coins is relevant because the total amount possible coins to exist is limited to Ƀ21 million, as defined by the original code. This limit itself is not able to prevent the great volatility of the exchange rate we have seen over the years, but it prevents seemingly unstoppable inflation we observe in most fiat currencies. There is some inflation, roughly 4% p.a., but this rate is declining, and by the time the limit is reached it is not unlikely there will be a deflating trend. It is also believed that Bitcoin will experience the deflationary spiral, that the limit induces hoarding of bitcoins, but because of the nature of bitcoin it is not an altogether plausible assumption to make. By hoarding bitcoins, one would solely do it in speculation, not because of a foreseeable beneficial use of the money in the future.

Before conceptualising Bitcoin is it necessary to get an understanding of the technology behind it, the blockchain.

2.2 The Blockchain

First introduced and conceptualized in 2008 by Satoshi Nakamoto, the blockchain is functioning as the cornerstone of cryptocurrencies. When it was first introduced, blockchain's potential was unclear, as it was initially made mainly for bitcoin as an open-source code. As of 12th March 2018, we have almost 1600 different cryptocurrencies (CoinMarketCap, 2018) and there are continuously more currencies popping up from all over the world. Since its inception, the community has found many ways to make use of the blockchain. Some people from technological communities have even proclaimed that this may be one of the biggest unintentional innovations ever made. The basic principle of how the technology works, is that it allows information to be stored and distributed, while it doesn't allow itself to be copied or altered (Rosic, 2017).

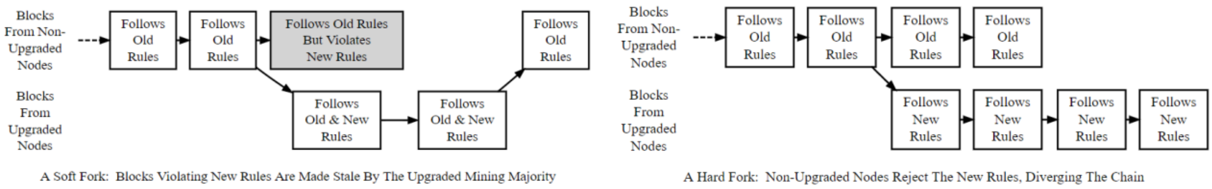
The identity of a transactor, masked by the private key, is kept secret by encryption, but is publicly accessed by a public key. The private key is essentially a person's wallet and is made out of a 256-bit number, often presented in hexadecimals. If the number of a private key is known, one has access to this particular wallet. That is why a public key is needed, to prevent others from accessing your wallet. The public key works like a pin-code for the private key. All private keys are able to provide the owner with a set of public keys, which is the key(s) that are available and shown to the public in the blockchain. A transaction between two wallets is performed by an encrypted digital signature on the transaction which is issued by the sending wallet. This signature is only possible to be decrypted by the receiving wallet – even a minor change will alter the whole procedure. The receiver is able to decrypt the transaction that is sent to his public key with his private key and receive the transferred money to his wallet.

The blockchain is considered a digital ledger that registers and records financial transactions. It is also completely incorruptible, like a non-configurative balance sheet. The information that is saved as a database on the blockchain is not saved in one single unit or database, but is received, copied and stored across the whole network in every node. A node is a computer that is connected to the blockchain system. There are two important statements we can draw from this. Firstly, the blockchain cannot be controlled by a single computer or unit, unless this single unit, or collaboration of units, stands for more than half of the processing power, called a 51% attack. Secondly, the concept has no single point of failure because everyone maintains the same records and continue the process. Since Bitcoin was introduced, there has never been any operational failures, only hacking attempts or human errors.

The blockchain system runs on a network of nodes. When a node connects to the system, a copy of the blockchain gets downloaded or updated on the computer. Every computer in the network is its administrator and user concurrently, and all these nodes validate the transactions together. They do this by solving mathematical puzzles – what we described as the mining process.

We know that the blockchain consists of several blocks containing data about all transactions done throughout Bitcoin's existence, but how is this done in practice? When a person is initiating a transaction, he sends the amount of money out of his wallet and gets the transmission confirmed by nodes in the network. Then the bitcoin gets confirmed, it is passed on to what is called the memory pool, or Mempool, and is contained there until it is included into the first new block that is mined by a computer and is finally added to the transactional history amongst all other transactions and is in principle confirmed to have successfully been recorded. This is effectively the bottleneck of the whole network, as the exact time for finding the next new block is uncertain. In ~66% of all cases the new block is found in 10 minutes or less, but there is also a small probability of 0.3% to wait more than 60 minutes for one block. The interval of block-creation follows a standard Poisson process with no memory and has the same probability of time-spending on each block every time a new block is getting solved (bitcoinwiki, 2017). This leads to that the Mempool may not be emptied within, in worst case, the hour, although this is rarely seen. The time for a transaction to be successful will vary because of this.

It must also be mentioned that there are very rare cases that the code may lead to some temporary difficulties, most seen in practice that two different computers may solve the same block at the same time and thus make a so-called fork. The network will then be divided between the two chains, because the chains are equally long and difficult to solve. Luckily the code is also functioning such that when the network is divided, and one computer solves a new block, regardless of which chain it has accepted, the network will accept this as the true chain, since it is the longest and most difficult, and we will regain equilibrium again. This type of fork is called a soft fork, because the software protocol is only rendering blocks invalid.



Picture 1 - Soft vs. Hard fork (Investopedia, 2016)

In comparison, a hard fork is a method of splitting the chain by revalidating previous versions of the software. This have been done in few, but more drastic cases when, for instance, there have been disagreements about an update of the code (Redman, 2017) or after security failures by hacking (Wong, 2017).

The world today may have a great use for the blockchain technology, especially with its benefits as being non-manipulable and the security it possesses. Even if the technology is useful most places, it may probably be needed most in developing countries, for instance, where corruption and swindling are enormous problems. Examples on where the blockchain may be used is digital storage, management, predictions, holding medical records, smart contracts, efficient transactions and a more advanced internet.

2.3 Functionalities of Bitcoin

First, to clarify the difference between the capitalized word Bitcoin and non-capitalized bitcoin, we cite bitcoin.org's definition:

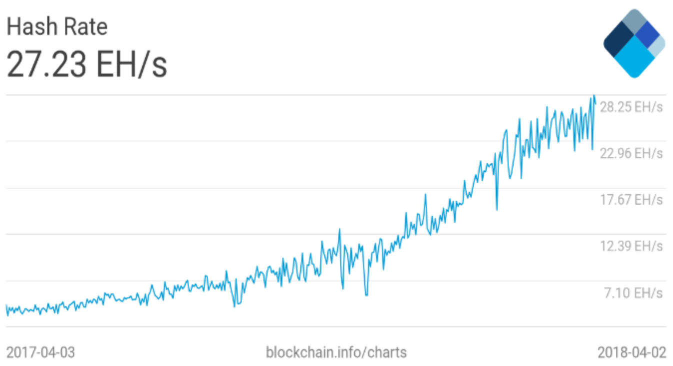
'Bitcoin - with capitalization, is used when describing the concept of Bitcoin, or the entire network itself. e.g. "I was learning about the Bitcoin protocol today."

bitcoin - without capitalization, is used to describe bitcoins as a unit of account. e.g. "I sent ten bitcoins today."; it is also often abbreviated BTC or XBT.' (bitcoin.org, 2018)

Also, hereafter, all references as of "today" refer to the 1st April 2018.

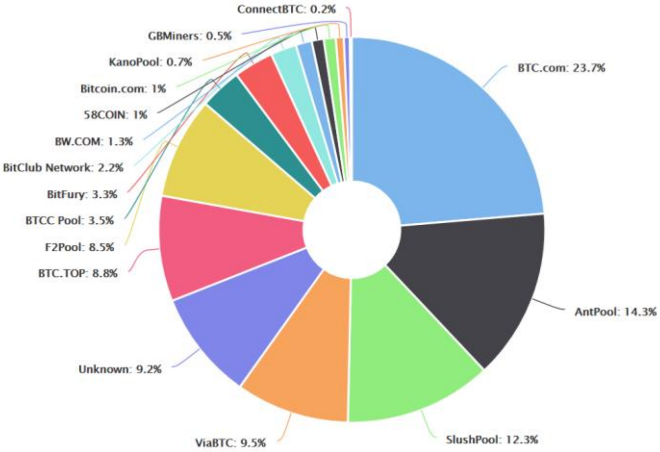
Bitcoins exchange rate was at $0.08 when Mt. Gox started their exchange 8 years ago. Today the rate is $7,400 – an incredible growth by a factor of 92,500 over those eight years is what most people would consider unthinkable. This is even after a dramatic decrease from December 2017 when the rate had an all-time peak of $19,666 at the exchange of Bitstamp, or a factor growth of almost 250,000. It should also be pointed out that just from the 1st December to the 17th, when the peak was observed, the rate more than doubled just during this short period, with daily returns of even +20%. Bitcoin's capitalized value today exceeds $115 billion, almost dropped to a third of the market capitalization from the peak four months ago (blockchain, 2018). This illustrates just a few important moments of Bitcoin's relatively short life of merely 9 years.

The efficiency Bitcoins mining process is rated by a computers hash rate, or hash power, which is the speed a computer is able to solve the algorithms. In practice, the hash rate is "*the measuring unit that measures how much power Bitcoin network is consuming to be continuously functional*" (Khatwani, 2018). This graph shows that today the total hash rate is 27.23 EH/s (exabits per second), which is about 27,230,000 TB/s (terabits per second), or 27,230,000,000,000,000,000 B/s. This number is so incredibly large it is hard to fathom – even the combined power of the world's 500 most powerful computers would only represent a fraction of the processing power of the Bitcoin mining-network (Hoffmann & Watchulonis, 2015).



Picture 2 - Hash rate development over the last year (blockchain, 2018)

Since the value of Bitcoin is as high as it is, as mentioned and shown later in part 4.2, it is not surprising that there has been many people that want to get hold of some of these coins the easy way, either by stealing directly from an owner of coins by hacking, taking advantage of persons that do not take their security seriously, or set up a fake exchange to exploit incautious people. Obviously, a wallet is more exposed when having a hot wallet rather than a cold wallet. A hot wallet is a wallet that is connected to the internet, while a cold wallet is not connected or offline. This is part of the reason banks are focusing on having such a strong digital security system, so that their customers' accounts are not exposed. Wouldn't the network be fairly easily compromised by hackers, scammers, or other attackers when the blockchain does not inhibit any security software or firewall? No, and it gets even harder over time. For an attacker to directly overcome the existing network in a 51% attack, he must be able to outpace all other computers in the network by implementing a "new" blockchain and stand for the majority of the hash processed, which is considerably difficult with regard to the existing hash produced by the network as presented above. Of course, it would be possible to perform this type of attack through a network of computers which has the combined majority of the power, but, as of today, the largest contributor of hash is the pool from BTC.com, which utilize 23,7% of the power, as we can see in picture 2. This is a remarkable rate, but for this pool to (theoretically) be able to take "control", their hash rate must almost be subject to an increase of over 110%. This is possible since the whole concept is rooted in the blockchain. All transactions done with bitcoins are registered and stored in the chain, which keeps records of these in full publicity for everyone to see.



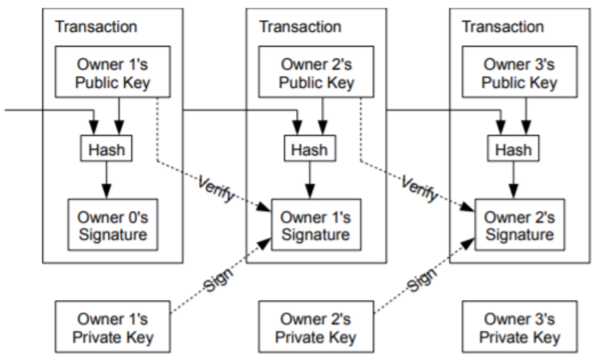
Picture 3 - Distribution of hash-contributors (blockchain, 2018)

2.4 History and exchange of Bitcoin

In 2013 there were three main contributors in the Bitcoin network that founded the first exchange and interest centre in downtown New York called the Bitcoin Centre NYC. The purpose of this establishment was not only to be a meeting centre for enthusiasts and stakeholders for discussion, education and encouragement of trading in bitcoins and cryptocurrencies rather than traditional money, but also to become a regulated exchange platform with live currency exchange – a "*fusion between Wall Street and Bitcoin*" (Cannucciari, 2016), and was thus fittingly placed next to the Wall Street stock exchange.

There has also been several online, private and, to some degree, unregulated exchanges for Bitcoin as well. For instance, Mt. Gox, Bitstamp, BTC-e, Kraken, CEX.io, and a handful others, are some of the non-physical places one is able to exchange in cryptocurrencies. The existence of these has not been flawless. Money laundering, embezzlement, fraud, DDoS attacks, hacking of hot wallets, exploitation have occurred in several of these exchanges and caused loss of values of millions of dollars. This is in addition to straight-out fake exchanges, personal hacks, and people losing their private keys or don't take security measurements when managing their wallets. Mt. Gox's failure and closure in February 2014 is plausibly the most infamous event with regard to exchanges, whereas allegedly ₿744,000 were stolen at the value of almost $500 million, estimated at that day's bitcoin price. It is still unclear exactly how many bitcoins that were stolen and all of the circumstances revolving the case, but ultimately this shows how important cyber security is when handling this type of business.

In Satoshi Nakamoto's white paper from 2008, *Bitcoin: A Peer-to-Peer Electronic Cash System* (Nakamoto, 2008), "he" first described his intentions of the concept of Bitcoin. This paper was written and published the year before Bitcoin was launched. His paper and its driven mindset has become a fundament for many newer cryptocurrencies as well. The main idea is to have an electronic cash system, peer-to-peer, without any third parties and at the same time use of digital signatures, so that extra fees and probability of being compromised are negligible and that the double-spending possibility is removed from transfers of money. Nakamoto's issue about a third party is that it is based on the trust-based model, which is a model exposed for exploitation and other human influences, transaction costs, mediating disputes and lacks practicality. His idea was to produce and present a system that is based on cryptography rather than human trust, which is un-falsifiable per se. There is of course some trust involved with respect to the possibility of overwriting the block chain when one actor is standing for more than half of the hash of the network, but the greater the network becomes with increasing processing power, the difficulty



Picture 4 - Verification/signature procedure from Nakamoto's paper of overthrowing the network and implementing a new "true" block chain grows exponentially for the attacker, making it an ever increasingly challenging task.

Bitcoin's price has been one of the big talks over the world, especially after last year. When Bitcoin was introduced by Nakamoto in 2009, and he performed the first Bitcoin transaction with Hal Finney in the amount of 50 bitcoins, the net worth of a coin was nothing more than around ¢3. This transaction is rooted, and is also the only transaction, in block 0, which is also called the genesis block. Both the price, or relative value, of Bitcoin and the volume traded in bitcoins have seen many dramatic changes over the years. We will study this more in detail in part 4.2.

Now that we have established some basic understanding of what Bitcoin is and how it works, we will proceed in the next part to present and discuss some of the literature that already exists on Bitcoin from a statistical, financial, or economical perspective.

3. Literature review

There does not exist much statistical literature on Bitcoin yet. We will however present some of the articles that have contributed to this area of research to establish a framework around the topic at hand.

The research article by Chu, Nadarajah and Chan's *Statistical Analysis of the Exchange Rate of Bitcoin* (2015) is the first article to use data to perform an analysis of the rate and use a selection of statistical methods to fit parametrical distributions to the log returns of the exchange rate. They used data from the exchange company Bitstamp, consisting of 951 observations from September 2011 to May 2014. They found that the five-parametric generalised hyperbolic distribution gave the best represented model of the logarithmical return of the exchange rate. As this article coincides much to ours, it has been natural to compare some of our results to what they have found – both restricting our data to the one they had to see if our methods provide matching results, and with our extended period to study the development and changes in the results. This article was a part of Chan's thesis for PhD (Chan, 2016).

There have also been a few other studies that have been attempting to have a fitted distribution to describe the exchange rate, both for Bitcoin and other cryptocurrencies. Osterrieder (2017) is the first to study statistical properties of cryptocurrencies by more than just Bitcoin and fiat currencies. He uses data only from June 2014 to November 2016 to fit seven distributions, the normal, Student's *t*, generalised *t*, hyperbolic, generalised hyperbolic, asymmetric normal-inverse Gaussian, and asymmetric variance gamma, for seven cryptocurrencies, Bitcoin, Ripple, Litecoin, Monero, Doge, Dash, and MaidSafeCoin. With regard to Bitcoin, which is most relevant in our case, he concluded that it exhibits heavy tails, and that the asymmetric Student's *t* distributions gives the best fit overall. This is despite other distributions give higher log-likelihood, but the difference is not very large, so it would be easier to describe the data by parsimony.

Collaboratively, Osterrieder, Chan, Chu, and Nadarajah (2017) follow up on each of their previous studies by applying the 15 distributions used by Chu, Nadarajah, and Chan to the seven cryptocurrencies used by Osterrieder. The conclusions concerning Bitcoin are very similar in the papers and has provided us with guidance in our own research. The generalised hyperbolic distribution is selected to give the best fit of the data. They used data from June 2014 to February 2017. They modelled the same data later with twelve GARCH models, as proposed by themselves (Chu, Chan, Nadarajah, & Osterrieder, 2017).

