A bit risky? A comparison between Bitcoin and other assets using an intraday Value at Risk approach

Kristian Vagstad
Ole Christian Andreas
Valstad

◉ NTNU
Det skapende universitet

MASTERKONTRAKT
- uttak av masteroppgave

1. Studentens personalia

Etternavn, fornavn **Vagstad, Kristian**	Fødselsdato **28. apr 1989**
E-post **vagstad@gmail.com**	Telefon **97721758**

2. Studieopplysninger

Fakultet **Fakultet for samfunnsvitenskap og teknologiledelse**	
Institutt **Institutt for industriell økonomi og teknologiledelse**	
Studieprogram **Industriell økonomi og teknologiledelse**	Hovedprofil **Investering, finans og økonomistyring**

3. Masteroppgave

Oppstartsdato **15. jan 2014**	Innleveringsfrist **11. jun 2014**
Oppgavens (foreløpige) tittel **A bit risky? A comparison between Bitcoin and other assets using an intraday Value at Risk approach**	
Oppgavetekst/Problembeskrivelse Bitcoin has gained a lot of media attention in recent time. Due to the volatile behaviour of the Bitcoin price, some investors have warned enthusiasts and the curious public alike that Bitcoin is a risky investment. The aim of this master thesis will thus be to evaluate the intraday risk in Bitcoin and also compare it to other assets. To do this we will estimate a log-ACD-ARMA-EGARCH model using tick-by-tick data. We will then use this model to do a Monte Carlo simulation and use the results from this to compute an intraday Value at Risk. By doing this we will be able to make a risk measurement that takes into account both the information gained from the time between trades and price movements.	
Hovedveileder ved institutt **Post doktor Peter Molnar**	Medveileder(e) ved institutt
Merknader **1 uke ekstra p.g.a påske.**	

4. Underskrift

Student: Jeg erklærer herved at jeg har satt meg inn i gjeldende bestemmelser for mastergradsstudiet og at jeg oppfyller kravene for adgang til å påbegynne oppgaven, herunder eventuelle praksiskrav.

Partene er gjort kjent med avtalens vilkår, samt kapitlene i studiehåndboken om generelle regler og aktuell studieplan for masterstudiet.

Trondheim 2504-14
....
Sted og dato

K.Vagskal
....
Student

mohur
....
Hovedveileder

Originalen lagres i NTNUs elektroniske arkiv. Kopi av avtalen sendes til instituttet og studenten.

NTNU
Det skapende universitet

MASTERKONTRAKT
- uttak av masteroppgave

1. Studentens personalia

Etternavn, fornavn **Valstad, Ole Christian Andreas**	Fødselsdato **06. des 1988**
E-post **OCValstad@gmail.com**	Telefon **95122279**

2. Studieopplysninger

Fakultet **Fakultet for samfunnsvitenskap og teknologiledelse**	
Institutt **Institutt for industriell økonomi og teknologiledelse**	
Studieprogram **Industriell økonomi og teknologiledelse**	Hovedprofil **Investering, finans og økonomistyring**

3. Masteroppgave

Oppstartsdato **15. jan 2014**	Innleveringsfrist **11. jun 2014**
Oppgavens (foreløpige) tittel **A bit risky? A comparison between Bitcoin and other assets using an intraday Value at Risk approach**	
Oppgavetekst/Problembeskrivelse Bitcoin has gained a lot of media attention in recent time. Due to the volatile behaviour of the Bitcoin price, some investors have warned enthusiasts and the curious public alike that Bitcoin is a risky investment. The aim of this master thesis will thus be to evaluate the intraday risk in Bitcoin and also compare it to other assets. To do this we will estimate a log-ACD-ARMA-EGARCH model using tick-by-tick data. We will then use this model to do a Monte Carlo simulation and use the results from this to compute an intraday Value at Risk. By doing this we will be able to make a risk measurement that takes into account both the information gained from the time between trades and price movements.	
Hovedveileder ved institutt **Post doktor Peter Molnar**	Medveileder(e) ved institutt
Merknader **1 uke ekstra p.g.a påske.**	

4. Underskrift

Student: Jeg erklærer herved at jeg har satt meg inn i gjeldende bestemmelser for mastergradsstudiet og at jeg oppfyller kravene for adgang til å påbegynne oppgaven, herunder eventuelle praksiskrav.

Partene er gjort kjent med avtalens vilkår, samt kapitlene i studiehåndboken om generelle regler og aktuell studieplan for masterstudiet.

Trondheim 29/04-14

Sted og dato

Student

Hovedveileder

Originalen lagres i NTNUs elektroniske arkiv. Kopi av avtalen sendes til instituttet og studenten.

NTNU
Det skapende universitet

SAMARBEIDSKONTRAKT

1. Studenter i samarbeidsgruppen

Etternavn, fornavn **Vagstad, Kristian**	Fødselsdato **28. apr 1989**
Etternavn, fornavn **Valstad, Ole Christian Andreas**	Fødselsdato **06. des 1988**

2. Hovedveileder

Etternavn, fornavn **Molnar, Peter**	Institutt **Institutt for industriell økonomi og teknologiledelse**

3. Masteroppgave

Oppgavens (foreløpige) tittel **A bit risky? A comparison between Bitcoin and other assets using an intraday Value at Risk approach**

4. Bedømmelse

Kandidatene skal ha *individuell* bedømmelse
Kandidatene skal ha *felles* bedømmelse ✓

Trondheim 25/04-14
Sted og dato

Kristian Vagstad

Hovedveileder

Ole Christian Andreas Valstad

Originalen oppbevares på instituttet.

Preface

This Master's thesis concludes the authors' Master of Science degrees in Industrial Economics and Technology Management at the Norwegian University of Science and Technology (NTNU), in the Spring of 2014. The thesis is written by Kristian Vagstad and Ole Christian Andreas Valstad under the guidance of postdoctor Peter Molnar.

The cryptocurrency Bitcoin has caught the public eye recently, our attention included. The concepts and possibilities proposed by the Bitcoin protocol could potentially revolutionize the way monetary transactions are conducted. However the high volatility in the Bitcoin exchange rate has so far been a barrier for mass adoption. We therefore feel that analyzing the potential risk due to this volatility is crucial in order to investigate the current viability of Bitcoin as a medium of transaction. Also relevant when choosing this as the subject was our personal enthusiasm for the topic – we have learned a great many things along the way, not only limited to what are presented in this thesis.

The thesis has been written and edited in Latex. Figures are made in Matlab 8, Python 2.7 or Microsoft Excel. We have coded all models discussed here in Matlab 8 and Python 2.7 ourselves. Data and source code can be obtained upon request.

We would like to thank our academic supervisor Peter Molnar at the Department of Industrial Economics and Technology Management at NTNU for his helpful guidance and Nelly Semitela for thorough proofreading. All errors in this thesis are our own.

Trondheim, June 6th 2014

Kristian Vagstad and Ole Christian Andreas Valstad*

---

*E-mail addresses: vagstad@gmail.com and ocvalstad@gmail.com

Sammendrag

Bitcoin er en lovende kryptovaluta som har mottatt mye oppmerksomhet i senere tid, men den høye volatiliteten i Bitcoin-prisen har så langt vært en barriere for allmenn utbredelse. Gitt måten Bitcoin-transaksjoner fungerer, vil brukere av Bitcoin eksponeres mot svingende valutakurser, selv på korte intradaglige tidshorisonter. Denne avhandlingen gjør en analyse av denne risikoen i Bitcoin, og foretar en sammenligning mot mer tradisjonelle aktiva som gull og valutakursen mellom euro og Amerikanske dollar. Den senere litteraturen på ultra-høy-frekvent dataanalyse danner grunnlaget for en intradaglig Value at Risk-modell (IVaR). Denne IVaRen er basert på en Monte Carlo simulering, der en log-ACD-ARMA-EGARCH modell er blitt brukt til å beskrive de underliggende data. Resultatene viser tydelig at sammenlignet med gull og euro, er Bitcoin-brukere utsatt for en vesentlig større risiko på intradaglige horisonter. Dette kan potensielt sette en stopper for videre utbredelse av Bitcoin som transaksjonsmedium. Etter forfatternes mening er derimot ikke denne risikoen stor nok til å oppveie de potensielle fordelene Bitcoin tilbyr.

A bit risky? A comparison between Bitcoin and other assets using an intraday Value at Risk approach

Kristian Vagstad and Ole Christian Andreas Valstad

June 2014

Abstract

The promising cryptocurrency Bitcoin has attracted a lot of attention recently, but the high volatility of the Bitcoin price has so far been a barrier to widespread adoption. Given the way Bitcoin transactions work, users will be exposed to exchange rate risk even for short intraday horizons. This paper analyzes this risk, and compares it to more traditional assets, namely Gold and the Euro/USD exchange rate. To do this we make use of the recent literature on market risk measures for ultra-high-frequency data to construct an intraday Value at Risk measurement. This IVaR is based on a Monte Carlo simulation, where a log-ACD-ARMA-EGARCH model is used to describe the price movements. The results clearly indicate that for the intradaily horizon, Bitcoin is far more risky than Gold and Euro, which may challenge the applicability of Bitcoin as a medium of transaction. However in our opinion this risk is not large enough to outweigh the potential benefits that Bitcoin offers.