Theodossiou (1998) gave a proof of the skewed generalized *t* distribution and provided a practical approach. The most interesting part for us in this paper was the way the distribution was fitted to model stock exchange index from USA, Canada and Japan. This research showed how this distribution works in practice as a very flexible and applicable model and gave proof that it gives a good fit in statistical analysis in finance. The paper provided a better baseline for financial risk management analysis and financial modelling like GARCH and EGARCH. We think this distribution will provide a reasonable fit to the exchange rate of Bitcoin, and, since it has not been included in the studies mentioned above, function as a worthy candidate in addition to the other distributions.

We believe this thesis will contribute to understanding of the phenomenon and provide more insight to it than the presented research has so far. This is mainly by studying the development over a substantially longer time-period than before. Some literature's time span is chosen to be shorter than necessary because it was aimed to compare several cryptocurrencies and their properties over the same time period rather than studying the longest available set of observations for each currency. It is interesting to see whether or not it turns out to have similar characteristics as the shorter periods, or if inclusion of earlier data makes a big difference in the modelling process.

This literature has given us good fundamental insight in theory and has guided us in our study of Bitcoins exchange rate. We hope to make a useful model for risk management analysis that can be used by others in future studies.

Now that the concept and existing literature is known, some preliminaries to the properties of the exchange rate will be set to be able to study the exchange rate further and perform the necessary analytics.

4. The data

4.1 The gathering of data

The data on Bitcoin's exchange rate to United States Dollar (USD) is gathered from Quandl, a platform that stores financial, economic, and alternative data as a source for analysts. The hotlink to the source from Quandl as given in the paper by Chu, Nadarajah, and Chan were currently unavailable, so in order to get hold of the data we had to receive it from Quandl's addon in Excel. This dataset was then imported to Stata and R as an Excel spreadsheet before we started analysing and processing it. The reasoning for choosing the rate from the exchange company Bitstamp is that the exchange has a significant and stable trading volume over a long time-period. Bitstamp was first founded in Slovenia, then outsourced to Britain for financial and legal purposes, and is apparently also located in Luxembourg from 2016 on. Their exact location of operation now is somewhat uncertain (Wikipedia, 2018). Bitstamp was the world's second largest trader by volume a while back, but it has remained, on average, the fifth largest over the last five years with noticeable decline in volume from January in 2017 (Cieśla, 2018). An argument on why not to choose to use other sources, for instance the Bitcoin Price Index, was due its shorter period of available data. It lacks in fact two years in comparison to Bitstamp's observations. This would in itself not be an immediate problem for us to use these data, but we seek to have a large number of observations in order to study the exchange rate more broadly and lengthily.

There are twenty-one observations that are lacking from the dataset collected from Quandl. In 2011 we lack data from the 30th September, the 1st, 2nd, 15th, 16th, 18th, 19th, 22nd, 23rd, and 27th October, the 2nd, 3rd, 7th, 23rd, and 27th November, and the 4th and 17th December. It is uncertain why there is no trading data on these dates. It is unlikely that these are ripples due the hack of Mt. Gox in June 2011, but since Bitstamp were founded and started business in August 2011, our best guess is that it is the consequence of some technical or practical issues regarding the early-day trading. From 2015 we also lack the data from 6th-8th January. This is the result of a hack on Bitstamp on the 5th January. The trading became suspended after the hacker got away with Ƀ18,866, valued at roughly $5 million, and they spent the consecutive days rebuilding the system and reconstructing their site (Kodrič, 2015). The total expenses, including loss of coins, costs, and reputation damage, is thought to exceed $8 million.

4.2 Exchange rate, volume and log-returns

The studied variable from Bitstamp's dataset is the close value of the exchange rate.

We will first consider the first four years, as the volatility in the exchange rate in the earlier years is not as clear when looking at the whole period at once. Until the start of 2013, the value had increased to just more than $13. With a rough estimate of a total growth of 400%, which is quite dramatic, it is not much compared to how it will develop in the near future. As we see from figure 1, the value over the next year exceeded $1000, for only to fall to a short oscillation between $200 and $300 during most of 2015. By the turn of the year, we see a more or less steady, but volatile, increase of the exchange rate, which continues to grow until the end of 2017. After three years of fall and rise, the rate reached $1000 again the 1st January 2017, and it practically exploded from here on, at least for a year. When observing figure 2, the daily exchange rate from the 1st January 2016 until the 1st April 2018, we see that the rate went from $10,000 at the 1st December 2017, peaked on the 17th at almost $20,000, and then plummeted down to $10,000 again on the 1st February 2018. Since then, the price has hovered between $7,000 to $11,000, not leaving much room for any certainty for where it will go next – evaporate, enrich, or actually maintain stability.

Figure 1 - Bitstamp's exchange rate from 11/09/2011 to 1/1/2017

Figure 2 - Bitstamp's exchange rate from 1/1/2016 to 1/4/2018

The traded volume is also interesting to observe. In figure 3, we see only the traded volume by Bitstamp over its period of trading. This is only a fraction of the total volume, as seen clearly from the graphs in appendix A. We will not dwell much on these plots, but it should be noticed the enormous difference the inclusion of the Chinese Yuan (CNY) makes, especially during Q4 in 2016 and the first days of January. Here are peaks showing daily trading in bitcoins up to

Ƀ14 million only in Yuan, which is a huge amount considering there only existed about 16 million bitcoins at that time. Nonetheless, the American Dollar (USD) is traded at most at 29th February 2016 with a volume of Ƀ955,000. Bitstamp has been trading at a daily volume of Ƀ10,000 on average but has also traded up to Ƀ137,000. Despite of not being the largest exchange in terms of volume, we regard Bitstamp of being representable for the majority of the exchanges' rates. This is because the rate does normally not vary much from exchange to exchange.

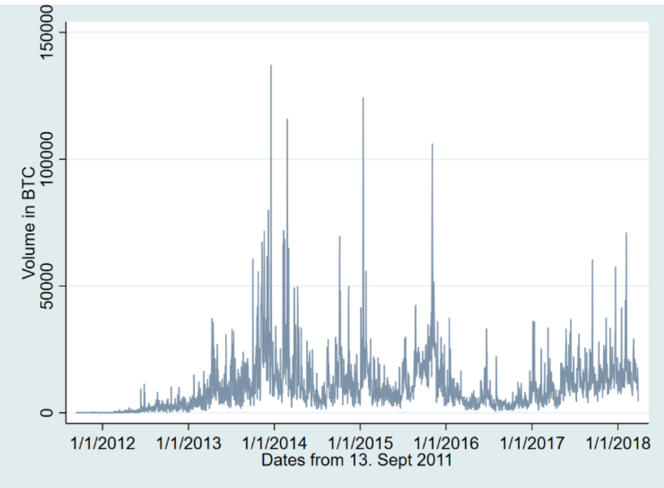


Figure 3 - Bitstamp's transaction volume

We will transform the closing rate to their logarithmical returns. This is preferable in our case, mostly because of normalisation, that the variable is viewed in a comparable metric, time-additivity, and numerical stability. For thorough reasoning, see Quantivity (2011).

To check the consistency of our data, will we compare our data to that of Chu, Nadarajah, and Chan. They are evaluating data from 13th September 2011 to 8th May 2014, so the shape of the log-returns in this period should be identical to figure 1 in their article. The plot of the raw exchange rate will obviously not show any particular issue with this, as the changes early on are minimal in general, but great relative to each other. It was not mentioned explicitly in their paper how to treat the observations that are missing, as it is a few ways to deal with this issue. After we calculated and plotted the log-return, we saw that, when comparing, that the data were differentiated by a couple of extrema, especially in the early period where we see a substantial lack of observations. This may or may not have any significant impact on our findings, but we will make effort to mimic the research as similarly as possible and must thus explore to find their method when it is not pointed out in the text.

We tried first to overlook the values that were lacking and treat them as zeros. Of course, this will prevent the possibility of performing certain statistical calculations, and will also be false in a way, because even if Bitstamp did not have any trades these dates, there was still some value to Bitcoin. It cannot just be treated as it had no value. We also tried to find and substitute the values with either data from another exchange, Bitcoins Trading average, or an average of the adjacent available data over the dates with missing values. Even if we made some seemingly usable plots for the log-return, the use of foreign data could render the absence of confidence in the results. We chose eventually to only remove the dates with missing values and exclude them from the dataset. It is important to not have any lacking observations, because this will both alter some results, or make us unable to perform a number of calculations. The result is a plot of 2,372 observations in total, a substantially larger timespan studied than in every other paper in this field. Without removing the zero-values, there would not have been too much of a difference visually in the plot in figure 4. The main difference would be the lack of the extreme, and highest, observation on the 28th October 2011, and a few other negative peaks in the early period.

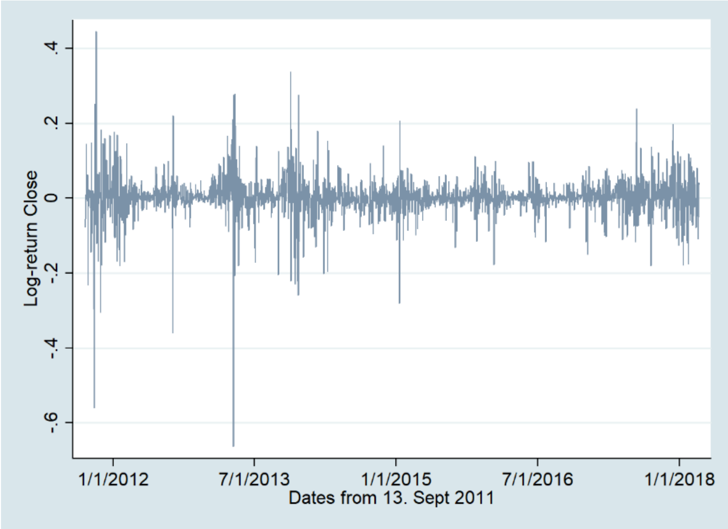


Figure 4 - Log-returns of Bitstamp's exchange rate

4.3 Statistical properties

We will not compare the log-returns to the USD exchange rate towards other major currencies or cryptocurrencies as done in many studies on cryptocurrencies in general, but rather see how Bitcoin's own exchange rate have changed with respect to comparable results found earlier. By performing the descriptive summary command in Stata, we obtain most of the same comparable statistics as presented in the paper by Chu, Nadarajah, and Chan. We believe the numbers are from a reasonable while back, in order to look at the development in the exchange rate. There is some change in most statistics, but it is also worth to notice that the maximum and minimum is unchanged from three years ago. This is also viewable when comparing figure 4 here to figure 1 in their paper.

The global maximum is 0.446, observed on the 28th October 2011, and the global minimum is -0.664, observed on the 11th April 2013. Followed by this, we have that the range of all log-returns is unchanged. Besides this, all other statistics have changed. The first general and probably most important change is that we see a much more concise dataset with less spread and variation.

Statistics	*C, N&C (2015)*	*N&F (2018)*
Minimum	-0.664	-0.664
First quartile	-0.012	-0.012
Median	0.004	0.002
Mean	0.005	0.003
Third quartile	0.025	0.020
Maximum	0.446	0.446
Interquartile range	0.037	0.032
Range	1.109	1.109
Skewness	-1.503	-1.343
Kurtosis	22.425	26.744
Standard deviation	0.069	0.053
Variance	0.005	0.003
Coefficient of variation	15.156	16.629

Table 1 - Summary statistics log-returns of the exchange rate of Bitcoin in 2015 versus 2018

The first and third quartile are closer to zero, respectively 0.0002 larger and 0.0046 smaller. As a result of this, the interquartile range is ~ 0.005 smaller, because there is less difference between the quartiles. The median and mean are both 0.002 smaller, such that we are even closer to a dataset with zero mean. We notice that the skewness is smaller, with an increase of 0.138. This increase indicates that the log-returns are less negatively skewed with the addition of the newer observations but is still left-skewed. Oppositely, we see that the kurtosis has grown considerably, by 4.669. This indicates that the “peakedness” of the distribution is even more substantial than before. It is not so strange as figure 4 shows the peaks are less volatile and different. Given this increase, it must be that we have even more observations clustered around the mean. The lower standard deviation and variation is again strengthening the indication of the observations’ log-return being drawn towards zero. This is also to be expected after looking at the log-return plot and in light of our already mentioned statistics. The coefficient of variation, or measurement of relative variability, is observed to have an increase of 1.473. This increase is despite both the standard deviation and the mean becoming smaller, so the reason of the increase in the coefficient of variation must be that the decrease in the mean of the observations is more sizeable relative to the decrease in the standard deviation.

For us to be able to estimate the parameters of the statistical models properly, it is necessary that the data possesses the property that all observations are *independently* and *identically distributed (iid)*. The assumption of *iid* must be maintained by the data in order to perform the fitting process, which we will explain in chapter 5 and 6.

As we are able to test only for randomness as tests with numerical results, these results will be presented in in chapter 6.1 alongside with the conclusion of the evaluation of observations being *iid*. The two other tests will be discussed in the following chapters.

4.4 Autocorrelation/Serial correlation

We will first study the autocorrelation function (acf) and partial autocorrelation function (pacf) of the log-returns of the exchange rate to check for correlation between the observed values in the dataset. This is supposedly a straight-forward process, but when we executed the acf-command on the log-return variable in Stata, we see very clearly that the plot is very different from what we were to expect from earlier findings. The only official research that has included this in the paper is the article by Chu, Nadarajah, and Chan, at least to our knowledge, and we will compare to what they have found. If their acf in figure 2 is compared to figure 5, the plotted function we got as output by the command ac, we begin to wonder whether or not we are studying the same phenomenon. There are three main issues about this comparison.

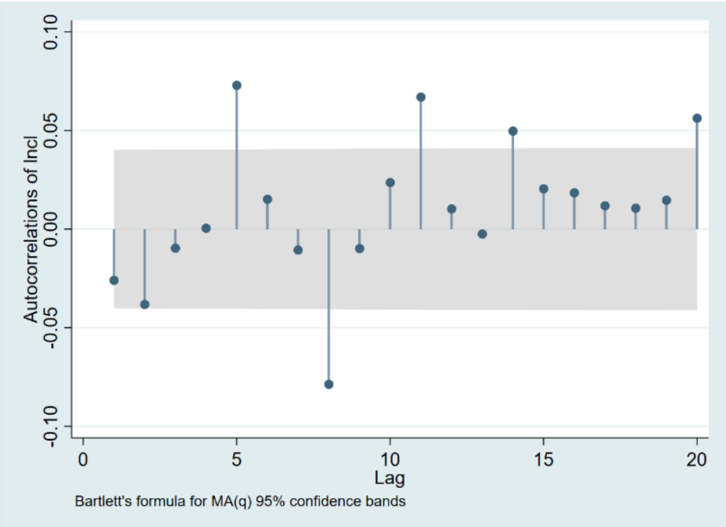


Figure 5 - Autocorrelation plot of bitcoins log-return by 20 lags

Firstly, the Bartlett's formula for $MA(q)$ 95% confidence bands is differing from $\sim \pm 0.04$ in our plot against $\sim \pm 0.21$ in the article's plot. It is not explicitly mentioned that they are using any other significance level that 5%, but even if lowered to 0.01%, the band would only be $\sim \pm 0.075$. This would remove any significant points outside the band, but still also differ from the compared plot. The reason may be that they have utilised another type of autocorrelation function with some properties, but if this is the case it is not mentioned explicitly in the article.

Secondly, the sizes of the autocorrelation at all lags, relative to the confidence interval, are all larger such that we do have five significant points outside the band. We observe these significands at time lag 5, 8, 11, 14, and 20. There is no reason to believe that these points would actually have any significant impact in particular, as they are seemingly pulled at random. Why would the price of Bitcoin five days ago have any impact on today’s price in a positive direction, while the price eight days ago would affect it negatively, and at the same time the four last days would be insignificant? On the other hand, it would not be directly unreasonable to assume the prices to be correlated to some degree in terms of time, but this would be easier to remark if the plot were to be more structured.

Thirdly, the direction of the autocorrelation at each lag is inconsistent with the ones of Chu, Nadarajah, and Chan. There is not much more to comment on this than implying the autocorrelation functions looks like they are from two completely different datasets, which is strange given we have exactly the same log-return plot and similar statistics.

Although significant lags, they are apparently random at some length. No clear pattern in lags is observed in the plot, except for possibly the four first autocorrelated lags of arising at a regular basis of three-days intervals. This is complimented by an almost significant point at lag two, which could, with some effort, be thought to have some pattern in fluctuations in the pattern low, high, low, high, high at the successive lags of 2, 5, 8, 11, and 14, followed by five non-significant days and a high at lag 20. Most likely will this prove to be occurred by chance. Despite of this is it arguable to actually say there are any correlations between time-lags in general. We will probably treat the dataset in principle as a white noise process.

There is not much research done in the field regarding the autocorrelation of Bitcoin, but we have found on the online forum stackexchange.com an anonym user asking for how to interpret serial autocorrelations and AR models (user3783846, 2018). Here is the plot for the autocorrelation first

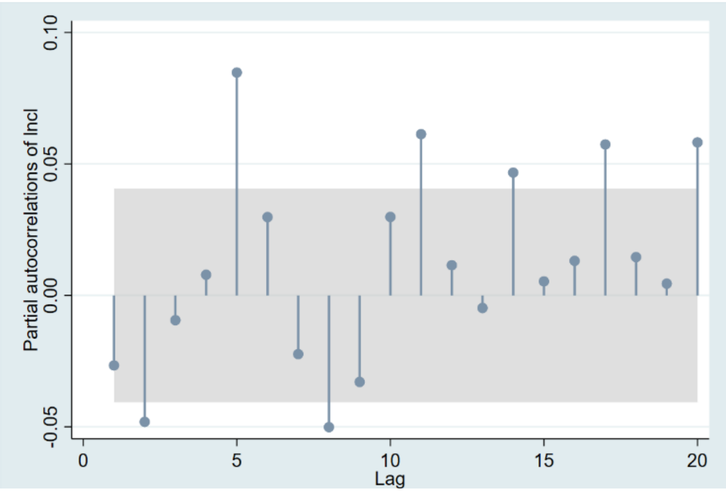


Figure 6 - Partial autocorrelation plot of bitcoins log-return by 20 lags

presented, and even if it includes 35 lags instead of 20 as we have done, it resembles our autocorrelation function quite reasonably with regard to the issues pointed out above. From the code provided at the end of the text, we see that the sample that has been referred to in the text, is exactly the same we have used from Quandl, providing Bitstamp's exchange rate. There are however some immediate problems regarding our comparison of these correlograms. The difference in number of included lags, 20 to 35, and the observed timespan, which is said to be over five years, should be considered when evaluating.

The pacf plotted in figure 6, and it supports our claims above. As the linear dependence is removed, two more significant lags are observed. We believe these plots are reflecting reality well, because it shows some serious impact on how observations are correlated, as expected from an unpredictable and volatile phenomenon. It may be that this simple autocorrelation function does not comprehend and explain the dynamics of Bitcoin properly, but because there is no clear or reasonable pattern in the autocorrelation, and we require the log-returns to be non-correlated, will this be left as it is for now.

4.5 Randomness

To test the randomness of observations in the log-return of the exchange rate, we will perform a series of tests, as done in the article by Chu, Nadarajah, and Chan. The seven tests we are to utilise are Bartels rank test (Bartels, 1982), Cox and Stuart's sign test (Cox & Stuart, 1955), the difference sign test (Brockwell & Davis, 2002), the Mark-Kendall rank test of statistical independence (Mann, 1945) (Kendall, 1955), Wald and Wolfowitz' test for independence and stationarity (Wald & Wolfowitz, 1943), the turning point test (Bienaymé, 1874), and the Ljung-Box test (Ljung & Box, 1978).

4.5.1 Test statistics and hypotheses

In all the properties and statistics are n the number of relevant observations in the data series, μ the expected value of the test statistic, and σ the standard deviation of the test statistic.

Bartels' test is a rank version of the traditional von Neumann's ratio test of 1941. To test for randomness, we have to use the two-sided version of the test, whereas the null hypothesis is randomness against the alternative hypothesis of the dataset being characterised by non-randomness. The test statistic calculated is $RVN = \frac{\sum(R_i - R_{i+1})^2}{\sum(R_i - \bar{R})^2}$, where $R_i = rank(X_i)$, $i = 1, 2, 3, \dots, n$.

Cox and Stuarts test is a sign test applied to data that are grouped in pairs with the *i*-th observation to its corresponding *j*-th observation in each half of the time-ordered data. By using the two-sided test, we check the null hypothesis of randomness against the alternative hypothesis of the dataset showing an upward or a downward trend. The test statistic calculated is $S = \sum w_{ij} h_{ij}$, where h_{ij} is a comparison between the *i*-th and *j*-th observation in the subseries and w_{ij} is the appropriate weight. We will always have $i < j$. We define the comparison variable as $h_{ij} = \begin{cases} +1 \ \ if \ \ y_i > y_j \\ 0 \ \ \ if \ \ y_i < y_j \end{cases}$. The number of observations is N, but in the test, we have that $n = \frac{N}{2}$ (rounded down if N is an uneven number) because of pairing, which in our case means that we have $n = 1185$ terms that are being tested for randomness.

The difference-sign test of statistical independence tests for if any successive differences in a data series is showing any sign for a trend. It tests the null hypothesis of the data series being *iid* against the alternative hypothesis of the data not being *iid*. The test statistic calculated is $D = \frac{(pd-\mu)}{\sigma}$, where pd is the number of positive differences in the data series, $\mu = \frac{(n-1)}{2}$, and $\sigma = \sqrt{\frac{(n+1)}{12}}$. This test is not referred to in our reference article.

The Mann-Kendall rank test of randomness tests for trend based on the number of increasing ordered pairs in a data series. It tests the null hypothesis of the data series being *iid* against the alternative hypothesis of the data not being *iid*. The test statistic calculated is $R = \frac{(pairs-\mu)}{\sigma}$, where $pairs$ is the number of increasing pairs in the data, $\mu = n\frac{(n-1)}{4}$, and $\sigma = \sqrt{n\frac{(n-1)(2n+5)}{72}}$.

Wald and Wolfowitz' runs test tests the null hypothesis that all observations in any sequence of the set is drawn independently from the same distribution and induces randomness. The alternative is that we have either different distributions, a trend, or dependent drawings of elements. Here we also use the two-sided test to test for randomness in general, and not only for trend. The test statistic calculated is $z = \frac{(R-E(R))}{\sqrt{V(R)}}$, with $R = \sum_{\alpha=1}^{N-1} x_\alpha x_{\alpha+1} + x_N x_1$, $E(R) = \frac{(S_1^2 - S_2)}{(N-1)}$ and $V(R) = \frac{S_2^2 - S_4}{N-1} + \frac{S_1^4 - 4S_1^2 S_2 + 4S_1 S_3 + S_2^2 - 2S_4}{(N-1)(N-2)} - E(R)^2$, with $S_r = \sum_{i=1}^{N} x_i^r$, $r = 1, 2, 3, 4$.

The turning point test tests for a data series' independence by comparing turning points in the series to the expected number as in an *iid* series. A turning point is found where an observation is either higher or lower than its corresponding observations, i.e. $x_{i-1} < x_i \ and$

$x_i > x_{i+1}\ or\ x_{i-1} > x_i\ and\ x_i < x_{i+1}$. It tests the null hypothesis of the data series being *iid* against the alternative hypothesis of the data not being *iid*. The test statistic calculated is $T = \frac{(tp-\mu)}{\sigma}$, where tp is the number of turning points present in the data series, $\mu = 2\frac{(n-2)}{3}$, and $\sigma = \sqrt{\frac{(16n-29)}{90}}$.

Ljung and Box' test is a parametric test that tests if any of a group of autocorrelations in a time series are different from zero. This test is testing randomness overall instead of randomness at every lag. Because this test has a very clear null hypothesis, but also a rather unspecified formulated alternative hypothesis, it is considered a portmanteau test. The null hypothesis for Ljung and Box' test tests the null hypothesis of the data series being *iid*, which implies zero correlation between samples and thus randomness, against the alternative hypothesis of the data not being *iid*, which implies there exists serial correlation. The test statistic calculated is $Q = n(n+2)\sum_{k=1}^{m}\frac{\hat{\rho}_k^2}{n-k},\ k = 1, 2, \dots, m, \dots$, where $\hat{\rho}_k^2$ is the estimated autocorrelation at lag k, or $\hat{\rho}_k = \frac{\sum_{t=k+1}^{n}\hat{e}_t\hat{e}_{t-k}}{\sum_{t=1}^{n}\hat{e}_t^2}$, where $\hat{e}_{1,}\dots,\hat{e}_n$ is the standardised residual, and m is the number of tested lags. Under the null hypothesis the test statistic follows chi-squared distribution with the degrees of freedom accounting for the number of included parameters in the model.

We also chose the Ljung-Box test in favour of the Box-Pierce test, because it is shown by earlier subsequent simulations to provide better statistics than the former one. It is also not too different, which makes it comparable to Box-Pierce. Also, while it is true that this test is parametric (in contrast to what claimed by Chu, Nadarajah, and Chan), it is actually usable in our case. This is because it normally would be required to test the residuals from an already fitted $ARMA(p,q)$ model, which are assumed of being a white noise process. It must be applied to a time series that is stationary, which our dataset also is assumed to be.

The estimated tests' p-values are presented in 6.1.

4.6 Dynamic moments

To further strengthen our assumptions for *iid*, we will now look at the development of the dataset's accumulated second, third, and fourth moment, namely its variance, skewness, and kurtosis. This is done to first make out some idea of whether there is any change in the variance over time, and thus decide whether the findings indicate the dataset is homo- or heteroskedastic, for then to establish some assumptions about the not-yet-selected distribution and its behaviour. In the article by Chu, Nadarajah, and Chan is the variance's development tested for by performing the Breusch-Pagan test for heteroskedasticity for linear regression models. Our issue with this is that they have not specified any regressive model to apply this test to. Instead of attempting to perform a replication of this performance, and since this test is only applicable as a post-regressive test, will we only evaluate the cumulative variance plot as an alternative to the test in regard to consider if we will come to the same conclusion of not being able to reject homoskedasticity in the dataset.

The p-value in Breusch-Pagans' test was estimated to be 0.403, which indicates strongly that we do not reject the tests' null hypothesis of homoskedasticity in the log-returns. In figure 7, we show the cumulative variance in our dataset, or, more precisely, how the log-return's variance develops and changes over the whole time-period. The first 30 days were burned out in order to remove extreme values and changes over the first period, and to have a more reliable plot to study.

Figure 7 - Cumulative variance

As expected, the first period of two years offers a relatively high variance compared to our value of reference presented under statistical properties (0.003). At the highest, 29th October 2011, the variance peaked at 0.021, seven times higher than the latest observed variance. The variance declined quickly after this, and, despite the two small jumps in April and December 2013, there has been a gradual downward trend overall. Surprisingly, it seems that it has more or less stabilised at a variance of 0.003 from around the beginning of 2016, ending today at the value of 0.0027943. For the sake of comparison, we also have the variance dated to the 14th May 2015 at 0.0047217, which is consistent with Chu, Nadarajah, and Chans presented result at 0.005. It is some variation over this time-period, but not much of significance. This would imply that, even though the variance may have been heteroskedastic earlier, the dataset is most likely to be described as homoscedastic. This is expected and in terms of our previous findings and is actually closing the gap between the exchange rate and that of other currencies, although it is still far to go.

In the following figures, 8 and 9, are the cumulative skewness and kurtosis plotted, respectively. The kurtosis is not plotted as excess kurtosis. In these plots we see more varying motion in those moments than in the variance. Both plots are consistent with our findings of a skewness of -1.343 and a kurtosis of 26.744 today.

Figure 8 - Cumulative skewness

By first studying the skewness' development and disregarding the first five observations that are less than -2, we see that there is no clear trend over the whole period. There is however three time-windows that show some periodical trends. Over the first one and a half year it is a downward trend from -0.5 to -1.3, then a sudden drop in mid-2013 followed by a one-and-a-half-year stable skewness at -2 before an increase in 2014. We can make out that it was kept stable around -1.5 from 2014 until mid-2016 and had a small decline before increasing somewhat a year later and through until today. What is common for the cumulative skewness over six years, is that all values are negative. This indicates that our data are left-skewed, and that the mass of observations are located to the right in the distribution rather than in the middle or to the left. Eventual extreme outliers will also assumedly be found more often to the left in the distribution than to the right.

Figure 9 - Cumulative kurtosis

The kurtosis does not seem to have any very extreme shifts in its early periods as the other moments have shown. Starting at 15 and hovering just above 10 for the first months, increasing sharply until 2013. Here we also observe an interesting part with sudden decline from 25 to 22 and a harsh rise to above 30 over just three days, before being followed by more unspecified patterns. There was an increase from 2014 and on until the start of 2017, when it was kept at roughly 30 and has declined to 27, where it is at today. Because the kurtosis is very high compared to the one of a normal distribution, it is called leptokurtic. What follows this observation is that we might expect the distribution to be heavy-tailed rather than light-tailed. The distribution is likely to be peaked and thus have most observations in the centred area, but also have quite a few more outliers than in a normal distribution.

Now that the statistical properties and characteristics of the data has been established, will we in the next part describe the methods the exchange rate of Bitcoin will be fitted to and evaluated by.

5. Re: Distributions fitted

5.1 Distribution candidates

We will, in this section, present the statistical distribution candidates that will be fitted to model the log-return of the exchange rate of Bitcoin. In the following distributions, X will be the denotation of a continuous random variable of mentioned log-returns. $f(x)$ and $F(x)$ will represent the probability density function (pdf) and the cumulative distribution function (cdf) of the variable X, respectively. While it has been shown that the generalized hyperbolic distribution gives a good fit of the data, the analysis will be redone and re-evaluated because of our previous findings of dissimilarities between the data, and to check if the results are still consistent. We have chosen to omit five of the proposed distributions Chu, Nadarajah, and Chan included in their paper, because the selective criteria were found to be pretty much similar relative to each other. The specifications of the pdfs of the distributions that are chosen to be included in this thesis are as following:

I. Normal distribution (Gauss, 1809):

$$f(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

for $x \in \mathbb{R}, \mu \in \mathbb{R}$ and $\sigma > 0$;

II. Logistic distribution (Johnson & Kotz, 1970):

$$f(x;\mu,\sigma) = \frac{e^{-\frac{x-\mu}{\sigma}}}{\sigma\left(1 + e^{-\frac{x-\mu}{\sigma}}\right)^2}$$

for $x \in \mathbb{R}, \mu \in \mathbb{R}$ and $\sigma > 0$;

III. Laplace distribution (Laplace, 1774):

$$f(x;\mu,\sigma) = \frac{1}{2\sigma} e^{-\frac{|x-\mu|}{\sigma}}$$

for $x \in \mathbb{R}, \mu \in \mathbb{R}$ and $\sigma > 0$;

IV. Student's *t* distribution (Gosset, 1908):

$$f(x;\mu,\sigma,v) = \frac{\mathrm{K}(v)}{\sigma}\left[1 + \frac{(x-\mu)^2}{\sigma^2 v}\right]^{\frac{-(1+v)}{2}}$$

for $x \in \mathbb{R}, \mu \in \mathbb{R}, \sigma > 0$ and $v > 0$, where $\mathrm{K}(v) = \sqrt{v}\,\mathrm{B}(\frac{v}{2},\frac{1}{2})$, and

$\mathrm{B}\left(\frac{v}{2},\frac{1}{2}\right) = \int_0^1 t^{\frac{v}{2}-1}(1-t)^{-\frac{1}{2}}\,dt$;

V. Exponential power distribution/generalised normal distribution (Subbotin, 1923):

$$f(x;\mu,\sigma,\beta) = \frac{\beta}{2\sigma\Gamma(1/\beta)} e^{-\left(\frac{|x-\mu|}{\sigma}\right)^{\beta}}$$

for $x \in \mathbb{R}, \mu \in \mathbb{R}, \sigma > 0$ and $\beta > 0$, where $\Gamma\left(\frac{1}{\beta}\right) = \int_0^\infty t^{\frac{1}{\beta}-1} e^{-t}\,dt$ is the gamma function;

VI. Skew normal distribution (Azzalini, 1985):

$$f(x;\mu,\sigma,\lambda) = \frac{2}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)\Phi\left(\lambda\frac{x-\mu}{\sigma}\right)$$

for $x \in \mathbb{R}, \mu \in \mathbb{R}, \lambda \in \mathbb{R}$ and $\sigma > 0$, where $\phi(\cdot)$ and $\Phi(\cdot)$, denote the pdf and the cdf of the standard normal distribution, respectively;

VII. Skew *t* distribution (Hansen B. E., 1994):

$$f(x;\mu,\sigma,\lambda,v) = \frac{\mathrm{K}(v)}{\sigma}\left[1 + \frac{(x-\mu)^2}{\sigma^2 v}\right]^{\frac{-(1+v)}{2}} + \frac{2\mathrm{K}^2(v)\lambda(x-\mu)}{\sigma^2}\ {}_2F_1\left(\frac{1}{2},\frac{1+v}{2};\frac{3}{2};-\frac{\lambda^2(x-\mu)^2}{\sigma^2 v}\right)$$

for $x \in \mathbb{R}, \mu \in \mathbb{R}, \lambda \in \mathbb{R}, \sigma > 0$ and $v > 0$, where ${}_2F_1(a,b;c;x) = \sum_{k=0}^{\infty}\frac{(a)_k(b)_k}{(c)_k}\frac{x^k}{k!}$, where $e_k = e(e+1)\ldots(e+k-1)$ denotes the ascending factorial;

VIII. Generalised *t* distribution (McDonald & Newey, 1988):

$$f(x;\mu,\sigma,v,\tau) = \frac{\tau}{2\sigma v^{1/\tau}\mathrm{B}(v,1/\tau)}\left[1 + \frac{1}{v}\left|\frac{x-\mu}{\sigma}\right|\right]^{-(v+1/\tau)}$$

for $x \in \mathbb{R}, \mu \in \mathbb{R}, \sigma > 0, v > 0$ and $\tau > 0$;

IX. Hyperbolic distribution (Barndorff-Nielsen, 1977):

$$f(x;\mu,\alpha,\beta,\delta)=\frac{\sqrt{\gamma}\,e^{\beta(x-\mu)}}{\sqrt{2\pi\alpha}\,\delta\,K_1(\delta\gamma)}\,\frac{K_{1/2}\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)}{\left(\sqrt{\delta^2+(x-\mu)^2}\right)^{-\frac{1}{2}}}$$

for $x\in\mathbb{R},\mu\in\mathbb{R},\delta>0,\alpha>0$ and $\beta>0$, where $\gamma=\sqrt{\alpha^2-\beta^2}$ and $K_\nu(\cdot)$ denotes the modified Bessel function of the second kind of order ν defined

$$K_\nu(x)\;=\begin{cases}\frac{\pi\csc(\pi\nu)}{2}[I_{-\nu}(x)-I_\nu(x)], & if\ \nu\notin\mathbb{Z}\\ \lim_{\mu\to\nu}K_\mu(x), & if\ \nu\in\mathbb{Z}\end{cases}$$

where $I_\nu(\cdot)$ denotes the modified Bessel function of first kind of order ν defined

$$I_\nu(x)=\sum_{k=0}^{\infty}\frac{1}{\Gamma(k+\nu+1)k!}\left(\frac{x}{2}\right)^{2k+\nu};$$

X. Generalised hyperbolic distribution (Barndorff-Nielsen, 1977):

$$f(x;\mu,\alpha,\beta,\delta,\lambda)=\frac{(\alpha^2-\beta^2)^{\lambda/2}\;e^{\beta(x-\mu)}}{\sqrt{2\pi}\,\alpha^{\lambda-\frac{1}{2}}\,\delta^\lambda\,K_\lambda(\delta\gamma)}\,\frac{K_{\lambda-\frac{1}{2}}\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)}{\left(\sqrt{\delta^2+(x-\mu)^2}\right)^{1/2-\lambda}}$$

for $x\in\mathbb{R},\mu\in\mathbb{R},\lambda\in\mathbb{R},\delta>0,\alpha>0$ and $\beta>0$.