Contents

1 Introduction

Bitcoin has been described by many as a promising concept, which might however facilitate criminal activity (Fagan, 2013) and is far to volatile. We will not get into the illegal and criminal discussion here, but in this paper we study the inherent risk that Bitcoin investors, consumers and merchants face due to the high volatility.

While most people who are interested in Bitcoin see a bright future and thus consider it a long-term investment, it has also attracted a lot of speculators, especially in the wake of huge price spikes as those seen in April, November and December of 2014. Labeled as the new gold rush, speculators have flocked into this uncharted territory causing great profits for some, agony for others. Figure 1 reveals that even on shorter-than-daily horizons – there is evidence of high intradaily volatility in the Bitcoin exchange rate. In 2013 the difference between the highest and lowest traded price in one day was greater than 10% 33 times and greater than 5% 73 times.

Some Bitcoin prophets call it the money of the future (Petroff, 2013) and urge everyone they know to get in, while others warn strongly against it. The usual critics is that Bitcoin has no underlying value, it is too risky and even that it is a Ponzi scheme (Trugman, 2014). While the former can be discussed (Weisenthal, 2013) and the latter is not true (D'Angelo, 2013), given the volatility of the Bitcoin price, there is a considerable risk in Bitcoin that cannot be denied (Washington State Department of Financial Institutions, 2014).

If Bitcoin is to be a serious alternative to bank transfers and credit cards, consumers will have to start to use it regularly. However, given the high volatility and the risk this poses to users and merchants, this might be hard to achieve. It is the Catch-22 of Bitcoin – the volatility hinder mass adoption, but until the user group is large enough, the volatility will remain high. The solutions here seems to lie in the solutions developed for the Bitcoin ecosystem – make them so advantageous that the users do not care about the added risk.

Given the way Bitcoin transactions work, discussed in Section 2 and further outlined in Appendix A.1, merchants and consumers using bitcoins will be exposed to intradaily exchange rate risk. Since transactions can take up to an hour or more to confirm, assessing the risk that owners and users of bitcoins are exposed to during this period is an interesting and important topic. To our knowledge such an analysis does not exist at the time of writing. We will therefore explore this subject by making an intraday Value at Risk model. Given that Bitcoin can be thought of as both a currency and a store of value, we will do the same analysis for two assets with these characteristics, namely Gold and the Euro/USD

exchange rate, and compare the results. By doing this we will be able to quantify the potential additional risk level Bitcoin users are exposed to compared to traditional assets. In addition this risk assessment will also be useful for investors with shorter-than-daily time horizons.

The rest of this paper is organized as follows. In order to better understand the background for the exchange rate risk in Bitcoin, Section 2 provides the reader with a basic understanding of Bitcoin. In Section 3 we provide a review of the literature covering the methodology and models applied in this study. Section 4 presents the data used and its characteristics. Section 5 describes the model, while Section 6 presents the estimated parameters of the model. The simulation and backtesting procedures are presented in Section 7, whereas Section 8 presents and discusses the results. Finally Section 9 concludes.
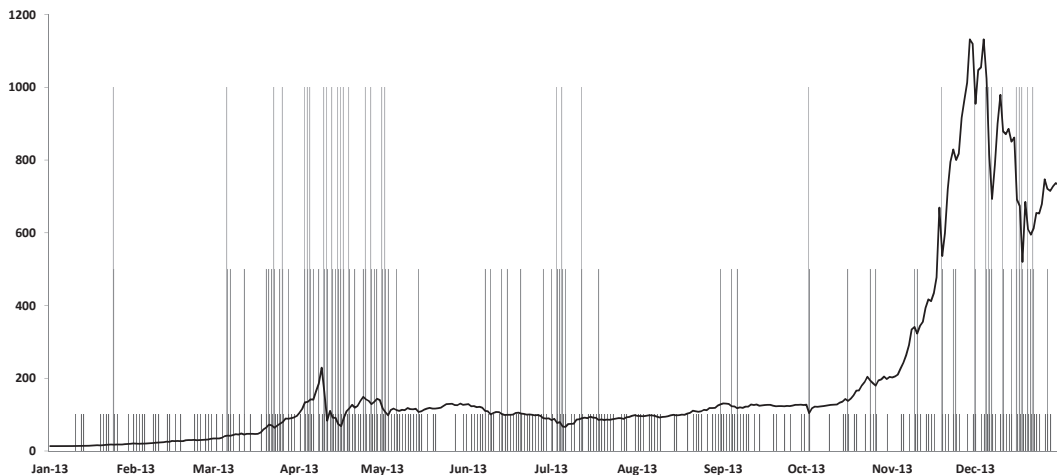
Figure 1: Bitcoin Price in 2013, where the lower columns marks days where the price difference between highest and lowest trade were greater than 1%, the middle columns mark days with greater than 5% difference and the high columns days with greater than 10% difference.

2   Bitcoin

In this section we present a basic introduction to Bitcoin. Some of the more technical aspects of the Bitcoin protocol are further outlined in Appendix A.1.

Bitcoin is a peer-to-peer payment system, first proposed in 2008 by Nakamoto[1].

---

[1]Satoshi Nakamoto might be one person or a group of people, the identity at the time of writing remains unknown.

The main difference between Bitcoin and other currencies is that the Bitcoin network does not require a trusted central authority for managing transactions and controlling the supply of money. Instead this is done collectively by the Bitcoin network. All Bitcoin transactions are kept in the block chain, a shared public ledger making everyone connected to the network able to see all transactions ever conducted. This is done by public key cryptography – anyone holding the public key can verify the validity of a bitcoin[2], but only the holder of the private key can use the bitcoin in a transaction. This chain-of-blocks structure also makes payments irreversible – if a transaction is considered verified, it is impossible to undo it[3]. In light of this Bitcoin can be thought of as the first widely accepted decentralized cryptocurrency, although the notion of integrating cryptography with electronic money was first attempted already in the late 1980s (Chaum et al., 1990).

The Bitcoin protocol also solves a more general problem known as the Byzantine Generals Problem (Lamport et al., 1982). In other words, Bitcoin poses a way to establish trust between otherwise unrelated parties over an untrusted network like the Internet[4]. Not only does this innovation enable Bitcoin to work as a currency and payment system, in general it also enables Internet users to safely and securely send digital property between each other. That way, everyone knows that the transaction has taken place and that no one can refute its legitimacy. Bitcoin as digital money is just one of many possible ways of utilizing this invention. Others can be digital signatures, digital contracts, digital keys and a lot more (Andreessen, 2014).

2.1 History of Bitcoin

Bitcoin was launched in January 2009 when version 0.1 of the Bitcoin client was released publicly. During its first year of existence, it was traded solely in private. The first known purchase using bitcoins, a pizza bought for 10 000 bitcoins (Mack, 2013), happened in May 2010. When the first Bitcoin exchange emerged a month later, the market could for the first time arrive at a consensus for the value of a bitcoin, the price being $0.08/BTC[5]. The development of the Bitcoin price since then is plotted in Figure 2. Not long after the first exchange had opened, others

---

[2]Technically "a bitcoin" does not exist, only the history of transactions between addresses. However, for simplicity we refer to it as "a bitcoin".

[3]This is what secures Bitcoin from the "double-spending" problem, the possibility of fraudsters spending their money several places at once.

[4]This has later been proved by Miller and LaViola (2014).

[5]BTC is the common symbol for bitcoin as a currency unit. Note also that Bitcoin with uppercase B refers to the protocol, while bitcoins with lowercase b refers to the currency.

emerged all over the world, allowing users to trade bitcoins for everything from Brazilian reals to Polish zloty. During the next two years adoption grew steadily, but among the general population it was still considered as an obscurity – most of the transaction volume was due to online gambling sites and dubious activities on the dark web (Meiklejohn et al., 2013).

Figure 2: Historical Bitcoin price in USD on logarithmic scale, source coindesk.com.

In 2013 Bitcoin moved from the underground of the Internet to the radar of Wall Street investors and governments all over the world[6]. The bank crisis in Cyprus in the spring of this year (Farrell, 2013) is said to be the main catalyst for the price peaking above $200/BTC, an all time high at the time. The market then collapsed, but still left the price well past previous levels. A summer of relative stability followed before Bitcoin yet again made the headlines. In October and November the price increased rapidly, fueled by extensive media coverage. On November 29th 2013 Bitcoin hit a record of $1242/BTC, higher than the price of one ounce of gold at the time. Meanwhile more and more merchants started accepting bitcoins as payment and Bitpay, a Bitcoin based automated payment processing system, announced that they had processed over $100 millions in transactions that year among more than 15500 merchants in 200 countries (Bitpay.com, 2013).

As a reaction to government regulations in China (Bradbury, 2013) and fear of interventions from governments in other countries, the price began to fall in December 2013 and early 2014. Later, a series of events led to a further price decrease

---

[6]Forbes probably summarized it better by calling it "Year Of The Bitcoin" (Forbes, 2013).

culminating when one of the biggest and oldest exchanges, namely Japan based Mt.Gox, declared bankruptcy in February 2014 (Southurst, 2014). In the following months the price has seemed to stabilize around $500/BTC and where Bitcoin goes from here remains unknown.