Most of these distributions are related or similar in many ways. If we, for instance, consider the family of skewed generalised *t* distributions, illustrated in picture 5 as done in the paper by Hansen, McDonald, and Newey (2010), it is clear that there are eight special cases that can be derived from the skewed generalised *t* distribution, which is a five-parametric distribution described by Theodossiou (1998) to be

$$f(x;\mu,\sigma,\lambda,p,\nu)=p\left[2c\sigma\nu^{\frac{1}{p}}B\left(\frac{1}{p},\nu\right)\left(\frac{|x-\mu+m|^p}{\nu(c\sigma)^p(\lambda sign(x-\mu+m)+1)^p}+1\right)^{\frac{1}{p}+\nu}\right]^{-1},$$

for $x\in\mathbb{R},\mu\in\mathbb{R},\lambda\in\mathbb{R},\sigma>0,\mathrm{p}>0,\nu>0$, and where *B* is the beta function. We also have that $c=\nu^{-\frac{1}{p}}\left[\sqrt{(3\lambda^2+1)\frac{B\left(\frac{3}{p},\nu-\frac{2}{p}\right)}{B\left(\frac{1}{p},\nu\right)}-4\lambda^2\frac{B\left(\frac{2}{p},\nu-\frac{1}{p}\right)^2}{B\left(\frac{1}{p},\nu\right)^2}}\right]^{-1}$ and $m=\frac{2c\sigma\lambda\nu^{\frac{1}{p}}B\left(\frac{2}{p},\nu-\frac{1}{p}\right)}{B\left(\frac{1}{p},\nu\right)}$, and are not parameters themselves.

This superclass distribution has not been included in previous studies, but we would like to include this as the eleventh distribution, because we believe it is reasonable to see if this five-parametric distribution can provide a better fit than its limited cases. Theodossiou (1998) himself also fitted this distribution to the exchange rate of Canadian Dollar and Japanese Yen to USD. Even though this is not necessary a very much used distribution in previous financial studies, it would be interesting to see how it performs in comparison to the other distributions.

From the skewed generalised *t* distribution we can represent the generalised *t* distribution if we set $\lambda = 0$; skewed student's *t* if $p = 2$; skewed exponential power if $\nu \to \infty$; student's *t* if $p = 2 \; and \; \lambda = 0$; exponential power if $\nu \to \infty \; and \; \lambda = 0$; skewed normal if $\nu \to \infty \; and \; p = 2$; Laplace if $\nu \to \infty, \; \lambda = 0, and \; p = 1$; normal if $\nu \to \infty, p = 2, and \; \lambda = 0$. These connections are illustrated in figure 5 ($q \equiv \nu$).

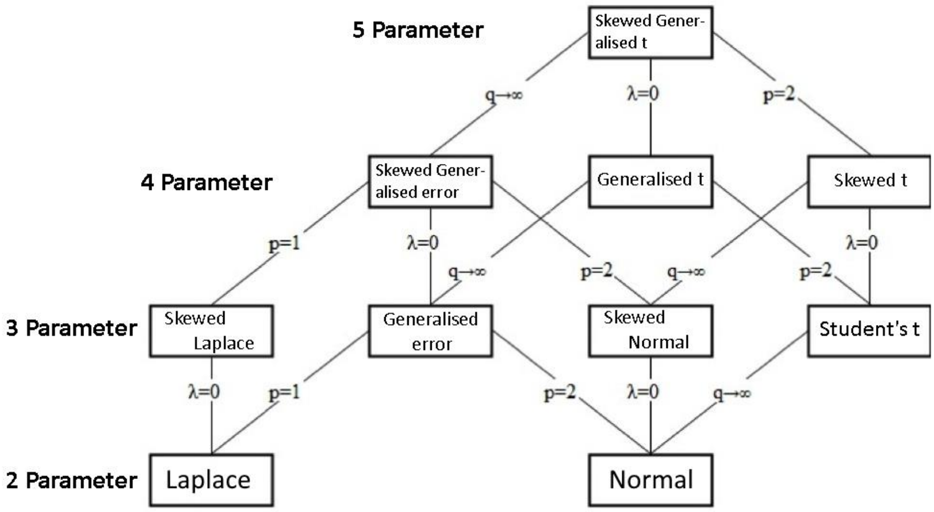


Figure 10 - Family of skewed generalised t distributions (Hansen, McDonald, & Newey, 2010)

Also, we have that the generalised hyperbolic distribution can represent the hyperbolic if $\lambda = 1$. Incidentally, the generalised hyperbolic is originally not a financial distribution, but rather one from physics. In the 70s, Barndorff-Nielsen studied the physics of blown sand, resulting in this particular distribution. It is somewhat commonly used in finance when modelling markets. This shows that it is necessary to retrieve knowledge from other fields in order to describe phenomena that are not possible to describe with known but limiting techniques.

From our set of distributions, there are light tailed distributions represented by the normal, logistic, Laplace, exponential power and skew normal distribution. A set of heavy tailed distributions are also represented by the Student's *t*, skew *t*, generalised *t*, skew generalised *t*, hyperbolic, and generalised hyperbolic distribution. The difference between light and heavy tailed distributions is that the heavy tailed distributions goes to zero much slower than the light tailed. We will experience many more outliers in the heavy tailed distribution that is holding more extreme values than what would normally be expected.

5.2 Maximum likelihood estimation

When estimating the fit of the distributions, we are using the method of maximum likelihood estimation (MLE). This method is used to maximise the likelihood function for the given data. The continuous distributions' likelihood is given by $L(\Theta) = \prod_{i=1}^{n} f(x_i; \Theta)$ or for the log-likelihood, that we require, $\ln L(\Theta) = \sum_{i=1}^{n} \ln f(x_i; \Theta)$, where Θ is the models parameter that maximises the function. Because the true value of the parameter is unknown, an estimate, $\widehat{\Theta}$, is used. We have that $\widehat{\Theta} = (\widehat{\Theta}_1, \widehat{\Theta}_2, \dots, \widehat{\Theta}_k)'$, that is the denoted maximum likelihood estimate. When n is large enough, we have that $\widehat{\Theta}$ converges to its true value Θ_0, which implies consistency with arbitrary precision.

In our process of estimating the MLE, we are using the `mle` estimation procedure in R, included in the basic `stats4` package. The coding is shown in appendix B.

5.3 Model selection criteria

Because some of the selected distributions are nested and others are not, we have to discriminate among them. This is done to be able to select the best distribution of choice regardless of one being very similar to another in other comparative terms. We have selected five criteria and two statistics in addition to the log-likelihood, whereas all of these are commonly used in model selection. In all criteria, the variable k is the number of independently estimated parameters within the model, or the so-called dimension. They are presented accordingly.

I. Akaike information criterion (AIC), defined by Akaike (1974):

$$AIC = 2k - 2 \ln L(\widehat{\Theta})\,;$$

II. Bayesian information criterion (BIC), defined by Schwarz (1978):

$$BIC = k \ln n - 2 \ln L(\widehat{\Theta})\,;$$

III. Consistent Akaike information criterion (CAIC), defined by Bozdogan (1987):

$$CAIC = -2 \ln L(\widehat{\Theta}) + k(\ln n + 1)\,;$$

IV. Corrected Akaike information criterion (AICc), defined by Hurvich and Tsai (1989):

$$AICc = AIC + \frac{2k(k+1)}{n-k-1}\,;$$

V. Hannan-Quinn criterion (HQC), defined by Hannan and Quinn (1979):

$$HQC = -2\ln L(\widehat{\Theta}) + 2k\ln\ln n\,;$$

VI. Kolmogorov-Smirnov statistic (KS), defined by Kolmogorov (1933) and Smirnov (1948):

$$KS = \sup_{x}\left|\frac{1}{n}\sum_{i=1}^{n} I_{[-\infty,x]}(X_i) - \Phi(x)\right|,$$

where *sup* is the supremum of the set of distances, $I_{[-\infty,x]}(X_i)$ is the indicator function, and $\Phi(x)$ is the MLE of $F(x)$, the cumulative distribution function;

VII. Anderson-Darling statistic (AD), defined by Anderson and Darling (1954):

$$AD = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\{\ln\Phi(x_i) + \ln[1-\Phi(x_{n-i+1})]\},$$

where $x_1 \leq x_2 \leq \cdots \leq x_n$ are the observed dataset in a sequence of increasing values.

The two last statistics, KS and AD, are different from the criteria. These will be evaluated by their respective p-values, which should be larger the higher chance the data are drawn from the distribution that is evaluated.

The log-likelihood is working somewhat different than the criteria because it is better used to discriminate amongst the distributions that are nested. The test, which is called the likelihood ratio test, was defined by Wilks (1938) and Cox and Hinkley (1974), and is performed for discrimination of distributions as follows:

Suppose we have two models – an alternative model that has a set of parameters, k_1, and a null model that has a set of parameters, k_2, where $k_2 < k_1$. The null is a particular case of the alternative. Their log-likelihood is denoted L_1 and L_2, respectively. We can estimate the test statistic D by the double of the difference in log-likelihoods:

$$D = -2\ln\left(\frac{L_2}{L_1}\right) \gtreqless \chi^2_{k_1-k_2,0.05},$$

where the degrees of freedom of the chi-square is the difference in number of parameters between the models and the significance level is chosen to be 5%. We use the $\chi^2_{r,\alpha}$ distribution to discriminate by choosing the alternative model if the chi-square is smaller than the test statistic, because if tested, the log-likelihood under the null (no difference in the models) is distributed like chi-square with $k_1 - k_2$ degrees of freedom (Wilks, 1938).

When presenting the values of the log-likelihood and the criteria in part 6.4, it is the distribution with the lowest values that provides the best modelling of the data. For the two statistics, that are evaluated by their respective p-values, it is the highest p-value that indicates the best modelling of the data. By using these criteria, we should have substantial proof to back our selection of distribution.

5.4 Value at risk and expected shortfall

After the best fitted distribution is selected will this distribution be assessed in an evaluation of the potential financial risk one is exposed to if one acquires bitcoins and holds them for a set period of time. We will in this evaluation look at two distinguished measures, the value at risk (VaR) and expected shortfall (ES). VaR is a model-dependent measure which is based on quantiles of a distribution and explained thoroughly and in context by Artzner et.al. (1998). ES, as proposed by Rockafellar and Uryasev (2000), uses the VaR to set a point on the distribution for so to measure the area below the point. As ES is a function of VaR is it natural to present them accordingly:

$$VaR_\zeta(X) = -\inf\{x \in \mathbb{R}: F_X(x) > \zeta\},$$

where $\zeta \in (0,1)$ is the confidence level set and F_X is the cdf of variable X. ζ is also setting the ζ^{th} quantile of the random variable. From this we can define the ES:

$$ES_\beta = -\frac{1}{\beta}\int_0^\beta VaR_\zeta(X)d\zeta = -\frac{1}{\beta}\Big(E\left[X\ 1_{\{X\leq x_\beta\}}\right] - x_\beta\big(P[X \leq x_\beta] - \beta\big)\Big),$$

where $0 < \beta < 1$, $x_\beta = \inf\{x \in \mathbb{R}: P(X \leq x) \geq \beta\}$ is the lower β-quantile and $1_B(x) = \begin{cases} 1\ if\ x \in B \\ 0\ else \end{cases}$ is the indicator function.

We will perform these measurements with different values of ζ so that we can get some indication of possible outcomes both in normal, extreme or highly unlikely (but possible) situations.

Now that we have presented the necessary prerequisites, we will continue in the next chapter by presenting our findings and results in terms of what we have presented above. These will also be discussed consecutively.

6. MLE estimations

All codes for the estimation procedures are shown in the appendix B.

6.1 Results of randomness tests

All p-values of the randomness tests are shown in table 2 as a comparison to the article by Chu, Nadarajah, and Chan. This is done to study the development in the log-returns of the exchange rate. All tests, except the Ljung-Box test, are non-parametric, as we have no assumption about distribution for now. All the tests are performed in R, with the tests Bartels, Cox-Stuart, difference sign, rank and turning point from the package `randtests` (Caeiro & Mateus, 2014), the Wald-Wolfowitz from the package `spgs` (Hart & Martínez, 2017), and Ljung-Box from the package `trend` (Pohlert, 2018).

The test statistics' p-values that are tested for here are summarized in the following table:

Test	*C, N&C (2015)*	*N&F (2018)*
Bartels (1982)	0.123	0.203
Cox and Stuart's sign (1955)	0.613	0.601
Difference sign	0.238	0.176
Rank	0.352	0.840
Wald and Wolfowitz' run (1940)	0.243	0.337
Turning point	0.129	0.188
Ljung and Box (1978)	0.302 (B&P)	0.324

Table 2 – P-values from the tests for randomness

First, we would like to digress and point out a matter regarding the Wald and Wolfowitz' test. It was, in the compared article, referred to Wald and Wolfowitz' paper, "*On a test whether two samples are from the same population"* (1940). This test and the one we have referred to in part 4.5 are in fact two quite different tests in terms of what they are actually testing for. In their tests from 1940, Wald and Wolfowitz are testing if two different independent variables are taken from the same distribution, while in the test from 1943 it is tested whether a sequence of variates satisfy the condition of randomness – which is exactly what we actually wish to test for. It is not mentioned explicitly how this test is performed by Chu, Nadarajah, and Chan, or what their test statistics are, so the procedure is uncertain. It may have been a mis-reference, a package or command in the program they got their results from giving both tests' results, that they deemed this test for being fit in this context, or just the wrong test performed. We can only speculate, but it is so that we have used the randomness test of 1943 nonetheless, and the comparison of those results may be somewhat erred.

All these tests are pointing towards the same conclusion of the log-returns being not significantly different from randomness. This have strengthened our belief of the dataset being iid. When comparing, we observe most p-values being relatively similar in both cases, but in particular the Cox and Stuart's sign test and the rank test stand out for being much more different – considerably more than 0.1 points in difference. The reason behind this is debatable, but both the Cox-Stuart test and the rank test do suggest that our more recent dataset has much more tendencies towards randomness than the shorter dataset. The other tests, except for Bartels and the difference sign test, also indicate this in our favour. Ergo, time is an insignificant factor for each observation.

By comparing all our findings on the log-returns of Bitcoin, both in terms of autocorrelation, randomness and heteroskedasticity, we will stand assured that the dataset does not have serial correlation and is both random and homoscedastic. This means that we can assume the data being independent and identically distributed (*iid*), because it lacks evidence against it. All these properties are necessary to be fulfilled for us to be able to attempt to fit a distribution to the exchange rate. Because of the behaviours of exchange rates in general, of them normally being skewed and heavy-tailed, we believe that there is needed a distribution that is somewhat more advanced than, say, a simple two-parametrical distribution as the normal or logistic. This was for instance done in a paper by Corlu & Corlu (2015), where they attempted to fit nine different currencies' exchange rate to the USD through four different flexible distributions, whereas one of these will be used in our fitting process in the next part. Their paper concluded that the generalised lambda distribution can be a good alternative for modelling.

6.2 Estimations of candidate distributions

It was necessary to utilise different packages and approaches to some distributions because of flexibility, possibilities in syntaxes and to get the right parameterisation.

We compared a restricted part of our dataset, set to 951 observations, to compare our method of work to the results presented in the article by Chu, Nadarajah, and Chan. This has proven in most cases to be sufficient, although some differences have shown to be substantial, both when we have given the start values of the MLE procedure to be the initials as in a standard case in each distribution, for instance in the normal distribution with a mean of 0 and standard deviation of 1, and when the start values were the presented results in the reference article. By observing this, we will have to emphasise that there is probably some difference in our approach to the estimation procedure and may therefore result in some estimations being very different from each other, even though the consistency of the estimated parameters is abided. We may be sure of the consistency because the log-likelihoods we got from our estimations are pretty much similar, deviating only by $+10$ to -12 at most in log-likelihoods at values at around $-1{,}500$. To clarify, similarity was only desirable when comparing the restricted dataset, not when estimating with the full dataset.