2.2   Trading process

There are two ways to obtain bitcoins: One can mine them, as explained in Appendix A.1, or one can buy them. There are numerous ways to buy bitcoins, the simplest being meeting someone who already owns bitcoins, giving them cash or transferring them money, and then they will transfer bitcoins to you. Sites like localbitcoins.com specialize in this kind of trading, by matching buyers and sellers in the same area. This kind of trading gives the highest degree of anonymity[7]. The most common way to obtain bitcoins though is through an exchange.

Bitcoin exchanges works very much like exchanges for other assets. You sign up[8], transfer money[9] and buy bitcoins from a seller. The choice of exchange depends heavily on your country of residence, as the sign up and money transferring process can be different depending on where you are located[10]. So even though Bitcoin presents an alternative to traditional fiat currencies, it is still closely tied to the existing financial infrastructure. In this article we investigate the price at two of the most popular exchanges, Bitstamp and Btc-e.

After you have acquired some bitcoins at your exchange account, you want to transfer them to your Bitcoin wallet. A Bitcoin wallet has the same purpose as other wallets, a place where your valuables are safely stored. Each wallet has one or more associated addresses. These work similarly to bank account numbers, you send bitcoins from the address of the exchange account to your wallet address. With bitcoins in your wallet, spending them comes down to sending from your wallet to merchants, or other people.

Merchants have several ways of accepting Bitcoin payments in their store or on their website. The simplest way is just to make users transfer bitcoins to the company's wallet, like a cash transfer. The price can either be quoted directly in

---

[7]The participants in such a deal do not need to leave any other digital footprint than the Bitcoin transaction.

[8]This process can be lengthy for some exchanges as it requires you to identify yourself and give a proof of residence.

[9]Bank transfer is the most common, although a lot of other exist, see The Bitcoin Wiki (2014b).

[10]Coinbase, a Bitcoin service provider based in the US, only allows bank transfers from US bank accounts.

bitcoins, or it can be priced in any currency and at the time of purchase converted to bitcoins using the current exchange rate. There is also the opportunity of using a Bitcoin payment service provider, such as BitPay. By doing this merchants can get paid in fiat currency, while still accepting bitcoins. This is done by letting the provider handle the Bitcoin part of the purchase, i.e. they find the current conversion rate, accept payment in bitcoins and then transfer fiat.[11].

Given the volatile history of the Bitcoin price and the non-zero block confirmation time previously discussed, there is a considerable risk involved in doing payments with bitcoins, since the exchange rate between Bitcoin and fiat could have changed quite a bit in that short time span. For the different payment models different parties are exposed to this risk. If the merchant prices his goods in bitcoins and is paid in bitcoins, this risk can be ignored[12]. If the merchant accepts payment in bitcoins, but the price is set by the current exchange rate, the merchant can choose – he can assume the risk by guaranteeing the exchange rate at the time of purchase[13], or let the buyer take it by setting it to the exchange rate at time the network verifies the transaction[14] By using a Bitcoin payment service provider, this risk is usually assumed by them. It is out of the scope of this paper to discuss who should bear this risk in general, our goal is to quantify it.

2.3   Advantages of Bitcoin

For many of the pioneers of Bitcoin the most important feature is how the trust model in Bitcoin is built, or rather the lack of it. Given that the Bitcoin protocol poses a way to establish trust between otherwise unrelated parties over an untrusted network like the Internet, users do not need to trust anyone when making a transaction. The need for providing sensitive information like credit card details in order for the transaction to proceed is removed, and giving the recipient any other personal information is voluntary. While one usually need to provide merchants with such information, having the opportunity to choose not to can be a big benefit, e.g. for anonymous donations. The advantage of not leaving credit card details seems obvious given the amount of stolen credit card cases in recent years (Riley et al., 2014). The possibility of doing business with someone who one neither trust nor know, can be enticing for many individuals.

This goes both ways, the privacy in Bitcoin is entirely up to the user. Bitcoin is

---

[11]Usually one have the option of getting paid 100% in fiat, 100% in bitcoins or a mix.

[12]However, if the merchant has any costs in fiat he is exposed to exchange rate risk anyway.

[13]Known as zero-confirmation – the merchant does not wait for the block chain to confirm the transaction.

[14]Details about how the Bitcoin network verifies transactions can be found in Appendix A.1

pseudo-anonymous – all transactions are publicly available, but the owners of the sender and receiver addresses are kept secret. However, it is worth noting that in practice it can be possible to identify owners of certain addresses, especially if the owners do not take care to avoid cases such as multiple usage of the same Bitcoin address. Meiklejohn et al. (2013) explore this subject in greater detail.

In traditional fiat systems people are dependent on a reliable government to manage the supply of money. However, as history shows, governments only seem to be trustable as long as it is to their benefit (Jones, 2014). The average life expectancy for a fiat currency is 27 years (Mack, 2011), and even the oldest and most successful fiat currency in existence, the British pound Sterling, has lost 99.5% of its initial value[15]. In the Bitcoin system, money supply is an entirely predictable process. Bitcoins are mined at a predefined decreasing speed until the maximum number of mined bitcoins has been reached, which is 21 millions. This removes the possibility for hyperinflation and makes Bitcoin rather deflationary[16]. Whether this is good or bad in the long run is widely discussed, see for instance The Bitcoin Wiki (2014a) for a discussion.

The trust issue is also relevant when it comes to being completely in control of your own money. If you have an account on PayPal for instance, and they decide that you broke their policy, they stand free to restrict your access to the account and thus limit your access to your own funds. This is what happened in 2010 when the secrets-and classified information-revealing organization WikiLeaks was shut out from the service (Moyer, 2010). One might also think that this only happens when you trust third parties, not the government. However, as previously mentioned, one of the catalysts of the rising Bitcoin interest in 2013 happened when the Cypriotic economy was in trouble, and the Central Bank, with the help of the EU and IMF, imposed a one-time bank deposit levy for domestic accounts (Bensasson and Georgiou, 2013). Situations like these can be avoided with Bitcoin. As long as the user is the only one with access to the private keys[17], no one can touch his bitcoins.

The Bitcoin system also has major advantages when it comes to speed. Because of the trust model used by the banks, bank transfer and international wire transfer can take a considerable amount of time to clear, usually only within the opening hours of the bank. With this centralized structure, there is a great need to make sure that the transaction indeed is trusted and valid before a confirmation can be made. The Bitcoin network takes care of this within minutes, and after a few

---

[15]The British pound was originally defined as 12 ounces of silver.

[16]The value of bitcoins will rise over time compared to inflationary assets such as fiat.

[17]However, if he trust a third party such as an online wallet provider or an exchange, the same situations can occur.

hours at the most[18] users can be completely sure that their transactions have gone through.

Another advantage of the decentralized structure of the Bitcoin network is the lack of bounds for when and where one can send money. Without the need for a centralized authority doing the authorization, sending money to a neighbor and to the other part of the world is the same short procedure, which can be done at any given time, not just within the opening hours of the banks. For some this can be very beneficial – payments from anywhere in the world can be accepted 24 hours a day, thus greatly lowering the boundaries between merchants and customers around the world.

The price one has to pay for this is not discouraging, for the most part it is free or close to – the transaction fee is still voluntary[19]. One might argue that credit card transactions seem to be both instantaneous and free. However, this is a service the merchants need to pay for, which in the end leads to higher prices for customers. The size of this fee can vary quite a bit, but around 3% does not seem to be unusual. For low margin retailers the credit card fees can account for as much as half of the profit (Harrison, 2014).

What is also worth to note is that because of the non-reversible transaction structure of Bitcoin, no chargebacks are allowed[20]. In 2013, online merchants paid $3.10 for each dollar of fraud losses incurred according to LexisNexis Group (2013). This amount includes fees and interest to financial institutions and costs related to replacing lost or stolen merchandise, in addition to the amount of chargebacks for which their company is held liable. A chargeback is a form of customer protection provided by the issuing banks, which allows cardholders to file a complaint regarding fraudulent transactions on their statement. The study done by Lexis-Nexis Group (2013) interestingly finds that 20% of fraudulent transactions are due to "friendly fraud", where a consumer makes an online purchase with their credit card, then issues a chargeback after receiving the purchase, claiming the purchase was never delivered or accounted for. Frauds due to identity theft resulting from lost credit card information made up another 17% of frauds in 2013, which happened to over 5% of all American consumers, causing losses of over $21 billions. So it is easy to understand why no chargebacks is an appealing feature for merchants, and that the anonymity of transactions which protects against identity fraud is beneficial to both parties[21].

---

[18]Usually a lot less.

[19]However, by including a small transaction fee one might reduce the confirmation time.

[20]Unless the recipient agrees.

[21]We are not implying that frauds are not happening in the Bitcoin economy, but we emphasize the fact that responsibility for avoiding fraud is in the hands of the person sending the bitcoins,

Lastly we remark an area that really seems to have a lot to gain from the advantages of the Bitcoin system, namely the remittance[22] industry. The World Bank (2014) estimates that the total value of remittances to developing countries in 2013 is at $404 billions. The average total cost of sending remittances was around 8.4% in the first quarter of 2014, but for some parts of the world it was much higher, e.g. the average total cost for sending remittances to Sub-Saharan Africa was as high as 12%. The potential Bitcoin has to impact these mechanisms is enormous, given the before-mentioned properties of the Bitcoin system.