There are several possible explanations for this deviation. It is mentioned in the article when presenting the maximum likelihood estimation that *"[t]he maximization was performed using the routine `nlm` in the R software package"*. This routine is different than our previously mentioned routine, `mle`, in terms of field of application. While `nlm` is an all-purpose routine in minimisation/maximisation processes that needs a thorough and explicit syntax in order to run the procedure, `mle` is, as its name suggests, a routine that is optimised for just maximum likelihood estimation. Both functions are included in two fundamental packages in R, `nlm` is found in the `stats` package while `mle` is found in the `stats4` package. There is nothing that implies any of the functions to be miss-coded in the source. We choose to believe that the method we have selected, that is optimised for MLE, is the method that provides the best and most accurate results for the procedure we are following.

Parameter estimations from all distributions are presented in table 3. They are shown in a comparison to the results from the article by Chu, Nadarajah and Chu, with each parameters' standard deviation in the brackets where it is found. The parameters are only compared to see the difference in distributions as a result of a progression in time, not as an attempt to check whether the results are similar or not.

From these estimates we see that there are many differences in the scaling in most of the estimations of the parameters in each distribution, but there is also much similarity in some cases. For instance, the estimate of the lambda in the skew normal distribution, which is an indicator for the shape of the distribution, was estimated to be practically zero, but we found it to be -0.0069 with a substantially smaller standard deviation of the parameter, while both sigma in the Laplace distribution are almost identical, with only a somewhat smaller standard deviation of the parameter.

Distribution	*C, N&C (2015)*	*N&F (2018)*
Normal	$\hat{\mu} = 4.534 * 10^{-3}(2.228 * 10^{-3})$ $\hat{\sigma} = 6.868 * 10^{-2}(1.581 * 10^{-3})$	$\hat{\mu} = 2.968 * 10^{-3}(1.089 * 10^{-3})$ $\hat{\sigma} = 5.304 * 10^{-2}(7.688 * 10^{-4})$
Logistic	$\hat{\mu} = 5.391 * 10^{-3}(1.540 * 10^{-3})$ $\hat{\sigma} = 2.892 * 10^{-2}(8.345 * 10^{-4})$	$\hat{\mu} = 3.517 * 10^{-3}(7.810 * 10^{-4})$ $\hat{\sigma} = 2.298 * 10^{-2}(4.104 * 10^{-4})$
Laplace	$\hat{\mu} = 3.753 * 10^{-3}(1.170 * 10^{-3})$ $\hat{\sigma} = 3.804 * 10^{-2}(1.241 * 10^{-3})$	$\hat{\mu} = 2.242 * 10^{-3}(5.580 * 10^{-4})$ $\hat{\sigma} = 4.307 * 10^{-2}(8.831 * 10^{-4})$
Student t	$\hat{v} = 1.389\ (1.026 * 10^{-1})$ $\hat{\mu} = 3.858 * 10^{-3}(9.195 * 10^{-4})$ $\hat{\sigma} = 2.134 * 10^{-2}(1.197 * 10^{-3})$	$\hat{v} = 1.002\ (1.090 * 10^{-4})$ $\hat{\mu} = 3.054 * 10^{-3}(5.551 * 10^{-4})$ $\hat{\sigma} = 4.321 * 10^{-1}(1.445 * 10^{-2})$
Exponential power	$\hat{\mu} = 3.996 * 10^{-3}(1.490 * 10^{-4})$ $\hat{\sigma} = 2.819 * 10^{-2}(1.368 * 10^{-3})$ $\hat{\beta} = 5.871 * 10^{-3}(2.982 * 10^{-2})$	$\hat{\mu} = 1.642 * 10^{-3}(4.252 * 10^{-4})$ $\hat{\sigma} = 4.981 * 10^{-2}(1.627 * 10^{-3})$ $\hat{\beta} = 6.266 * 10^{-1}(2.079 * 10^{-2})$
Skew normal	$\hat{\mu} = 4.534 * 10^{-3}(3.256 * 10^{-1})$ $\hat{\sigma} = 6.868 * 10^{-2}(1.597 * 10^{-3})$ $\hat{\lambda} = 6.006 * 10^{-9}(5.942)$	$\hat{\mu} = 1.415 * 10^{-3}(1.149 * 10^{-3})$ $\hat{\sigma} = 5.290 * 10^{-2}(7.674 * 10^{-4})$ $\hat{\lambda} = -6.858 * 10^{-2}(1.608 * 10^{-2})$
Skew *t*	$\hat{\mu} = 9.774 * 10^{-4}(1.865 * 10^{-3})$ $\hat{\sigma} = 2.133 * 10^{-2}(1.206 * 10^{-3})$ $\hat{\lambda} = 1.639 * 10^{-1}(9.492 * 10^{-2})$ $\hat{v} = 1.379\ (1.015 * 10^{-1})$	$\hat{\mu} = 4.015 * 10^{-3}(9.628 * 10^{-4})$ $\hat{\sigma} = 1.733 * 10^{-1}(2.906 * 10^{-2})$ $\hat{\lambda} = 2.716 * 10^{-2}(2.306 * 10^{-2})$ $\hat{v} = 1.016\ (5.425 * 10^{-3})$
Generalised *t*	$\hat{\mu} = 3.026 * 10^{-3}(1.186 * 10^{-3})$ $\hat{\sigma} = 2.310 * 10^{-2}(3.695 * 10^{-3})$ $\hat{\tau} = 9.471 * 10^{-1}(1.541 * 10^{-1})$ $\hat{v} = 3.042\ (1.423)$	$\hat{\mu} = 2.000 * 10^{-3}(4.721 * 10^{-4})$ $\hat{\sigma} = 5.919 * 10^{-2}(5.189 * 10^{-3})$ $\hat{\tau} = 1.003(7.785 * 10^{-2})$ $\hat{v} = 3.190\ (7.124 * 10^{-1})$
Skewed exponential power	$\hat{\mu} = 4 * 10^{-3}(1.507 * 10^{-4})$ $\hat{\sigma} = 2.812 * 10^{-2}(1.366 * 10^{-3})$ $\hat{p} = 5.842 * 10^{-1}(2.963 * 10^{-2})$ $\hat{\alpha} = 4.936 * 10^{-1}\ (1.298 * 10^{-2})$	$\hat{\mu} = 4.547 * 10^{-3}(3.450 * 10^{-4})$ $\hat{\sigma} = 5.040 * 10^{-2}(8.769 * 10^{-4})$ $\hat{p} = 7.511 * 10^{-2}(1.194 * 10^{-3})$ $\hat{\alpha} = 6.151 * 10^{-1}\ (2.456 * 10^{-3})$
Hyperbolic	$\hat{\mu} = 3.023 * 10^{-3}(6.032 * 10^{-4})$ $\hat{\delta} = 1.068 * 10^{-5}(9.945 * 10^{-3})$ $\hat{\alpha} = 2.628 * 10^{1}(1.207 * 10^{1})$ $\hat{\beta} = 5.185 * 10^{-1}(3.492 * 10^{-1})$	$\hat{\mu} = 1.734 * 10^{-3}$ $\hat{\delta} = 4.335 * 10^{-6}$ $\hat{\alpha} = 3.285 * 10^{1}$ $\hat{\beta} = 6.644 * 10^{-1}$
Skewed generalised *t*		$\hat{\mu} = 4.452 * 10^{-3}(6.384 * 10^{-4})$ $\hat{\sigma} = 5.877 * 10^{-2}(3.138 * 10^{-3})$ $\hat{\lambda} = 5.672 * 10^{-2}(9.101 * 10^{-3})$ $\hat{p} = 9.747 * 10^{-1}(3.202 * 10^{-2})$ $\hat{v} = 3.409(1.637 * 10^{-1})$
Generalised hyperbolic	$\hat{\mu} = 2.948 * 10^{-3}(8.964 * 10^{-4})$ $\hat{\delta} = 1.217 * 10^{-2}(2.578 * 10^{-3})$ $\hat{\alpha} = 7.731(1.517)$ $\hat{\beta} = 3.447 * 10^{-1}(5.186 * 10^{-1})$ $\hat{\lambda} = -1.39 * 10^{-1}(1.112 * 10^{-1})$	$\hat{\mu} = 2.027 * 10^{-3}$ $\hat{\delta} = 1.015 * 10^{-2}$ $\hat{\alpha} = 1.219 * 10^{1}$ $\hat{\beta} = 3.769 * 10^{-1}$ $\hat{\lambda} = -5.255 * 10^{-2}$

Table 3 - Comparison of the fitted distributions: estimated parameters and their standard errors.
Left side has the sample period from 13/09/2011 to 08/05/2014 while the right side is from 13/09/2011 to 01/04/2018.

These observations imply that even though some estimated parameters are consistent in terms of relative scale to each other, many estimations show that the behaviour of the exchange rate has changed during the last four years.

6.3 Model selection criteria and fit

The estimated results for the criteria are presented in table 4.

If the three two-parametric distributions are considered, it is clear that the Laplace distribution is the best choice over the normal and logistic, with the best values in both log-likelihood, AIC, BIC, AICc, CAIC, and HQC, and the highest p-values. In the case of the three three-parametric distributions, we see the same is the case for the exponential power distribution, which is only challenged by the Student's *t*. The skewed normal distribution does not show a good fit, which was to be expected due the poor fit of the normal distribution as well. In the case of the four four-parametric distributions there is a close fit by both the generalised *t* and the skew exponential power distribution, while the skew *t* and hyperbolic are considered inferior to the two former ones. Although they are very similar, it is clear that the generalised *t* is providing the best fit with the lowest -ln L, AIC, BIC, AICc, CAIC, and HQC, while the generalised *t* has the highest AD and skewed exponential power has the highest KS. Lastly, in our two five-parametric cases, the skew generalised *t* distribution is showing lower values in all criteria and higher p-values compared to the generalised hyperbolic distribution.

It is clear, when comparing all distributions collectively, that the distributions with more parameters are most often to be favoured to the ones with fewer parameters because of their flexibility and applicability. Three exceptions are the skewed normal, which gives a poorer fit than both the logistic and Laplace, the skewed *t*, that is approximately as good as the Student's *t* (which makes it worse because of parsimony), but worse than the exponential power, and the

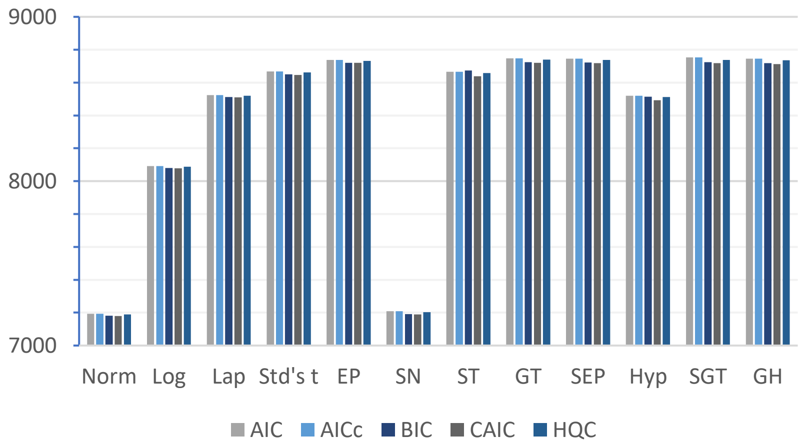


Figure 11 - Bar chart of estimated criteria

Distribution	*-ln L*	*AIC*	*AICc*	*BIC*	*CAIC*	*HQC*	*KS*	*AD*
Normal	-3,598.537	-7,193.074	-7,193.069	-7,181.532	-7,179.532	-7,188.872	~0	~0
Logistic	-4,048.014	-8,092.028	-8,092.023	-8,080.486	-8,078.486	-8,087.826	7.77e-16	~0
Laplace	-4,263.819	-8,523.637	-8,523.632	-8,512.095	-8,510.095	-8,519.435	1.83e-08	4.37e-07
Student's *t*	-4,336.130	-8,667.512	-8,667.502	-8,650.199	-8,647.199	-8,661.210	2.19e-04	2.97e-05
EP	-4,371.797	-8,737.593	-8,737.583	-8,720.280	-8,720.279	-8,731.291	0.00574	0.005787
Skew normal	-3,607.651	-7,209.303	-7,209.293	-7,191.990	-7,188.990	-7,203.001	~0	~0
Skew *t*	-4,337.059	-8,666.118	-8,666.101	-8,674.118	-8,639.034	-8,657.715	0.04161	0.01031e
Generalised *t*	-4,377.849	-8,747.697	-8,747.680	-8,724.613	-8,720.613	-8,739.294	0.05645	0.03114
SEP	-4,376.622	-8,745.243	-8,745.226	-8,722.159	-8,718.159	-8,736.840	0.02133	0.08613
Hyperbolic	-4,264.327	-8,520.654	-8,520.637	-8,513.112	-8,493.570	-8,512.251	1.87e-07	1.80e-06
SGT	-4,381.647	-8,752.942	-8,752.917	-8,724.087	-8,719.087	-8,738.337	0.6309	0.4760
Generalised hyperbolic	-4,378.072	-8,746.144	-8,746.119	-8,717.289	-8,712.289	-8,735.640	0.1089	0.1648

Table 4 - Log-likelihoods, criterions, and p-values for the statistics

hyperbolic, which performs even poorer than the Laplace. The distribution that have proven to provide the over-all best fit to the log-returns of Bitcoin's exchange rate to USD is the skewed generalised *t* distribution. The criteria are compared visually in figure 11.

It is not possible to compare the criteria to any other results directly, because they are dependent on the number of observations in the dataset. If a dataset contains more observations, this will have a greater impact on the criteria. This is reassuring to observe if we actually do compare, because one would expect that a larger number of observations will be able to explain the phenomenon better than with fewer numbers of observations. This is what we have proven to have an example of here. What is actually comparable is the relative difference between the values of the evaluated distributions. We see that the pattern of distributions that provides the poorest fit is similar. The normal and skew normal distributions have the worst fits, followed by logistic, Laplace and hyperbolic, then Student's *t* and skewed *t*. These are distributions with noticeable worse performance than the last five ones, meaning they are pretty much out of the equation in describing the data, unless strict parsimony is desired. We see the same tendencies when comparing, only seeing the hyperbolic performing somewhat worse compared to the Laplace, because of stronger penalties from the larger sample size. Besides this, most observations seem to be relatively similar.

By separating all nine estimated distributions in the skewed generalised *t* family and organise them in groups by number of parameters, there is an apparent difference in how well the distribution is fitted depending on which branch that is looked at. If we consider figure 10, we can see that the most well-fitted distributions are located to the left rather than the right. This would imply that the parameter from the skewed generalised *t* distribution that has the least impact on how well the distribution fits is mainly lambda, the skewness parameter.

The logistic distribution is not nested to any of the other estimated distributions, as it is related to secant distributions instead of skewed generalised *t* or generalised hyperbolic distributions. The generalised hyperbolic distributions are harbouring the student's *t* distributions if we set $\lambda = -\frac{\nu}{2}, \delta = \sqrt{\nu}$ and $\alpha = \beta = 0$, or the Laplace with scale parameter 1 if $\lambda = \alpha = 1$ and $\beta = \delta = 0$, but we disregard these connections when considering nesting, for simplicity. The criteria have discriminated amongst these distributions well.

In the figure 11 is a histogram with the fitted probability density function of the skewed generalised *t* distribution superimposed to show visually how the estimated distribution fits the observed data. In figures 12 and 13 are also the probability plot and the quantile plot included.