3 Background

3.1 Risk

So far we have stated that we want to investigate the riskiness Bitcoin investors and users are faced with compared with that of other comparable assets. First of all we have to define what we actually mean by risk. On the general level Hansson (2014) states that "... the word "risk" refers, often rather vaguely, to situations in which it is possible but not certain that some undesirable event will occur". More financially related is Jorion (2007)'s definition, who defines risk as "the volatility of unexpected outcomes". This last definition is common when regarding financial assets, risk does not necessarily mean that something bad is more likely to happen, it just refers to a case where the potential outcome is less certain. Because of this some risk is necessary if performance over risk-free is to be achieved (Lamb, 2011).

A common way of assessing the risk of an asset, would therefore be to look at the volatility of the asset. Volatility is a measure for the dispersion of the returns for a given asset, i.e. it gives us an indication to how much uncertainty there is in the returns. With a lower volatility, the range the returns are expected to be in is much lower than with a higher volatility. We see from our previous definitions that this gives rise to a conflict – while Jorion's definition fits well with volatility as risk measurement, Hansson specifically mentions "undesirable events", i.e. in our case negative returns. Given that we are interested in the potential negative outcomes of using Bitcoin, we use Hansson's definition.

---

because only the person who owns the bitcoins has access to them.

[22]A transfer of money by a foreign worker to an individual in his or her home country.

3.2   Value at Risk

A natural risk measurement for us to use is therefore Value at Risk (VaR). Value at Risk is defined as the worst loss over a target horizon that will not be exceeded with a given level of confidence (Jorion, 2007). This fits well with our definition of risk. Other advantages that using VaR gives us is that we get an intuitive interpretation of the risk level – which is also accessible for non-financial people, as well as giving us a quantifiable risk level which can be used in a comparison to other assets.

There are three basic approaches to calculate Value at Risk, but within each approach there is still room for a lot of variation (Damodaran, 2014). We will not go into detail here, but shortly mention the three:

Parametric VaR
    The first one can be called Parametric VaR, where the estimate is computed analytically by making assumptions about the risk factor return distributions, e.g. a normal linear VaR.

Historical Simulation VaR
    The historical simulation model makes no assumptions about the distributions, but rather use a vast amount of historical data to estimate VaR. The main assumption here is that all possible future movements have already happened, and thus that historical distribution will be the same as the distribution of future returns.

Monte Carlo VaR
    The core of the Monte Carlo VaR approach is to run Monte Carlo simulations for the proposed models and assumptions enough times, so that we are fairly sure that the simulated distribution gives a proper estimate of the real distribution. On the basic level this can be quite similar to the normal linear VaR, but the flexibility of the Monte Carlo approach is far greater than the parametric.

For our purposes the Monte Carlo method seemed like the obvious choice. The flexibility it offers compared to the others is invaluable, given that we explicitly want to model the time between trades, as described in Section 5.

However, when using VaR there are several things we need to keep in mind. While the interpretation of VaR is quite intuitive, it may also lead to a false level of security. For instance if the VaR at the 99th percentile is interpreted as the maximum we can possibly lose, we may feel much more secure than we really are, given that 1% of the time this level actually should be exceeded. This also

leads to the second point – even though VaR gives us a loss that we, with the given confidence level, are certain will not be exceeded in the chosen time horizon, it gives us no information about what happens if we actually exceed this level[23]. VaR thus gives no information about the maximum possible loss we stand to suffer. Because of these (often overlooked) limitations, it is important for risk managers to treat VaR as just another tool in the toolbox, and not the only way to measure the risk level they are exposed to (Linsmeier and Pearson, 2000). However, for our purposes it functions well as an intuitive and quantifiable level of risk.

3.3   Ultra-High-Frequency data

There has been a growing interest in studying ultra-high-frequency data as such data have become increasingly accessible to researchers. The quest of studying ever higher frequencies has been a quest for better information. In this sense we think of the frequency of a time series as the informativeness of the process that creates it. The more often it is observed, the more information can be extracted from it and better decisions can be taken with respect to it. In this paper we follow the distinction between high-frequency data (HF data) and ultra-high-frequency data (UHF data) as stated by Zivot (2005). HF data refer to observations on variables recorded on daily or intradaily time scales where the observations are often, but not necessarily irregularly spaced over time. Whereas UHF data are tick-by-tick or quote-by-quote time stamped observations and are thus irregularly spaced in time by definition.

The nature of UHF data makes it challenging to model and analyze both econometrically and statistically. Since the observations are irregularly spaced in time, the number of observations over a fixed time interval will be random. One can of course aggregate this data up to regular time intervals, however this will incur a loss of information. Further trading activity will vary substantially between different assets. Secondly one must also take care to perform proper cleaning and correction of data, since UHF data often will contain errors, for instance improper records or data gaps. Also UHF data will typically have intraday and/or inter-week patterns in activity. Furthermore issues such as bid-ask bounce may distort some statistical models. Lastly the number of observations can be stupendous, it is not uncommon to have more than 10 000 observations per trading day.

---

[23]In order to assess this one can look at Conditional Value at Risk, which measures the expected loss given that the VaR level has been exceeded.

3.4   Intraday Value at Risk

Much effort has been put into creating sophisticated Value at Risk models for daily or longer time horizons, encouraged by the legal obligations banks and insurance companies are put into by the Basel and Solvens regulations. For shorter time periods however the field is much more unexplored. However, given the relatively recent access to intradaily data, it has finally started to grow.

Intraday VaR (IVaR) models was first proposed by Giot (2005), where he uses intradaily returns to estimate a VaR model. Two distinct ways to model the intradaily returns are already here made clear. UHF data is naturally irregularly spaced, while most normal VaR techniques use regularly spaced intervals. One therefore has the choice to either convert the UHF data to regular intervals in order to use the standard framework, or to make a VaR model based on the irregular spaced returns. Giot does both, by making conditional parametric VaR models with a normal GARCH, Student GARCH and RiskMetrics approach for the regular spaced interval, and by utilizing a log-ACD model for the irregular spaced data. While these are not directly comparable, Giot concludes that the Student GARCH model performs best.

Giot and Grammig (2006) uses the same conversion from UHF data to regular returns, but follow quite a distinct approach, in that they introduce the notion of an Actual VaR that models the potential price impact incurred by liquidating a portfolio. The further approach is to use a reconstructed real time order book from the automated auction marked Xetra taken at 10- and 30-minutes frequencies, to compute returns that depend on volume. These are then used to compute a VaR for each level, and by comparing these they are able to quantify the relative liquidity risk premium within a day.

A third way of computing the returns for an intraday Value at Risk is seen in Dionne et al. (2009). The approach is a development of the log-ACD VaR model proposed by Giot (2005), but while Giot uses the log-ACD model to estimate the volatility of price durations, Dionne et al. specifically models the whole trading sequence. More specific, they model the time between trades using a log-ACD model, and then use an ARMA-EGARCH model for the returns between trades that takes into account the duration between trades. By using this combined model in a Monte Carlo simulation, they are able to simulate the trading process, and are then free to divide into regular spaced intervals in order to assess the VaR in the regular framework. This differs from Giot's approach, where the log-ACD VaR model is based on irregular spaced intervals, and then converted by an approximation in order to be compared to the other models. The backtesting of these reveal that they perform quite differently. While Giot's log-ACD disappoints, Dionne

et al.'s model performs quite well at most intervals and confidence levels[24]. Other differences in the models lie in the specification of duration as trade durations and not price durations, as well as utilizing Monte Carlo simulation in order to get a model that is easier to convert to regular intervals. According to Dionne et al. the advantage of this model is thus that the risk measure has a higher information content due to the modeling of each trade, and not just the ones that lead to a price difference. What is also worth mentioning is that by following this approach, after one has estimated the model, a VaR can be computed for any time horizon, without re-estimation of the model parameters.

Colletaz et al. (2007)'s work is similar, however they apply the VaR methodology in price events times, rather than calendar time. The result is an irregularly spaced intraday Value at Risk, that both forecast the timing for the next price event and the corresponding level of risk summarized by the VaR forecast. The proposed model is applied on two NYSE traded stocks with relatively good results. Tse and Liu (2013) calculate an IVaR on ten NYSE traded stocks, and find that results can be improved by modeling price movements and durations jointly. Overall we see that the number of papers investigating the field of intraday Value at Risk is not very large, however the reported results seem to be satisfactory for most models.

3.5   The Autoregressive Conditional Duration model

The features of UHF data described earlier combined with the increasing availability of such data has encouraged a number of contributions dealing with the challenges of accurately model UHF data in the domain of econometrics. Engle and Russell (1997) proposed the Autoregressive Conditional Duration (ACD) model with the objective of modelling the time between events, like trades or quote updates that happen irregurarly and randomly throughout market hours. It is essentially an ARMA process with nongaussian innovations. The ACD model assumes that the durations follow a process similar to the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model developed by Bollerslev (1986). Duration in this context is simply the time between trades. As for GARCH models, ACD models are typically called for when the data exhibits a high dependency from one data point to another. This feature is most commonly referred to as clustering and is often observable in ultra-high-frequency data. This means that short durations tend to be followed by short durations, and similarly long durations are often followed by long durations. Just as GARCH models aim to remove dependence in squared returns, ACD models aim to remove dependence

---

[24]Their model is applied on three stocks traded on the Toronto Stock Exchange.

in the durations. Here we will provide the general setup for ACD models, and provide a short review of the most commonly used.

Statistically we view ultra-high-frequency data as a point process, which we denote N. In the general sense a point process is a random collection of points (observations) falling in some space. In UHF finance the space in which we record observations is simply a portion of the real line, because the data is temporal ("of time"). Such a point process can be described completely by the random intervals between the points as they are ordered in a natural way. The arrival times of the point process, represent events like a trade or a quote update. Events then have different characteristics associated with them, like price and volume, and we refer to them as marks.

A temporal point process N may alternatively be described by a counting process $N(t)$, where for any time t between 0 and T, $N(t)$ is the number of points occurring before time t. We then consider a sequence of arrival times of events $\{t_0, t_1, ..., t_n, ...\}$ where $0 = t_0 < t_1 < ... < t_n < ...t_{N(T)}$, where $t_{N(T)}$ is the last observed point in the sequence. In the most general setting, simultaneous arrival times are allowed, but as described in 4.3 we have excluded such events. In the literature this is standard procedure, see for instance Dufour and Engle (2000-05). $\{z_0, z_1, ..., z_n, ...\}$ denotes the sequence of marks associated with the arrival times.

Further let $x_i = t_i - t_{i-1}$ be the ith duration between the events happening at times t_{i-1} and t_i. Then the sequence of durations have non-negative elements all greater than zero. Durations and marks can then be considered as a joint sequence:

$$\{(x_i, z_i), i = 1, .., T\}$$

The ith event in the joint sequence then has a density conditional on all previous durations, up to and including x_{i-1}. It is referred to as the information set available at time t_{i-1} denoted $\mathscr{F}_{t-1}$. The joint density is then:

$$(x_i, z_i)|\mathscr{F}_{t-1} \sim f(x_i, z_i|\breve{x}_{i-1}, \breve{z}_{i-1}; \Theta_f) \tag{3.1}$$

Where Θ_f denotes the set of parameters and $\breve{x}_{i-1}$ and $\breve{z}_{i-1}$ the sequence of X and Z up to observation $(i-1)$. This joint density can be expressed as the product of the marginal density of durations and the conditional density of the marks given the durations, formally:

$$f(x_i, z_i|\breve{x}_{i-1}, \breve{z}_{i-1}; \Theta_f) = g(x_i|\breve{x}_{i-1}, \breve{z}_{i-1}; \Theta_x)q(z_i|x_i, \breve{x}_{i-1}, \breve{z}_{i-1}; \Theta z) \tag{3.2}$$

Where g is the marginal density of the duration x_i and q the conditional density of the mark z_i

The ACD model imposes a generalized autoregressive structure to the conditional duration. The conditional expected duration is given by:

$$\psi_i = E(x_i|\mathscr{F}_{i-1}) = \psi_i(\breve{x}_{i-1}, \breve{z}_{i-1}) \tag{3.3}$$

So the expected conditional duration, ψ_i is measurable with respect to $\mathscr{F}_{i-1}$. Further, it is assumed that all the temporal dependence in the sequence of durations are captured by the expected conditional duration, so we have:

$$\frac{x_i}{\psi_i} = \epsilon_i \tag{3.4}$$

The epsilons are independent and identically distributed random variables with $E(\epsilon_i) = 1$. The sequence of ϵ_i is often referred to as standardized durations and Eq. (3.4) the mean equation. This assumption implies that the marginal density of duration x_i is $g(x_i|\breve{x}_{i-1}, \breve{z}_{i-1}; \Theta_x) = (x_i|\psi_i; \Theta_x)$, i.e. the conditional expected duration captures all temporal dependence of the duration process.

The density function for ϵ with parameters Θ_ϵ, $p(\epsilon, \Theta_\epsilon)$ must satisfy $g(x_i|\mathscr{F}_-; \Theta) = \psi_i^{-1}p(x_i/\psi_i; \Theta_\epsilon)$. Maximum likelihood estimates of $\Theta = (\Theta_x, \Theta_\epsilon)$ can be obtained by performing numerical optimization after specifying a distribution of ϵ and using the log-likelihood function given by:

$$L(\Theta) = \sum_{i=1}^{N(T)} (\log(p(\frac{x_i}{\psi_i}; \Theta_\epsilon) - \log(\psi_i)) \tag{3.5}$$

The ACD model proposed by Engle and Russell (1998) is a linear parametrization of Eq. (3.3) where the conditional expected duration depends on m past durations and q past expected durations. It is commonly referred to as the ACD(m,q) model and is given by:

$$\psi_i = \omega + \sum_{j=1}^{m} \alpha_j x_{i-j} + \sum_{j=1}^{q} \beta_j \psi_{i-j} \tag{3.6}$$

There have been numerous extensions to the original ACD(m,q) model. Worth mentioning here is the log–ACD model of Bauwens and Giot (2000), which unlike the standard ACD(m,q) does not need non-negativity constraints to ensure

positivity of conditional durations. Fernandes and Grammig (2006) introduces augmented ACD (AACD) models that allow for asymmetric responses to small and large shocks. To allow for different regimes of heavier or slower trading period, regime switching ACD models allows for greater flexibility, see for instance the TACD model by Zhang et al. (2001).

To further expand the range of possible methods to model durations, several choices exist for the distribution of the standardized durations, ϵ_i. In the simplest case, the exponential distribution is considered, and models using this distribution are often referred to as EACD models. However this implies that the conditional hazard rate is flat (independent of time) which is easily rejected in most studies, see for instance Feng et al. (2004) or Lin and Tamvakis (2004). The hazard rate is the rate of likeliness to observe an event of interest in the next time unit given that the event is not observed up until that point. However this issue was already addressed by Engle and Russell (1998) and they therefore used the Weibull distribution (WACD) for increased flexibility. To allow for even greater flexibility Grammig and Maurer (2000) argued for the use of the Burr distribution (Burr-ACD), while Lunde (1999) proposed the use of the generalized gamma distribution (GACD). They both allow for a U-shaped hazard rate, to capture that the hazard rate tends to increase for short durations and decrease for long durations. Even though the research on ACD models has been extensive since its inception, the quest to find best practices for the specification of the model and the distribution of ϵ has not yet reached its conclusion. In our case where the expected conditional durations enter the conditional variance equation as an explanatory variable, the form of ACD model and distribution of ϵ is of great importance.

3.5.1   Durations and volatility modeling

Durations do not necessarily need to be defined as the time between every transaction, however more information may be gained by looking only at arrival times of certain events. For instance it can be a certain change in the price, or the time necessary for a given amount of volume to be traded or the time it takes to observe a given return. It then follows that we only are interested in retaining the arrival times that are thought to carry some special important information, a procedure that is most commonly referred to as "thinning the point process" (Hautsch, 2012).

We turn to the concept of price durations first, which was introduced by Engle and Russell (1997). It is defined as the duration between events for which the price has changed. A point process for price arrival times is obtained by selecting only those points for which the price has changed. However one is not particularly interested in small price changes, more so the times necessary for the price to change by a

given amount C. Price durations in itself is a simple and easy-to-interpret risk measure, the longer the time it takes to change the price, the less volatility exists. Giot (2005) makes use of price durations to construct an intraday Value at Risk model.

The importance of price durations comes from its direct relationship with volatility. As demonstrated in Engle and Russell (1998), by applying the ACD model to price durations, a model for the inverse of volatility is obtained by using the conditional expected price durations via the conditional hazard function. Formally it is given by:

$$\sigma^2(t|\mathcal{F}_{i-1}) = \left(\frac{C}{P(t)}\right)^2 h(x_i|\mathcal{F}_{i-1}) \tag{3.7}$$

Where $\sigma^2(t|\mathcal{F}_{i-1})$ is referred to as the conditional instantaneous volatility of the price, $P(t)$ is the price at transaction time t_i and $h(x_i|\mathcal{F}_{i-1})$ is the conditional hazard function of price durations. If an EACD model is used, $h(x_i|\mathcal{F}_{i-1})$ becomes $1/\psi_i$, and volatility can be modeled by using:

$$\sigma^2(t|\mathcal{F}_{i-1}) = \left(\frac{C}{P(t)}\right)^2 \frac{1}{\psi_i} \tag{3.8}$$

This link between price durations and volatility makes it possible to test various market microstructure hypotheses in a straightforward way by including additional explanatory variables into Eq. (3.6), so that:

$$\psi_i = \omega + \sum_{j=1}^{m} \alpha_j x_{i-j} + \sum_{j=1}^{q} \beta_j \psi_{i-j} + y^T z_{i-1} \tag{3.9}$$

Where as before expected conditional durations depend on past m durations, past q expected durations, but now also a vector of non-negative exogenous variables z_i. Relevant variables commonly used in the literature include the trading intensity and the average volume per trade, see for instance Bauwens and Giot (2000), Bauwens and Veredas (2004) or Bubak and Zikes (2006).