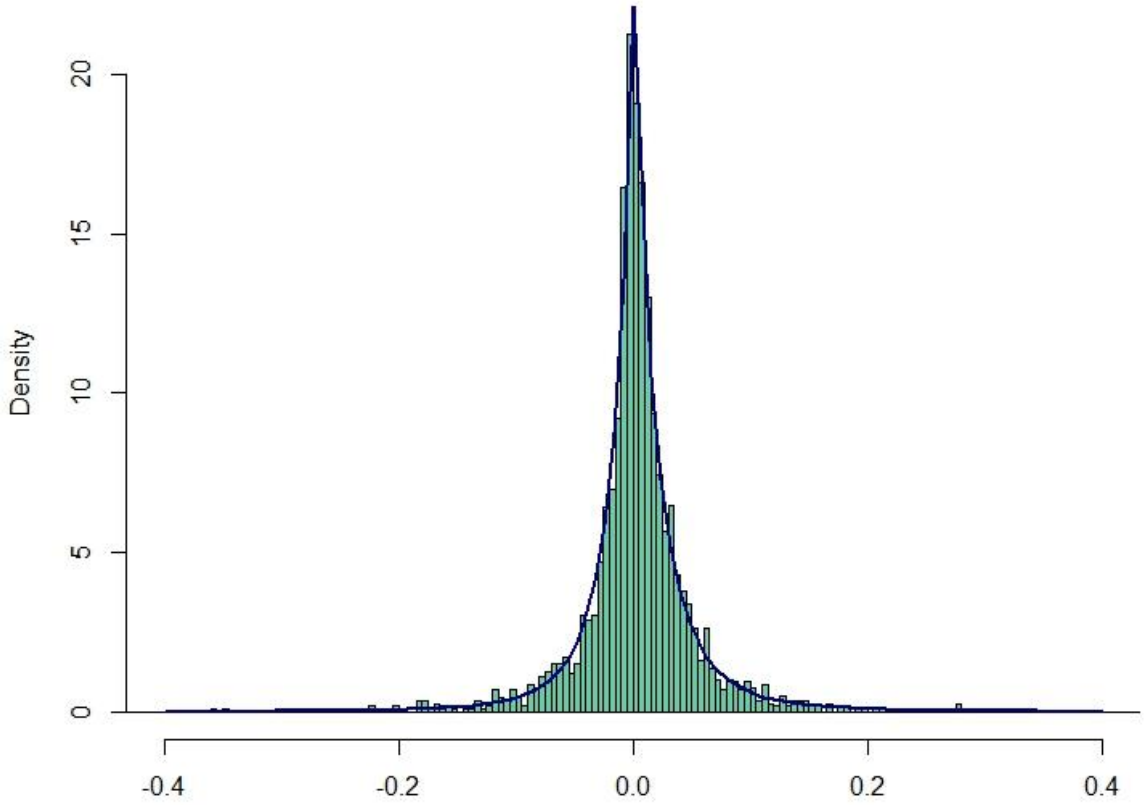


Figure 12 - Skewed generalised t fitted to the log-returns of the exchange rate from 13/09/2011 to 01/04/2018

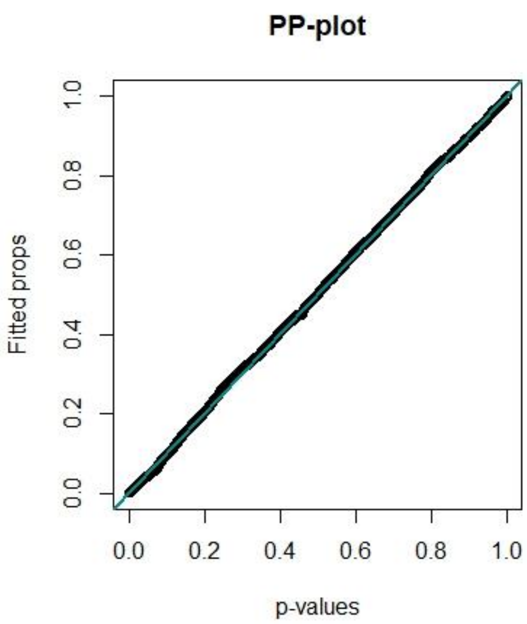


Figure 14 - PP-plot

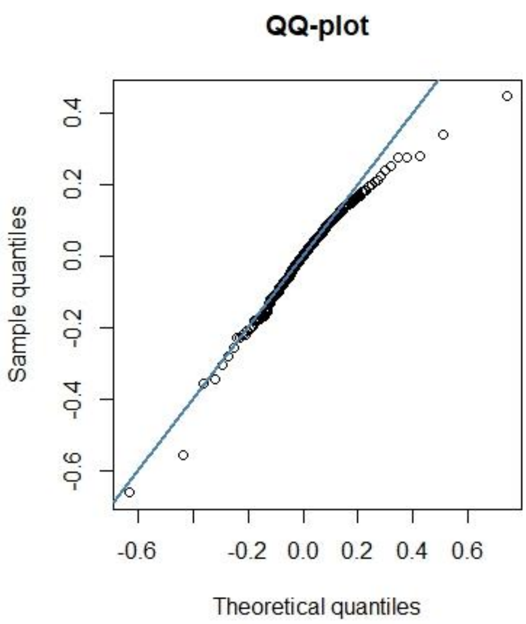


Figure 13 - QQ-plot

From these figures, it is clear that the model is reasonably fitted. In the histogram, the fitted distribution is covering the shape of the figure quite well, with a few outliers seen in the tails for the most part. This is also seen in the two plots, especially in the QQ-plot, which is better at pointing out how the distribution fits to the tails. There are some disturbances in the edges of the plot, especially on the topside, so this must be taken into consideration if we are to evaluate extreme observations through this distribution. The line in the quantile plot is drawn through quantiles 0.01 and 0.99. We see that the most observations are found on the line. Somewhat surprising do we see that the minimum of the dataset is found on the quantile line, so the observation may not be as farfetched as one probably would suspect. The simple PP-plot is showing that, for simple probabilities, the skewed generalised *t* distribution is performing well, since the $(0,1)$-line is not deviating from the observed probabilities.

The key estimations of financial risk are given in the next chapter.

7. Value at risk and expected shortfall

The measures of financial risk, VaR and ES, will be estimated and evaluated through the best fitted distribution and checked by the probability ρ through

$$\frac{\hat{p}}{2\hat{c}\hat{\sigma}\hat{\nu}^{\frac{1}{\hat{p}}}B\left(\frac{1}{\hat{p}},\hat{\nu}\right)}\int_{-\infty}^{\widehat{VaR}_{\rho}}\hat{p}\left[\left(\frac{|x-\hat{\mu}+\hat{m}|^{\hat{p}}}{\hat{\nu}(\hat{c}\hat{\sigma})^{\hat{p}}\left(\hat{\lambda}sign(x-\hat{\mu}+\hat{m})+1\right)^{\hat{p}}}+1\right)^{\frac{1}{\hat{p}}+\hat{\nu}}\right]^{-1}dx=\rho,$$

and

$$\widehat{ES}_p=\frac{1}{p}\int_0^p VaR_q\,dq,$$

respectively, where $\hat{c}=\hat{\nu}^{-\frac{1}{\hat{p}}}\left[\sqrt{\left(3\hat{\lambda}^2+1\right)\frac{B\left(\frac{3}{\hat{p}},\hat{\nu}-\frac{2}{\hat{p}}\right)}{B\left(\frac{1}{\hat{p}},\hat{\nu}\right)}-4\hat{\lambda}^2\frac{B\left(\frac{2}{\hat{p}},\hat{\nu}-\frac{1}{\hat{p}}\right)^2}{B\left(\frac{1}{\hat{p}},\hat{\nu}\right)^2}}\right]^{-1}$ and $\hat{m}=\frac{2\hat{c}\hat{\sigma}\hat{\lambda}\hat{\nu}^{\frac{1}{\hat{p}}}B\left(\frac{2}{\hat{p}},\hat{\nu}-\frac{1}{\hat{p}}\right)}{B\left(\frac{1}{\hat{p}},\hat{\nu}\right)}$. The two measures are plotted against p in figures 14 and 15. The estimated values themselves at different selected p-levels are presented in table 5. We choose to focus on the outlying cases, as these give better indication on the degree of volatility.

The VaR and ES show that the data are behaving as it is to be expected. The VaR has a sharp and steep decline and incline at the extreme probability levels at each side that follows from data with high volatility. This is clear both from the plot, and also by looking at the table, that the financial risk involved is drastically increasing in the extreme cases. For instance, from $p=0.1$ to $p=0.01$ there is a difference of 0.10975, while from $p=0.0001$ to $p=0.00001$

P	*VaR*	*ES*
0.1	$-3.915 * 10^{-2}$	$-1.001 * 10^{-1}$
0.01	$-1.489 * 10^{-1}$	$-1.921 * 10^{-1}$
0.001	$-3.700 * 10^{-1}$	$-2.852 * 10^{-1}$
0.0001	$-8.138 * 10^{-1}$	$-3.790 * 10^{-1}$
0.00001	-1.702	$-4.733 * 10^{-1}$
0.9	$5.011 * 10^{-2}$	$-8.024 * 10^{-3}$
0.99	$1.772 * 10^{-1}$	$2.276 * 10^{-3}$
0.999	$4.338 * 10^{-1}$	$4.132 * 10^{-3}$
0.9999	$9.483 * 10^{-1}$	$4.409 * 10^{-3}$
0.99999	1.979	$4.447 * 10^{-3}$

Table 5 - VaR and ES estimates for the skewed generalised t distribution for the given exchange rate and sample period

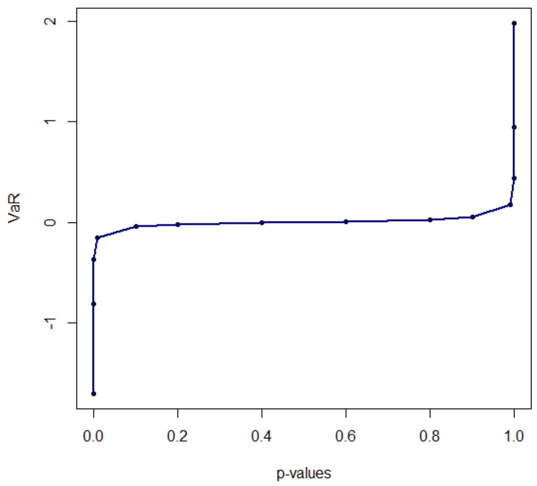


Figure 15 - Fitted VaR to p-values

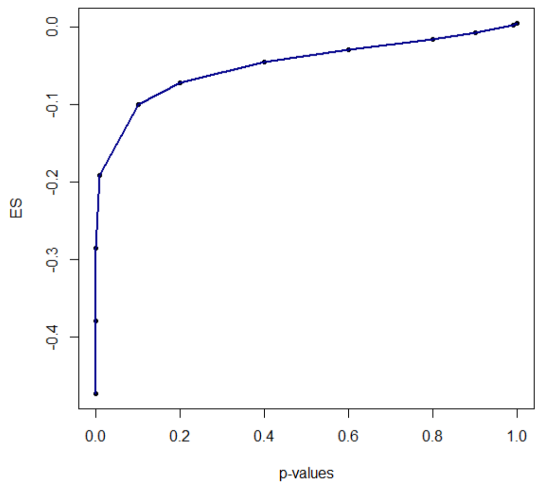


Figure 16 - Fitted ES to p-values

the difference is 0.8882.The plotted ES shows also that there is a sharp decline in the left side of the asymtotic function, and an increasing, but flattening, value to the right that is expected to fall short if bad comes to worse.

It is important to remember that all these estimations are done with the log-returns and does not represent any concrete form of currency themselves.

Now that we have presented our results, we will sum up our findings in the conclusion and give some remarks on these. There will also be some proposals for how to use this thesis as a base for future research, or if these findings will be described better by other methods when the data series has grown even larger.

8. Conclusions

The exchange rate of Bitcoin to the USD has now been analysed and modelled by eleven frequently used distributions in finance. We found that the skew generalised *t* distribution gives the best fitted model of the data, decided by its favourable values in log-likelihood, AIC, AICc, and especially in the p-values of the Kolmogorov-Smirnov and Anderson-Darling statistics. It is challenged by the generalised *t* distribution on the statistics BIC, CAIC and HQC, but since these criteria are not decisive when evaluating nested distributions very well, these are overruled by the log-likelihood in this case. It is also clear, from the p-values of the statistics, that after the skewed generalised *t*, the generalised hyperbolic distribution is the only distribution providing a satisfying fit of the model above the significance level of 5% in both statistics, while the generalised *t* and skewed exponential power distributions are significant in only one statistic each.

Predictions of the financial risk by Value at Risk and Expected Shortfall have also been calculated and shown in table 5. Here we see for instance that there is a probability of 1 percent that the log-returns of the exchange rate will be either smaller than -0.1489, or, by the same probability, be larger than 0.1772. We see that it is highly unlikely, but possible, to observe a log-return smaller than -1.702 or larger than 1.979. This would be formidable, especially since the most extreme values observed by today have “only” been -0.644 at the lowest and 0.446 at the highest, and these were observed at an early, and very volatile, period of Bitcoin.

We will present our findings in a relevant and empirical context, where we can compare our results to some other perspectives of the topic.

There have been many occasions where prominent financial or economical persons have deemed Bitcoin a fraud or a failed currency experiment, as CEO of JPMorgan, Jamie Dimon (Martin, 2018), or Nobel-winning economist Robert Shiller (Rooney, 2018). On the other side we see a bank in Argentina using bitcoins as cross-border payments (De, 2018) or a crypto mining rig supplying company from Beijing, Canaan Inc, attempting to raise $2 bn., supposedly for innovations and development in AI and blockchain (John & Hughes, 2018).

There are plenty of examples of sides taken in the act of speculations around Bitcoin. Obviously, this is part of whether we would think of a person as risk-averse or risk-seeking. Because of Bitcoins extreme volatility, there are many persons who are willing to risk money in order to have the chance of gaining more, while others keep away from this probable uncertainty to their wealth. It is still such that most people see Bitcoin as an asset, an investment opportunity rather than a currency. Although unclear for now, this might change in the near future.

According to analysts at Barclays are Bitcoin's development similar to that of a disease – it had an outbreak and became spread through the spoken or written word about it, but as it peaked in late-2017 it reached the immunity threshold and the potential host population has decreased in size as most people now know about it, and the majority of "infections" has reached its high (Lam, 2018).

Bitcoin have many different aspects that may seem contradictory, but essentially it is uncertainty that reigns, and peoples attitude towards the new and unknown. At the same time, it seems that the value of Bitcoin is kept somewhat stable, over the last weeks at least, and may have achieved some sort of equilibrium, at least temporary. Still, the dynamic rate requires a more complex explanation than most other currencies it can be compared to if it is to be considered a currency rather than an asset.

A proposal for future work to use non-parametric distributions or semiparametric distributions still stands, but it would also be interesting to check whether other parametric distributions would have a better fit if later dynamics of the exchange rate start behaving unpredictably. GARCH models have already been assessed to the rate, but only for a short time period. This could be done more in-depth, with focus on Bitcoin rather than several cryptocurrencies at once.

Besides exercising this fitting procedure repeatedly, one may use the information provided here to perform newer predictions for the exchange rate and further study the dynamics and development, as it is sure not to be disregarded or overlooked for the foreseeable future.

Bitcoin as a phenomenon may not necessarily be in the early phase any longer, but the concept and mechanics of the blockchain has still much potential to obtain. A concluding remark by Nassim Taleb stands as strong today as it did five years ago:

> *«Bitcoin is the beginning of something great: a currency without a government, something necessary and imperative. But I am not familiar with the specific product to assert whether it is the best potential setup. And we need a long time to establish confidence…»*

(Taleb, 2013)

9. References

Akaike, H. (1974, December). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), pp. 716-723. doi:10.1109/TAC.1974.1100705

Anderson, T. W., & Darling, D. A. (1954, December). A Test of Goodness-of-Fit. *Annals of Mathematical Statistics, 49*(268), pp. 765-769. doi:10.1080/01621459.1954.10501232

Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1998, July 22). *Coherent Measures of Risk.* Retrieved May 11, 2018, from ETHzürich: https://people.math.ethz.ch/~delbaen/ftp/preprints/CoherentMF.pdf

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics, 12*, 171-178.

Barndorff-Nielsen, O. (1977, March 25). Exponentially decreasing distributions for the logarithm of particle size. *The Royal Society, 353*(1674), 401-409. doi:10.1098/rspa.1977.0041

Bartels, R. (1982). The rank version of von Neumann's ratio test for randomness. *Journal of the American Statistical Association, 77:377*, 40-46. doi:10.1080/01621459.1982.10477764

Bienaymé, J. (1874). Sur une question de probabilités. *Bulletin de la Société Mathématique de France, 2*, 153-154. doi:10.24033/bsmf.56

bitcoin.org. (2018). *Some Bitcoin words you might hear*. Retrieved from bitcoin.org: https://bitcoin.org/en/vocabulary

bitcoinwiki. (2017). *Confirmation*. Retrieved from bitcoinwiki: https://en.bitcoin.it/wiki/Confirmation

blockchain. (2018). Blockchain Charts. blockchain.info.

Bozdogan, H. (1987). Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. *Psychometrika, 52*(3), pp. 345-370. doi:10.1007/BF02294361

Brockwell, P. J., & Davis, R. A. (2002). *Introduction to Time Series and Forecasting.* New York: Springer Texts in Statistics, Springer-Verlag.

Caeiro, F., & Mateus, A. (2014, November 16). Testing randomness in R. *randtests*. R package version 1.0.

Cannucciari, C. (Director). (2016). *Banking on Bitcoin* [Motion Picture]. Retrieved from bitcoin.org: https://bitcoin.org/bitcoin.pdf

Chan, S. (2016). *Some contributions to statistical modeling in finance*. Retrieved from The University of Manchester - School of Mathematics: http://www.maths.manchester.ac.uk/~saralees/PhDthesis_stephen_chan.pdf

Chu, J., Chan, S., Nadarajah, S., & Osterrieder, J. (2017). GARCH Modelling of Cryptocurrencies. *Journal of Risk and Financial Management, 10*(4). doi:10.3390/jrfm10040017

Chu, J., Nadarajah, S., & Chan, S. (2015). Statistical Analysis of the Exchange Rate of Bitcoin. *PLoS One, 10*(7). Retrieved from https://doi.org/10.1371/journal.pone.0133678

Cieśla, K. (2018). *Bitcoin trading volume*. Retrieved from data.bitcoinity.org: https://data.bitcoinity.org/markets/volume

CoinMarketCap. (2018). *All Cryptocurrencies*. Retrieved from coinmarketcap.com: https://coinmarketcap.com/all/views/all/

Corlu, C. G., & Corlu, A. (2015). Modelling exchange rates returns: which flexible distribution to use? *Quantitative Finance, 15*(11), 1851-1864. doi:10.1080/14697688.2014.942231

Cox, D. R., & Hinkley, D. V. (1974). *Theoretical Statistics.* London: Chapman and Hall.