Trading intensity is defined as the number of transactions during a price duration, divided by the value of this duration. A high number of trades over a short duration leads to high trading intensity, and consequently a low number of trades over a long duration leads to low trading intensity. The hypothesis is that an increase in trading intensity should make the market maker revise his/her quotes

more often, leading to shorter price durations. Easley and O'Hara (1992) discusses this relationship in greater detail. The average volume per trade defined as the average of the volume of trades made during a price durations, is thought to be an indication of informed trading. Again turning to Easley and O'Hara (1992), the idea is that a higher volume indicates informed trading[25] and leads to higher volatility.

For comparison purposes, price durations is not directly applicable since a good choice of C depends on which asset one look at[26]. Therefore one may specify the durations on the returns directly as the time it takes to observe a log return of at least R. The log returns are then calculated continuously for each consecutive transaction from an initial price. We refer to this concept as return durations.

4   Data

4.1   General

The chosen data period runs from January 1st 2014 – February 28th 2014. The sample is divided into an estimation and a validation period, where we use January for estimation and February for validation. The number of observations varies quite a bit between the assets, Bitstamp has the lowest number of observations with 873 560 in the sample period, while Gold has the highest with over 10 million observations.

We have obtained our data from two different sources. Bitcoincharts.com has been used for Bitcoin data, while Sierra Chart has been used for Euro and Gold. Sierra Chart is a software program for technical analysis of financial markets that allows access for a limited number of tick data time series. From Sierra Chart we downloaded data from Forex Capital Markets, which is an online broker based in the United States. It provides services to their clients through an online trading platform and their offices all over the world, from New York to Sydney. It uses direct market access that allows clients to trade at the best price at any given time, taking prices from a number of major banks and institutions.

From Bitcoincharts we downloaded tick by tick data for two exchanges[27], Bitstamp

---

[25]In Easley and O'Hara (1992) the idea is that informed trades trade only when there are information events, whether good or bad, that influence the price

[26]For instance for Bitcoin it can be sensible to set C at a couple of dollars, but that does not make much sense in the Euro/USD exchange rate

[27]We have included two exchanges in our analysis, because the price, volume and trade dynamics can be quite different across the various Bitcoin exchanges. Using two of them, we can

and Btc-e. Even though Bitcoincharts provides data from a long list of Bitcoin exchanges, we chose these two for the following reasons: After the previously mentioned downfall and collapse of Mt.Gox, Bitstamp has emerged as the market leader in terms of trading volume. Based in United Kingdom, it allows customers from all over the world to trade bitcoins in exchange for US dollars, and is regarded as a secure and trusted exchange in the Bitcoin community. After coming online in August 2011, Btc-e provides Bitcoin trading in US dollars, Euro and Russian Rubles. Measured in number of trades per day, it is one of the largest Bitcoin exchanges in the world.

4.2   Trading hours

The normal approach to only regard the time between opening and closing as trading hours, and remove all observations that fall outside this interval. Bitstamp and Btc-e have customers from all over the world, and are open 24 hours 7 days a week. As Figure 3 shows, the trading intensity seems to be fairly high and stable throughout the day. It is therefore obvious that we must include 24 hours of trading each day, including weekends.

For Gold and Euro, 24 hours of data for each day excluding weekends is also available, however the trading activity varies considerably throughout the day. We therefore tried to find an estimated trading day for each asset, consisting of the most active part of the day. We divide the day into 48 half-hour intervals, and choose the first interval of the trading day as the first interval where the number of transactions is larger than the average number of transactions. The last interval in the trading day will then be the last interval where this is the case. This resulted in a 9 hour trading day for Euro and a 18 hour trading day for Gold.

4.3   Zero durations

A drawback of most ACD models existing in the literature is that they do not permit durations equal to zero. This is due to the fact that the distributions used for durations are not defined at zero, with the exception of the exponential distribution, which is rarely regarded as the best choice. The smallest time increment in the collected data is one second, and it therefore follows that transactions within one second have the same timestamp. For dealing with zero durations we followed the most common approach in the literature as used by Engle and Russell (1998) and Dionne et al. (2009). An average price weighted by volume is computed, and

---

be more confident about our results.
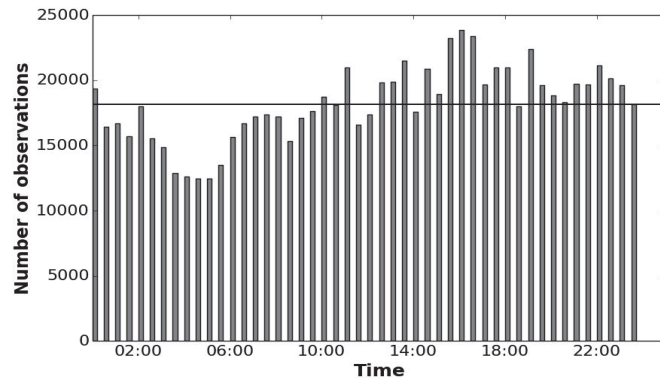
(a) Bitstamp
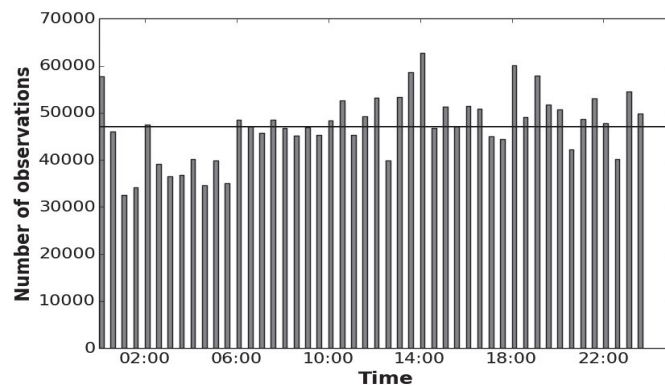
(b) Btc-e

Figure 3: Columns shows the average number of observations for a given half hour interval. The horizontal line marks the average number of observations across all intervals.

all other observations within that time stamp are removed. The argument for this approach is that we can consider these observations as split transactions, originating from large orders broken into smaller ones. The percentage of zero durations as reported in Table 1 may seem dramatically high for some of the datasets, but are not uncommon findings in the literature, for instance Engle and Russell (1998) reports around two third of the observations in their IBM dataset.

Remember from Section 3.3 that UHF data often contains errors. Our collected data was no exception, and we carefully removed suspicious observations. For instance we found observations with negative volume or volume equal to zero, and these were simply removed from the datasets. Similarly, consecutive observations with extreme price changes, typically when an observation has an artificially low price compared to the previous one and then reverting back to the "normal" price on the next observations, were also removed.

Table 1: Percentage of zero durations.

	Percentage of zero durations
Bitstamp	54.87%
Btc-e	66.51%
Euro	83.15%
Gold	86.01%

In our datasets we also take notice of abnormal events like the flash crash on Btc-e on February 10th 2014, where a single user dumped thousands of bitcoins and sent the price plunging from over \$600 to \$102 in a matter of minutes. While the rest of the Bitcoin ecosystem seemed largely unaffected, it caused massive turmoil on Btc-e over the next couple of hours before reverting back to match a price consistent with the rest of the market. It is obvious that including this short period in our analysis would distort the seasonal adjustment procedure, not to mention it would be next to impossible to incorporate in any model, and was therefore removed. Further we follow the common practice of eliminating all interday (overnight) durations in Gold and Euro because they would distort the results.

4.4   Data characteristics

After filtering the raw data each asset ends up with the descriptive stats for the durations in the estimation sample as reported in Table 2.

Table 2: Descriptive stats, durations.

	Mean	Median	Std.dev	D	Skew	Kurtosis	Max	Min	R_1	$Q(15)$	Obs
Bitstamp	13.50	6.0	23.78	1.76	6.14	80.55	898	1	0.26	53 070	196 546
Btc-e	6.55	3.0	10.85	1.66	4.98	46.04	312	1	0.25	97 729	404 003
Euro	1.80	1.0	1.91	1.06	5.69	72.66	97	1	0.15	34 970	439 849
Gold	2.22	1.0	3.04	1.37	8.17	271.93	317	1	0.19	83 348	617 372

We see that the different assets exhibit different features. Btc-e for instance has a total of 404 003 observations which corresponds to about 13 032 trades per day on average whereas Euro has a total of 439 850 which equals 19 124 trades per day on average. Since Bitcoin exchanges are open on weekends, Bitstamp and Btc-e have a total of 31 trading days in our estimation sample, while Euro and Gold have 23 trading days. This and the fact that we consider different trading hours for each asset as described earlier, make the descriptive stats hard to compare directly. R_1 is the autocorrelation coefficient of order 1. The Ljung-Box test statistic $Q(15)$ for 15 lags test the null hypothesis that the first 15 autocorrelations are zero. The

associated 95% critical value for this test is 24.99, which is much lower than the test statistics in Table 2, indicating presence of autocorrelation.

Even though we notice quite different descriptive statistics across our assets, they all have in common that durations are overdispersed – the standard deviation is greater than the mean. The dispersion ratio (D) is reported in the table. The majority of papers working with ACD models also reports this phenomena, see for instance Veredas et al. (2001), Bauwens (2006) and Dionne et al. (2009). This suggests that the exponential distribution is not appropriate for the distribution of trade durations. Further we note that all assets exhibit strong autocorrelation in durations.