Cox, D. R., & Stuart, A. (1955). Some Quick Sign Tests for Trend in Location and Dispersion. *Biometrika, 42*(1/2), 80-95. doi:10.2307/2333424

De, N. (2018, May 22). *A Bank In Argentina Is Now Using Bitcoin for Cross-Border Payments*. Retrieved from coindesk: https://www.coindesk.com/bank-argentina-just-added-bitcoin-cross-border-payments/

Embrechts, P., Kaufmann, R., & Patie, P. (2005, October). Strategic Long-Term Financial Risks: Single Risk Factors. *Computational Optimization and Applications, 32*(1-2), pp. 61-90. doi:10.1007/s10589-005-2054-7

European Central Bank. (2015). *Virtual currency schemes - a further analysis.* Frankfurt am Main: European Central Bank. doi:10.2866/662172

Gauss, C. F. (1809). *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium.* Hamburg: Friedrich Perthes and I.H. Besser.

Gosset, W. S. (1908, March). The probable error of a mean. *Biometrika*(6), 1-25. doi:10.1093/biomet/6.1.1.

Hannan, E. J., & Quinn, B. G. (1979). The Determination of the order of an autoregression. *Journal of the Royal Statistical Society, 41*(2), pp. 190-195. Retrieved from http://www.jstor.org/stable/2985032

Hansen, B. E. (1994, August). Autoregressive Conditional Density Estimation. *International Economic Review, 35*(3), 705-730. doi:10.2307/2527081

Hansen, C., McDonald, J. B., & Newey, W. K. (2010). Instrumental Variables Estimation With Flexible Distributions. *Journal of Business & Economic Statistics, 28*(1), pp. 13-25. doi:10.1198/jbes.2009.06161

Hart, A., & Martínez, S. (2017). Statistical Patterns in Genomic Sequences. *spgs*. R package version 1.0-1.

Hoffmann, T., & Watchulonis, M. (Directors). (2015). *Bitcoin: The End Of Money As We Know It* [Motion Picture].

Hull, J. C. (2015). *Risk Management and Financial Institutions* (4th ed.). Wiley.

Hurvich, C. M., & Tsai, C.-L. (1989, June 1). Regression and time series model selection in small samples. *Biometrika, 76*(2), pp. 297-307. doi:10.1093/biomet/76.2.297

Investopedia. (2016). Soft Fork and Hard Fork. Investopedia.

John, A., & Hughes, J. (2018, May 22). *As bitcoin world wobbles, mining rig company plans $2 billion IPO*. Retrieved from Reuters: https://www.reuters.com/article/us-canaan-ipo/as-bitcoin-world-wobbles-mining-rig-company-plans-2-billion-ipo-idUSKCN1IN01X

Johnson, N. L., & Kotz, S. (1970). *Distributions in statistics: continuous univariate distributions* (Vol. 2). Boston: Houghton Mifflin Company.

Kendall, M. G. (1955). *Rank Correlation Methods* (2nd ed.). Oxford, England: Hafner Publishing Co.

Khatwani, S. (2018). *Explaining Hash Rate Or Hash Power In Cryptocurrencies*. Retrieved from Coinsutra: https://coinsutra.com/hash-rate-or-hash-power/

Kodrič, N. (2015). *BITSTAMP IS OPEN FOR BUSINESS - BETTER THAN EVER!* Retrieved from Bitstamp: https://www.bitstamp.net/article/bitstamp-is-open-for-business-better-than-ever/

Kolmogorov, A. (1933). Sulla Determinazione Empirica di um Legge di Distribuizione. *Giornale Dell'Instituto Italiano Degli Attuari, 4*, pp. 83-91. doi:10.1007/978-1-4612-4380-9_9

Lam, E. (2018, April 10). *Why Bitcoin Behaves Like the Flu*. Retrieved from Bloomberg: https://www.bloomberg.com/news/articles/2018-04-10/bitcoin-is-a-disease-in-barclays-model-that-says-prices-peaked

Laplace, P.-S. (1774). Mémoire sur la probabilité des causes par les évènements. *Mémoires de l'Academie Royale des Sciences Presentés par Divers Savan*(6), 621-656.

Ljung, G. M., & Box, G. E. (1978). On a measure of lack of fit in time series models. *Biometrika, 65*(2), 297-303. doi:10.1093/biomet/65.2.297

Mann, H. B. (1945). Nonparametric Tests Against Trend. *Econometrica, 13*(3), 245-259. doi:10.2307/1907187

Martin, K. (2018, January 9). *JPMorgan Chase CEO Jamie Dimon regrets saying Bitcoin is a 'fraud,' but still isn't interested in it*. Retrieved from FOXBusiness: https://www.foxbusiness.com/markets/jpmorgan-chase-ceo-jamie-dimon-regrets-saying-bitcoin-is-a-fraud-but-still-isnt-interested-in-it

McDonald, J. B., & Newey, W. K. (1988, December). Partially Adaptive Estimation of Regression Models via the Generalized t Distribution. *Econometric Theory, 4*(3), 428-457.

Nakamoto, S. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System.* Retrieved from bitcoin: https://bitcoin.org/bitcoin.pdf

Osterrieder, J. (2017). The Statistics of Bitcoin and Cryptocurrencies. *Advances in Economics, Business and Management Research, 26*. Retrieved from https://pd.zhaw.ch/publikation/upload/211430.pdf

Osterrieder, J., Chan, S., Chu, J., & Nadarajah, S. (2017, April 7). *The Statistical Analysis of Cryptocurrencies.* doi:10.2139/ssrn.2948315

Pohlert, T. (2018). Non-Parametric Trend Tests and Change-Point Detection. *trend*. R package version 1.1.0.

Quantitivity. (2011, February 21). *Why Log Returns*. Retrieved from Quantivity: https://quantivity.wordpress.com/2011/02/21/why-log-returns/

Redman, J. (2017). *Bitcoin Cash Network Completes a Successful Hard Fork*. Retrieved from Bitcoin.com: https://news.bitcoin.com/bitcoin-cash-network-completes-a-successful-hard-fork/

Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *The Journal of Risk, 2*(3), pp. 21-41. doi:10.21314/JOR.2000.038

Rooney, K. (2018, May 21). *Bitcoin could be just another failed currency experiment, Nobel-winning economist Robert Shiller warns*. Retrieved from CNBC: https://www.cnbc.com/2018/05/21/bitcoin-could-be-another-failed-currency-robert-shiller-says.html

Rosic, A. (2017). *What is Blockchain Technology? A Step-by-Step Guide For Beginners*. Retrieved from Blockgeeks: https://blockgeeks.com/guides/what-is-blockchain-technology/

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), pp. 461-464. doi:10.1214/aos/1176344136

Smirnov, N. V. (1948). Table for estimating the goodness of fit of empirical distributions. *Annals of Mathematical Statistics, 19*(2), pp. 279-281. doi:10.1214/aoms/1177730256

Subbotin, M. T. (1923). On the Law of Frequency of Error. *Matematicheskii Sbornik,* 296-301.

Taleb, N. (2013). *IAm Nassim Taleb, author of Antifragile, AMA*. Retrieved from reddit: https://www.reddit.com/r/IAmA/comments/1aoi0s/iam_nassim_taleb_author_of_antifragile_ama/c8zb3d8/

Theodossiou, P. (1998, December). Financial Data and the Skewed Generalized T Distribution. *Management Science, 44*(12), pp. 1650-1661. Retrieved from http://www.jstor.org/stable/2634700 user3783846. (2018). *How to interpret serial autocorrelations and AR models*. Retrieved from StackExchange: https://stats.stackexchange.com/questions/325226/how-to-interpret-serial-autocorrelations-and-ar-models

Wald, A., & Wolfowitz, J. (1940). On a Test Whether Two Samples are from the Same Population. *Annals of Mathematical Statististics, 11*(2), 147-162. doi:10.1214/aoms/1177731909

Wald, A., & Wolfowitz, J. (1943). An Exact Test for Randomness in the Non-Parametric Case Based on Serial Correlation. *Annals of Mathematical Statististics, 14*(4), 378-388. doi:10.1214/aoms/1177731909

Wendel, J. G. (1961). The non-absolute convergence of Gil-Pelaez' inversion integral. *Annals of Mathematical Statistics, 32*(1), pp. 338-339. doi:10.1214/aoms/1177705164

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

Wilks, S. S. (1938). The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *The Annals of Mathematical Statistics, 9*(1), pp. 60-62. doi:10.1214/aoms/1177732360

Wong, J. I. (2017). *A coding error led to $30 million in ethereum being stolen*. Retrieved from Quartz: https://qz.com/1034321/ethereum-hack-a-coding-error-led-to-30-million-in-ethereum-being-stolen/

Zhu, D., & Galbraith, J. W. (2010, August). A generalized asymmetric Student-t distribution with application to financial econometrics. *Journal of Econometrics, 157*(2), 297-305.

Zhu, D., & Zinde-Walsh, V. (2009, January). Properties and estimation of asymmetric exponential power distribution. *Journal of Econometrics, 148*(1), 86-99.

10. Appendices

10.1 Appendix A – BTC transactions by currency

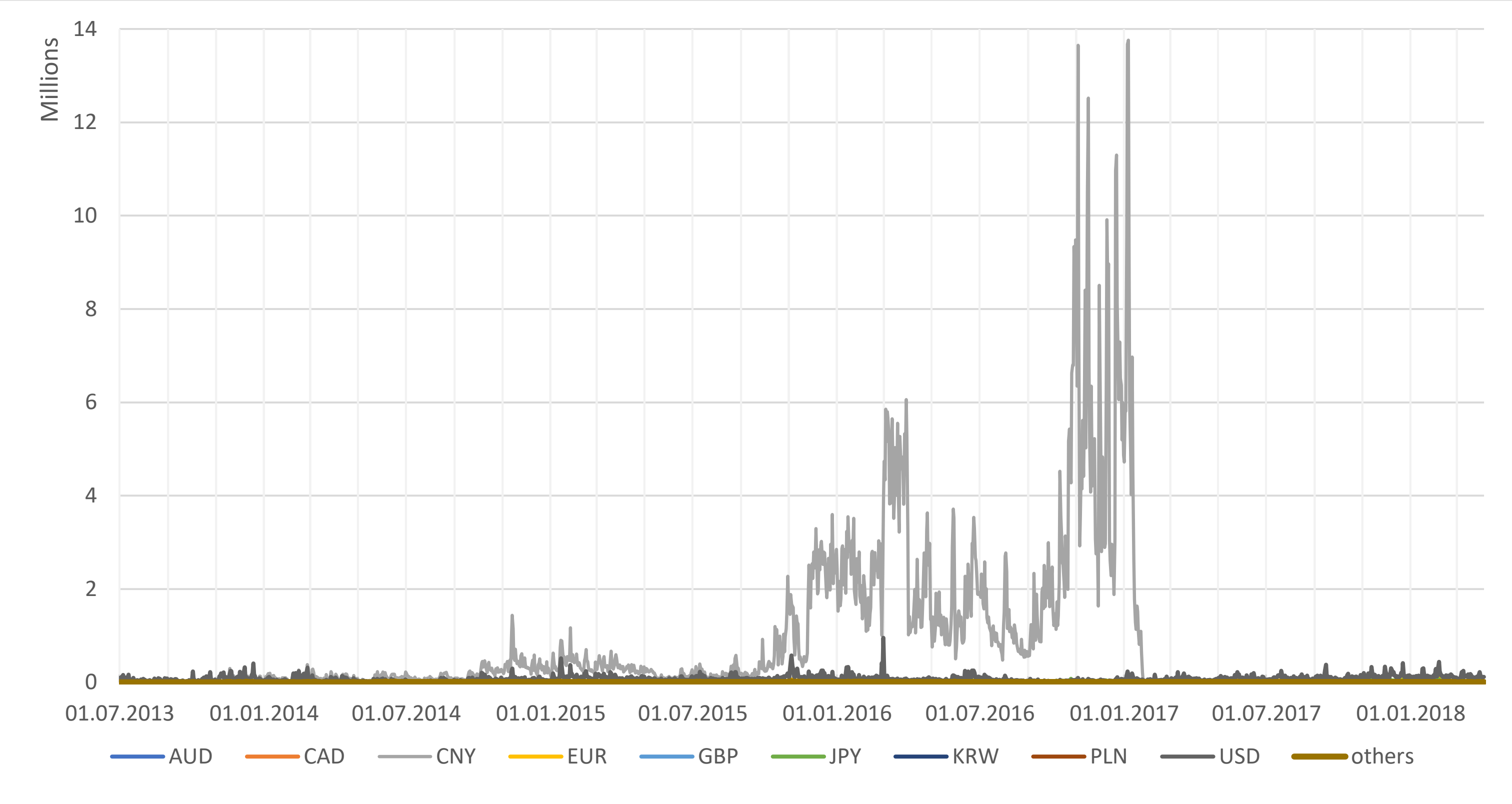


Figure A 2 - Bitcoin trading volume, all currencies (Cieśla, 2018)

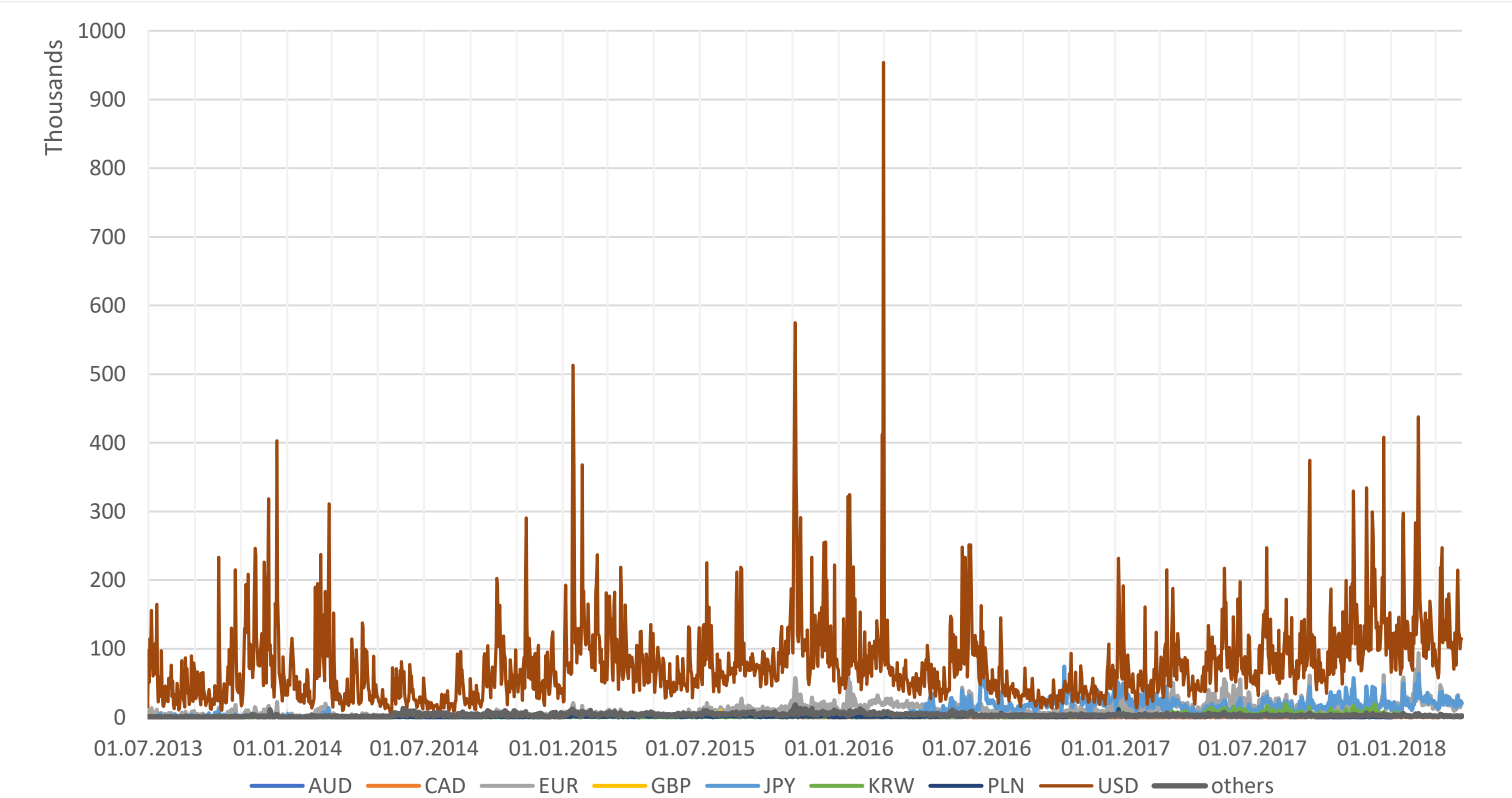


Figure A 1 - Bitcoin trading volume, currencies except Chinese Yuan (Cieśla, 2018)