Table 3: Descriptive stats, returns.

	Mean	Median	Std.dev	Skew	Kurtosis	Max	Min	R_1	$Q(15)$	$Q_2(15)$
Bitstamp	4.49E-07	0.0	0.0016	-0.01	22.41	0.036	-0.042	-0.27	15 726	42 738
Btc-e	1-18E-07	0.0	0.0014	-0.09	13.98	0.022	-0.033	-0.36	53 348	129 703
Euro	-3.33E-08	0.0	1.87E-05	4.31	500.6	0.002	-0.001	0.34	52 580	29 314
Gold	1.80E-08	0.0	4.41E-05	-39.56	10262	0.003	-0.012	0.37	94 043	64 305

Table 3 report the descriptive stats for estimation sample returns. We use transaction prices to calculate log returns, since bid-ask quotes were not available. We note the high autocorrelation the returns seem to exhibit, as pictured by the autocorrelation coefficients and the $Q(15)$ test statistic. The test statistic $Q_2(15)$ also strongly suggests that the volatility in the different return series are time varying. We also notice the difference in skewness and kurtosis across the assets, with Gold and Euro having the strongest evidence of a leptokurtic distribution. Btc-e on the other hand has quite much less kurtosis, however it still has excess kurtosis compared to the normal distribution. Regarding the skewness we can see some differences – while for Bitstamp and Btc-e it is close to zero, Gold has a highly negative skewness and Euro has a positive one. Also noting (and it comes as no surprise) is that both Bitstamp and Btc-e have a much higher standard deviation than Euro and Gold, with Bitstamp being the highest and Euro the lowest.

Previous works in the literature have reported a large proportion of zero returns. Again turning to Engle and Russell (1998), who reports that about two-thirds of observations were without price change. Gorski et al. (2002) and Dionne et al. (2009) report similar proportion of zero returns. For our data we find the highest proportion of zero returns in Bitstamp and Btc-e with 18% and 17% respectively. While other studies have removed zero returns from the sample, see for instance Darolles et al. (2000), we keep them in ours as the proportion is relatively low and zero returns also yield information according to the efficient market hypothesis[28].

---

[28]Price changes are due to new information arriving on the market, and in the absent of new

4.5   Seasonal data adjustment

It has been widely reported in the literature that high-frequency transaction data exhibits strong seasonality within a trading day. Typically one would observe higher trading activity at the beginning and at the end of the trading day[29]. Naturally, there will be less activity around lunchtime. Ignoring such intraday patterns could potentially distort any estimation results. The most common procedure is the one originally employed by Engle and Russell (1998). The procedure consists in decomposing the data into a deterministic part based on the time of day, and a stochastic part modeling the dynamics in the data. For the durations this give:

$$x_i = \tilde{x}_i s(t_{i-1}) \tag{4.1}$$

Where $\tilde{x}_i$ denotes what is referred to as "diurnally adjusted" durations in Engle and Russell (1998). $s(t_i)$ denotes the seasonal factor at time t_i. Eq. (3.3) for the conditional durations now become:

$$\psi_i = E(\tilde{x}_i | \mathscr{F}_{i-1}) s(t_{i-1}) \tag{4.2}$$

Optimally one would estimate the parameters both in Eq. (4.1) and (4.2) jointly by the maximum likelihood estimator in Engle and Russell (1998), however typically it is more common to apply a two-step procedure due to the difficulty of convergence. So first the raw durations are adjusted, then the ACD model(s) are estimated on "diurnally adjusted" durations, $\tilde{x}_i$.

Similarly returns are decomposed in the same manner, and we refer to them as "diurnally adjusted" returns.

$$r_i = \tilde{r}_i s(t_{i-1}) \tag{4.3}$$

Anatolyev and Shakin (2007) and Dionne et al. (2009) take the adjustment step even further by accounting for day-of-week patterns. In Anatolyev and Shakin (2007) it is argued that ignoring interday variations in the data will make it more difficult to estimate any viable model. We follow this procedure, that is we take into account both interday and intraday seasonality. Further explained by Engle

---

information or the new information is insufficient to move the price, traders will continue to trade at the current price.

[29]The reasoning behind this is that many traders will participate in the opening to take advantage from overnight news, and similarly want to close their position before the end of the day in order to avoid being exposed to news published outside their trading hours.

and Russell (1998) and Giot (2005), if seasonality and intraday variations are features in the dataset, estimating an ACD model on the original data would distort the results. First we remove interday seasonality by using a multiplicative form of adjustment as both durations and squared returns are always positive:

$$x_{i,\text{inter}} = \frac{x_i}{\overline{x_s}} \tag{4.4}$$

$$r_{i,\text{inter}} = \frac{r_i}{\sqrt{\overline{r_s^2}}} \tag{4.5}$$

Where $\overline{x_s}$ is the average duration for weekday s and $\overline{r_s^2}$ corresponds to the average of squared returns for weekday s given that observation i belongs to weekday s.

Table 9 reports the average duration and squared return across all weekdays. For Bitcoin, durations tend to be longer on weekends and in the middle of the week. A possible explanation lies in how Bitcoin exchanges work. First of all, since Bitcoin trading is still closely tied to the existing financial infrastructure, as discussed in Section 2.2, no new fiat currency will enter the exchanges on weekends. Since the bank transfer process can be lengthy, money transferred in the beginning of the week will not enter before the end of the week, and similarly money transferred before the weekend will not enter the exchange before the start of next week. The average total volume across weekdays, reported in Table 10, corresponds well to the interday duration pattern. Squared returns is also higher on the days with lower durations and higher volume. Euro and Gold tend to have shorter durations and higher squared returns closer to the end of the week, although the differences are a lot smaller than for Bitcoin.

In the literature several approaches are used for the specification of the seasonal factor s(t). Engle and Russell (1998) propose using piecewise linear or cubic splines. Zhang et al. (2001) estimate regressions based on kernels, while Ghysels et al. (2004) uses linear regressions. Other specifications include the autoregressive system with diurnal dummy variables of Dufour and Engle (2000) and the local linear regressions of Anatolyev and Shakin (2007). Here we use an approach proposed by Bauwens and Giot (2000). The interday durations, $x_{i,\text{inter}}$, are divided by the expected interday duration depending on the time of day. The expected interday duration is found by averaging the interday durations over thirty-minute intervals, and fitting a function using cubic splines so that the expected interday duration can be found given any time of day. The same procedure is done for returns. Further we estimate a cubic spline function for each day of the week, assuming

that intraday patterns can vary across different days. The computation of "diurnally adjusted" durations and "diurnally adjusted" returns is therefore done the following way:

$$x_{i,\text{intra}} = \frac{x_{i,\text{inter}}}{E(x_{i,\text{inter}}|\mathscr{F}_{t-1})} \tag{4.6}$$

$$r_{i,\text{intra}} = \frac{r_{i,\text{inter}}}{\sqrt{E(r_{i,\text{inter}}^2|\mathscr{F}_{t-1})}} \tag{4.7}$$

This is a practical approach, because it allows us to reintroduce seasonality in the simulated returns to get a viable economical interpretation of the results. However one need to careful not to capture seasonal effects that are there due to a few extreme observations rather than a consistent pattern. Therefore we tried to smooth the spline functions as much as possible. The log-ACD-ARMA-EGARCH model is applied on the deseasonalized data, but the IVaR is later computed on the simulated data after reintroducing seasonality.

Descriptive statistics for the diurnally adjusted estimation samples are reported in Table 4 and Table 5. Since we do not normalize the spline function the scale of the adjusted data is quite different from the raw data. Even though the scale is different, the time series properties are not much affected, for instance the strong autocorrelation in both returns and durations is still present. However it is noteworthy that the extreme excess kurtosis present in raw Euro and Gold returns are now present at a much lesser degree. We also note that the Ljung-Box(15) test for the squared returns is above the critical value for all assets, which indicates presence of heteroscedasticity in the returns. Figure 9 illustrate this quite clearly. The same also holds true for durations as illustrated by Figure 10, and we note that Bitstamp and Btc-e generally have higher autocorrelation coefficients than Euro and Gold.

Table 4: Descriptive stats, diurnally adjusted durations.

	Mean	Median	Std.dev	D	Skew	Kurtosis	Max	Min	R_1	$Q(15)$
Bitstamp	1.03	0.47	1.69	1,64	5.16	53.88	48.97	0.02	0.24	104 787
Btc-e	0.88	0.36	1.43	1.63	4.92	45.33	40.46	0.07	0.25	230 362
Euro	0.90	0.59	0.89	0.99	5.32	61.21	38.31	0.32	0.13	48 184
Gold	0.96	0.64	1.13	1.18	5.87	71.28	41.99	0.25	0.15	94 766

It is clear from Table 5 that given the excess kurtosis, the distribution of the returns does not resemble the normal distribution perfectly. However, by assuming conditionally normal returns we still obtain satisfactory results when backtesting

Table 5: Descriptive stats, diurnally adjusted returns.