10.2 Appendix B – R script

```
####Importing and preparing the dataset
library("haven")
library("stats4")
dataset_full <- read_dta("~/data_full.dta")
View(data_full)
attach(data_full)
data_full<-data_full[-1, ] ##Remove first line with no log-ret, so no
missing values.

####Tests for randomness
library("randtests")      ##provides      bartel.rank-,      cox.stuart-,
difference.sign-, rank-, and turning.point.test
bartels.rank.test(lncl)
cox.stuart.test(lncl)
difference.sign.test(lncl)
rank.test(lncl)
turning.point.test(lncl)
detach("package:randtests", unload=TRUE)
library("spgs") ##provides lb test
lb.test(lncl)
detach("package:spgs", unload=TRUE)
library("trend") ##provides ww test
ww.test(lncl)
detach("package:trend", unload=TRUE)

####Fits on distribution
##Setting up functions
library("sgt")
NLL.sgt <- function(m, s, l, p1, q1) {
     -sum(dsgt(lncl, mu=m, sigma=s, lambda=l, p=p1, q=q1, log=TRUE))
    }
NLL.logis <- function(mu, s) {
     -sum(dlogis(lncl, location=mu, scale=s, log=TRUE))
    }
detach("package:sgt", unload = TRUE)

##Making MLE variables
library("sgt")
library("stats4")
fit.norm <- mle(minuslogl = NLL.sgt, start = list(m=0.004534, s=0.06868),
method = "Nelder-Mead", fixed = list(l=0, q1=Inf, p1=2))
fit.studt <- mle(minuslogl = NLL.sgt, start = list(m=0.003858, s=0.02134,
q1=1.389), method = "Nelder-Mead", fixed = list(l=0, p1=2))
```

```
 fit.logis  <-  mle(minuslogl  =  NLL.logis,  start  =  list(mu=0.005391,
s=0.02892), method = "Nelder-Mead")
 fit.lapl <- mle(minuslogl = NLL.sgt, start = list(m=0.003753, s=0.03804),
method = "Nelder-Mead", fixed = list(l=0, q1=Inf, p1=1))
 fit.exppow <- mle(minuslogl = NLL.sgt, start = list(m=0.003996, s=0.02819,
p1=0.5871), method = "Nelder-Mead", fixed = list(l=0, q1=Inf))
 fit.skewnorm  <-  mle(minuslogl  =  NLL.sgt,  start  =  list(m=0.004534,
s=0.06868, l=0.000000006006), method = "Nelder-Mead", fixed = list(q1=Inf,
p1=2))
 fit.skewt <- mle(minuslogl = NLL.sgt, start = list(m=0.0009774, s=0.02133,
l=0.1639, q1=1.379), method = "Nelder-Mead", fixed = list(p1=2))
 fit.gent <- mle(minuslogl = NLL.sgt, start = list(m=0.003026, s=0.0231,
p1=0.9471, q1=3.042), method = "Nelder-Mead", fixed = list(l=0))
 fit.skexppow <- mle(minuslogl = NLL.sgt, start = list(m=0.004, s=0.02812,
l=0.4936, p1=0.5842), method = "Nelder-Mead", fixed = list(q1=Inf))
 fit.skgent <- mle(minuslogl = NLL.sgt, start = list(m=0, s=1, l=0, p1=2,
q1=3), method = "Nelder-Mead")detach("package:sgt", unload = TRUE)
 detach("package:stats4", unload = TRUE)

 ##Estimating the paramters and the distribution's log-likelihood
 summary(fit.norm)
 summary(fit.logis)
 summary(fit.lapl)
 summary(fit.studt)
 …

 ##Criterions

 AIC(fit.norm)
 AIC(fit.logis)
 etc...
 #NB: Rest criterions were calculated by hand

 ##KS and AD-test
 ks.test(unique(lncl),  psgt,  mu=0.002968,  sigma=0.05304,  lambda=0,  p=2,
q=Inf) #Normal
 ks.test(unique(lncl), plogis, location=0.003517, scale=0.02298) #Logistic
 ks.test(unique(lncl),  psgt,  mu=0.002242,  sigma=0.04307,  lambda=0,  p=1,
q=Inf) #Laplace
 ks.test(unique(lncl),  psgt,  mu=0.003054,  sigma=0.4321,  lambda=0,  p=2,
q=1.002) #Student  t
 …

 ad.test(unique(lncl),  psgt,  mu=0.002968,  sigma=0.05304,  lambda=0,  p=2,
q=Inf) #Normal
 ad.test(unique(lncl), plogis, location=0.003517, scale=0.02298) #Logistic
```

```
 ad.test(unique(lncl), psgt, mu=0.002242, sigma=0.04307, lambda=0, p=1, q=Inf) #Laplace
 ad.test(unique(lncl), psgt, mu=0.003054, sigma=0.4321, lambda=0, p=2, q=1.002) #Student t
 …

 ##VaR and ES

 library("cvar")
 VaR(qsgt, x=c(0.1, 0.01, 0.001, 0.0001, 0.00001, 0.9, 0.99, 0.999, 0.9999, 0.99999), mu=0.004452, sigma=0.05877, lambda = 0.05672, p=0.9747, q=3.409)
 ES(qsgt, x=c(0.1, 0.01, 0.001, 0.0001, 0.00001, 0.9, 0.99, 0.999, 0.9999, 0.99999), mu=0.004452, sigma=0.05877, lambda = 0.05672, p=0.9747, q=3.409)
 detach("package:cvar", unload = TRUE)

##Plotting a skewed generalised t distribution

par(mfrow=c(1,2))

hist(lncl, col = "mediumaquamarine",breaks = 160, freq=FALSE, xlim = c(-0.4,0.4), main = "")

curve(dsgt(x, mu=0.004452, sigma=0.05877, lambda=0.05672, p=0.9747, q=3.409), add=TRUE, lwd=2, col="navy")

qqnorm(dsgt(lncl, mu=0.004452, sigma=0.05877, lambda=0.05672, p=0.9747, q=3.409), main="", col="navy")

qqline(dsgt(lncl, mu=0.004452, sigma=0.05877, lambda=0.05672, p=0.9747, q=3.409), col="royalblue")


##Plotting VaR and ES

plot(probs, VaRvals, pch=20, ylab="VaR", xlab= "p-values")

lines(smooth.spline(probs, VaRvals, df=14), col="darkblue", lwd=2)

plot(probs, ESvals, pch=20, ylab="ES", xlab= "p-values")

lines(smooth.spline(probs, ESvals, df=14), col="darkblue", lwd=2)
```

10.3 Appendix C – Stata DO-file

```
 import excel "C:\Users\Joaki\OneDrive\Master\BTC2.xlsx", sheet("Sheet1")
firstrow

 //Labeling and renaming
 ren BCHARTSBITSTAMPUSDCLOSE Close
 ren BCHARTSBITSTAMPUSDHIGH High
 ren BCHARTSBITSTAMPUSDLOW Low
 ren BCHARTSBITSTAMPUSDOPEN Open
 ren BCHARTSBITSTAMPUSDVOLUMEBTC Volume_BTC
 ren BCHARTSBITSTAMPUSDVOLUMECURR Volume_USD
 ren BCHARTSBITSTAMPUSDWEIGHTEDPRI Wgh_price
 ren DATE Date
 la var Close "Close price"
 la var High "High price"
 la var Low "Low price"
 la var Open "Open price"
 la var Volume_BTC "Volume in BTC"
 la var Volume_USD "Volume in USD"
 la var Wgh_price "Weighted price"
 la var Date "Date from 13. Sept 2011"

 tsset Date //Applying time format to "Date"-variable

 format Date %dd/n/CY //Formating dates to non-US standard

 drop if Close==0 //Remove values that are lacking for non-trading days

 //Generating time variable for configuring log-returns
 gen lncl = ln(Close [_n]) - ln(Close [_n-1])

 //Labeling and renaming for convenience
 la var lncl "Log-return CL"
 twoway line lncl Date

 sum lncl, d //Detailed summarize of lnop's statistical properties

 //Calculating the autocorrelation and partial autocorrelation of lncl

 ac lncl, lags(20) //Insufficient?
 pac lncl, lags(20)

 //Use and "manipulate" syntax _asrol_ for calculating cumulative
 //(variance, skewness,) and kurtosis of lncl

 asrol lncl, stat(sd) w(Date 2372) //var sd2372_lncl is generated
 gen var_lncl = (sd2372_lncl)^2
 twoway line var_lncl Date

 //Import calculated skewness and kurtosis from excel
 twoway line skw_lncl Date
 twoway line krt_lncl Date
```

10.4 Appendix D – Reflection notes: Joakim Nilgard

This reflective note is done over the thesis Filip and I have written together this semester, which has been a challenging and demanding, but also rich, learning and very interesting to work on. The thesis is about Bitcoin, both the concept and the money itself, and our path of finding a distribution that may fit its logarithmic returns on the exchange rate to USD. At first it may seem a little farfetched, trying to fit something non-stationary into something that requires stationarity, but as we have come to understand – it is just more to it than to throw the data into an ARCH or GARCH model, which would probably be intuitive for some. What we wanted to do, was to find an underlying distribution of said log-returns, apply some information to our findings, and propose a fundament or starting point for further research on the topic. We found that the best distribution to describe the data was the skew generalised *t* distribution, which is a complex, five-parametric model. We applied this estimated model to two financial measurements, which gave indications on how volatile the exchange rate has been and still is.

Before embarking on the following sections where I will link the thesis topic to three themes, I must specify what exactly is the topic of the thesis to discuss. Although general, I will do these evaluations with regard to Bitcoin as a concept because I do not find it reasonable to discuss either internationality, innovation or responsibility in terms of exchange rate, statistical tests or distributions. Technology or currencies will be too broad, and the blockchain or cryptography is not the studied topic (even if it would be interesting to study this further on).

Bitcoin relates to international trends and is affected by (and probably affects as well) many factors, such as technology, politics, the markets demand, ideology, media, sociological, and legal forces. From what I have seen, there is a lot going on concerning Bitcoin already from the day Satoshi Nakamoto proposed the concept of virtual, cryptographic money. As it is partially discussed in the thesis, it is important to separate between Bitcoin itself and its code, and the blockchain, even though Bitcoin is built upon the blockchain and cannot exist without it. With this in mind is it easier to remove technology as a related factor, because it is only the blockchain that is a technological breakthrough, while Bitcoin is only another currency/mean of trading/asset etc.

There have been many political debates about Bitcoin, and many policies have taken form with it in regard. Policies restricting availability, trading opportunities, exchange mediums have been adopted. Taxation has been introduced if you are at all allowed by your bank to trade in cryptocurrencies. There have been many cases of fraud, embezzlement, Ponzi schemes and cheating with Bitcoin, although it seems for me that it has been seen at like more severe cases of the crime than if one were to use fiat-money. There has also been a case where a politician in the US suggested a policy about establishment of crypto-exchanges that were adopted, for then only to quit as a politician and establish a consulting firm that specialised in this subject and earned a lot of money on this. Most of the men that were in the core of the development if Bitcoins code have been arrested, being accused of many different things (whereas not all accusations are completely credible). Why is this? Is it because the Bitcoin-community tried and are trying to change something fundamental in the governmental system like relieving the need for banks or third-parties in transactions and thus removing enormous funding for the state? Is it because of the insecurity about the new phenomenon that has arisen, and not being sure about how to handle it? Is it because of an ideological cause that don't support the "omniscience" of the government and the wish for a freer world with less surveillance and power to politicians? Is it because of the hardship of traceability of transactions that may induce criminality and an attempt to limit this? There are many interesting aspects regarding the scope this comprises internationally, but as this can be discussed indefinitely, I'll leave it at this.

When one would think of innovation as something Bitcoin can offer, I fail to see exactly how this would be a topic to discuss. This is because, as it is discussed more in detail in the thesis, Bitcoin is not a new innovation itself – the blockchain is! Bitcoin is only based on this new, *innovative* technology, but is not new itself, because it can be perceived as something between a currency and an asset. That is not innovative and does not possess the ability to grow into something innovative either. Bitcoin may indirectly offer some innovative solutions, for instance in terms of money transfers or online and store payments, but this is still due the technology of the blockchain and could be done with most other digital money if the need was present. I see innovation as a description that is closely related to inventions, creative or practical solutions to problems or conveniences, or to make either something completely new or remaking already existing innovations into something different. It is something that is needed for a society to be progressive and developing, to bring out a competitive edge, to stimulate demand, to modernise and bring technology forth. What fulfils these requirements is the blockchain, which is applicable to almost any instance by configuration, not another type of asset or currency, which we already have countless examples of from before. Therefore, I do not see innovation as a relevant topic for Bitcoin.

The same applies for responsibility, I cannot see that responsibility is a topic to discuss regarding something that is not personal, organisational or in any kind a legal entity that has any responsibilities itself. Bitcoin is a software that may be used by practically anyone that has a computer connected to the internet. To have something to discuss, I will take a step back and view the usage of Bitcoin as a topic. As the term responsibility involves aspects like duty, accountability, and control, it is not something that Bitcoin itself can be described as, but rather those who use it.

The use of Bitcoin is to some extent discussed in the thesis, in chapter 2, where different kinds of mis-use of the technology has been attempted. Mis-use involves for instance hacking, embezzlement, scamming, and exploitation. It is most often much uncertainty regarding new phenomena such as Bitcoin, and with it comes different ideas on how to exploit this for own profit. Because Bitcoin, as a precursor to cryptocurrencies, had to stand for the adaption of those to the modern world, there have been difficulties in this process. It has followed that several legislatives in many countries were adapted or founded, and many regulations have been narrowing the applied freedom that were originally intended for Bitcoin, at least juridically.

It is also another aspect that should be addressed, and that is who is it actually that the users of Bitcoin should be responsible towards? Bitcoin was founded partially as a rebellion towards the existing monetary system, so it would be weird that one should be held accountable by the governing system one is rebelling against. Responsibility should at least be exercised when it comes to cyber security for your own money, because there is no one else to watch over them for you. This applies also for exchange companies that are holding money on behalf of others.

It is suggested that a discussion of the thesis findings and conclusions related to the core areas is included in this note. I don't think it is anything that is relatable between a statistical analysis and internationalisation, innovation or responsibility, maybe except being accountable for having all calculations done properly such that false information is not spread.

As a concluding remark, I would like to express my partial confusion, and mild frustration, over having to write this note, because, in my opinion, this is not aimed for students of finance to dwell on. To reflect on these topics is more in line for students studying concepts that are more organisational, empirical, or more physical in certain ways, not for those who are only studying on the theoretical plane. It does make little sense discussing this in our case, to reflect generally over the thesis, or have other, more relevant, topics to discuss would be much more informative and constructive for the student and the school.

10.5 Appendix E – Reflection notes: Filip Filipovic

This reflection note explains how this thesis is connected to the themes School of Business and Law at the University of Agder thinks are highly valued and relevant for students coming out from school to professional life. These themes are internalisation, innovation and responsibility, which is understandable considering worlds development in general where pretty much no borders exist anymore, looking in technological and innovative way.

The theme for this master thesis is Bitcoin and the analysis of its logarithmic returns on exchange rate against American dollar. There have not been many attempts in financial world to try and make a type of analysis we have conducted, so our goal was to make a model that can be used for further analysis in this field, since the concept of Bitcoin is relatively new and there are many aspects of it that have not been conceptualized and discovered yet. In the thesis we present different statistical distribution we used to fit the data to, and by the help of different criteria we chose the best fitted model. After all tests skew generalized t distribution showed up to be the best fit for our data set. I like to think that the results presented are more precise then the few other analyses, because we use the samples from the begging of Bitcoins trading start until April first of this year. It is by far the biggest data set looked at in this type of research, that we know of.

Bitcoin is very special phenomenon that most people haven't got enough knowledge to comprehend. The whole concept of Bitcoin is built on the one of the biggest innovations of our time, the blockchain. Bitcoins in its self is not that big of an innovation, because the virtual currency has existed for last thirty years. From innovative standpoint bitcoins cryptography part combined with blockchain makes it a ''bulletproof'' technology that can make life easier in many countries with underdeveloped economy and it can be used as a prevention against corruption. Since the bitcoin is still in its relatively early phases the full potential is probably still unknown, but possibilities are many. One thing is pretty sure though, and that it that bitcoin is not what is it without blockchain and that's why bitcoin from a standalone point can't lead towards innovation on its own but rather is an innovation.

Internalisation wise bitcoin can be regarded as an important tool in removing barriers in financial and banking world. Specially today when deals are made internationally across continents. Bitcoin reduces the speed of transactions and transaction cost and negates the middle man concept. The bitcoin makes everything more efficient and safe compared to banking, even though banking is considered safe, but the impact on internalisation might not be of the importance.

When it comes to responsibility, there is a big discussion regarding bitcoin. The biggest issues in the beginning were where people used bitcoin to order illegal substances and other illegal services on dark web without anyone being able to trace it back. Some of the bitcoin wallets have been traced back and connected (allegedly) to some firms and private persons but that is not an easy job. Governments have shut down some of the dark web sites, but it is not easy to completely block all of them and new ones pop up all the time. Here is where personal responsibility comes in, for everyone to not miss use the anonymity bitcoin provides.

Further issues have come up when European directive for personal information database has come out, where private persons can get all information saved about them by different companies and web sites deleted. The problem is that everything happening while using bitcoin gets registered in the blockchain and saved permanently so here is the problem for another discussion that must be solved regarding bitcoin.

This whole concept is still new considering financial world, and people have just touched tip of the iceberg in discovering his phenomenon. Governments and financial institutions should get together and make regulations and directives regarding the subjectivity of bitcoin and plan if or how to implement its use in the everyday life as easy as possible. For now, most of the world looks at the bitcoin as an asset instead of currency so maybe they should start there and regulate it as an acceptable currency.

These reflections are made within a context of bitcoin, so the standpoint view is different than usual, and it could be connected to some other values than the one mentioned. Through the study program with specialisation in finance I have learned to look at everything from different positions because things are always different and should be taken on with critical view. The society has a long and hard path to uncovering this exciting world of bitcoin and a lot of research must be done to establish theoretical theory that supports this whole concept. Maybe in some years the whole different story can be written considering these same values.