	Mean	Median	Std.dev	Skew	Kurtosis	Max	Min	R_1	$Q(15)$	$Q_2(15)$
Bitstamp	0.0003	0.0	0.86	-0.06	19.08	17.39	-18.67	-0.27	15 728	41 743
Btc-e	0.0001	0.0	0.86	-0.08	13.20	11.34	-20.85	-0.36	53 407	124 461
Euro	-0.0026	0.0	1.03	-0.01	12.01	19.83	-18.31	0.33	48 321	93 069
Gold	0.0005	0.0	1.02	0.15	23.50	32.55	-22.48	0.34	72 732	61 537

the model, especially for Bitstamp and Btc-e, indicating similarities in the lower tail of the distribution[30]. Also, by assuming conditional normality, it is well known that some degree of unconditional excess kurtosis is generated (Bollerslev et al., 1992).

5 Model

5.1 Intraday Value at Risk

Several studies have indicated that in order to model the price process of an asset correctly, one might need more information than what are generated by the price innovations alone, see Diamond and Verrecchia (1987), Admati and Pfleiderer (1988) and Easley and O'Hara (1992) among others. One of the theories is that the time between trades influences volatility, however the result of this influence is not agreed upon. While Easley and O'Hara associates high trading activity and thus low duration with low volatility, the study done by Admati and Pfleiderer finds the opposite. What they do agree upon though is that duration affects volatility. By including this in the volatility specification one can therefore gain insights otherwise not found.

Therefore there seems to be an advantage in specifically including the trade duration in the volatility specification. By making this joint model of durations and returns one are better than before able to distinguish between the effect duration and return have on the risk measure. Dionne et al. (2009) also finds that by ignoring the effect of trade durations the estimated risk level may be underestimated. In order to capture this relationship, our approach will therefore follow that of Dionne et al. closely.

Consider the trading process defined by the price paid in trade i at time t, p_t, and the time between trade t_i and t_{i-1}, defined as the duration x_i. The corresponding

---

[30]We have also tested with assuming a conditionally Student t specification, however we did not get any better results. Similar results can be seen Dionne et al. (2009)'s analysis.

return r_i is now defined as the continuously compounded return: $r_i = log(\frac{p_i}{p_{i-1}})$. A path is defined as the trading process from $i = 1, ...I$, where I is the last trade in the period. This path now consists of irregular spaced returns, so called tick-by-tick data.

A regular interval trading process is defined as always having the same duration between observations, which means that for a regular path $t_i - t_{i-1} = T$. In order to convert a path with irregular spaced observations into one with regular intervals, we use the following procedure. Starting with the first regular interval, the irregular spaced observations that belong to this interval are the n first observations, such that the sum of the those durations do not exceed the length of the regular interval, and if you add the next duration, the sum will exceed the interval length. The same will be true for the rest of the intervals: The first duration that belongs to each interval is the first duration that was not included in the previous intervals, and the last duration will be the last duration in the path that can be included without pushing the cumulative duration for that interval past the interval length. If we denote the first observation in the regular interval k as $\tau(k)$, and denote the number of observations in this interval before the cumulative duration is bigger than the interval length as $n(k)$, the rest of the observations can be put into regular intervals by setting

$$\sum_{i=\tau(k)}^{n(k)} \leq T \quad \text{and} \quad \sum_{i=\tau(k)}^{n(k)+1} > T \tag{5.1}$$

Since each duration has a corresponding return, and we are dealing with log returns, the return y_k for for each regular interval will be

$$y_k = \sum_{i=\tau(k)}^{n(k)} r_i \tag{5.2}$$

Now consider an IVaR model modeled as a time series, where each interval k has one VaR-level. For a given confidence level $1 - \alpha$, the IVaR is formally defined as

$$Pr(y_k < -\text{IVaR}_k(\alpha)|\mathscr{G}_k) = \alpha \tag{5.3}$$

where $\mathscr{G}_k$ denotes the information set that includes the transaction history (durations and returns) up to time t_{k-1}. The IVaR level will now be the corresponding α-quantile of the conditional distribution of y_k:

$$- \text{IVaR}_k = Q_k(\alpha | \mathscr{G}_k) \tag{5.4}$$

To obtain the conditional distribution of y_k, we chose a Monte Carlo simulation approach. In short this consists of making models for both the duration and returns, and simulate possible paths that the price can take. By doing this a sufficiently number of times, we can get the IVaR level for each regular interval by following the previously discussed procedure.

5.2   The log-ACD-ARMA-EGARCH model

Remember from Section 3.5 that durations and their associated marks have a joint density as stated by Eq. (3.1). In our framework the marks of interest are the log returns from one time of arrival to the next. We therefore let the joint density in Eq. (3.1) be replaced by:

$$(x_i, r_i) | \mathscr{F}_{t-1} \sim f(x_i, r_i | \breve{x}_{i-1}, \breve{r}_{i-1}; \Theta_f) \tag{5.5}$$

Following the expression of Eq. (3.2) the log-likelihood function for a sample of observations x_i, r_i with $i = 1, ...T$ is then:

$$\mathscr{L}(\Theta_x, \Theta_r) = \sum_{i=1}^{T} \left[\log g(x_i | \breve{x}_{i-1}, \breve{r}_{i-1}; \Theta_x) + \log q(r_i | x_i, \breve{x}_{i-1}, \breve{r}_{i-1}; \Theta_r) \right] \tag{5.6}$$

Several possibilities for g and q can be chosen. In the next subsections we follow the approach of Dionne et al. (2009) and make use of the log-ACD model for the durations x and a general ARMA-EGARCH setup to model the returns r.

5.3   Log-ACD

From its most general form, several variations of the ACD model can be formulated, by different specifications for the conditional durations and the density $p(\epsilon)$. As originally proposed by Bauwens and Giot (2000) we use the log-ACD model to ensure non-negative durations without the necessity of imposing non-negativity constraints. This makes the log-ACD(m,q) model, where the autoregressive equation is specified on the logarithm of the conditional duration ψ_i, a quite flexible

approach. Further we make use of the generalized gamma distribution for the density of standardized durations, $p(\epsilon)$. This allows for a humpshaped hazard rate. We refer to this model as the log-GACD(m,q) model.

In our log-GACD(m,q) model, the expected conditional duration is given by the equation:

$$\psi_i = \exp\left(\omega + \sum_{j=1}^{m}\alpha_j\left(\frac{x_{i-j}}{\psi_{i-j}}\right) + \sum_{j=1}^{q}\beta_j\log(\psi_{i-j})\right) \tag{5.7}$$

Following the literature in Section 3.5, we also make the assumption that all the temporal dependence in the durations are captured by the expected conditional durations, as given by:

$$\frac{x_i}{\psi_i} = \epsilon_i \tag{5.8}$$

The generalized gamma distribution for the standardized durations ϵ is given by:

$$p(\epsilon|h,k) = \frac{h}{\Gamma(k)s}\left(\frac{\epsilon}{s}\right)^{kh-1}\exp\left(-\left(\frac{\epsilon}{s}\right)^h\right) \tag{5.9}$$

Where Γ denotes the gamma function[31], h and k denote scale parameters and

$$s = \frac{\Gamma(k)}{\Gamma(k + \frac{1}{h})} \tag{5.10}$$

We note that if $k = 1$, the gamma distribution in Eq. (5.9) reduces to the distribution used in the Weibull ACD model, and further if $h = 1$ it reduces to the one used in the Exponential ACD model. The value of the parameters also affects the hazard rate, $\lambda(t)$, as derived in Glaser (1980):

---

[31]Such that $\Gamma(k) = \int_0^\infty s^{k-1}\exp(-s)ds$.

$hk - 1 < 0 :$

$$h \leq 1 \Rightarrow \lambda(t) \qquad\qquad\qquad\qquad\qquad \text{decreasing}$$
$$h > 1 \Rightarrow \lambda(t) \qquad\qquad\qquad\qquad\qquad \cup\text{-shaped}$$

$hk - 1 > 0 :$

$$h \geq 1 \Rightarrow \lambda(t) \qquad\qquad\qquad\qquad\qquad \text{increasing}$$
$$h < 1 \Rightarrow \lambda(t) \qquad\qquad\qquad\qquad \text{inverted } \cup\text{-shaped}$$

$hk - 1 = 0 :$

$$h = 1 \Rightarrow k = 1 \Rightarrow \lambda(t) \qquad\qquad \text{constant, exponential density}$$
$$h < 1 \Rightarrow \lambda(t) \qquad\qquad\qquad\qquad\qquad \text{decreasing}$$
$$h > 1 \Rightarrow \lambda(t) \qquad\qquad\qquad\qquad\qquad \text{increasing}$$

Further we let:

$$\xi_i = \psi_i s = \psi_i \frac{\Gamma(k)}{\Gamma(k + \frac{1}{h})} \tag{5.11}$$

Finally the conditional log-likelihood of the log-GACD(1,1) model on a set of observed durations $x = (x_1, ... x_n)$ is given by:

$$\mathscr{L}(\Theta_x | x) = \sum_{i=2}^{N} \left(\log(h) + (hk - 1)\log\left(\frac{x_i}{\xi_i}\right) - \log(\Gamma(k)\xi_i) - \left(\frac{x_i}{\xi_i}\right)^h \right) \tag{5.12}$$

5.4   ARMA - EGARCH

Given that we use transaction prices, the literature strongly suggest that because of market microstructure effects such as nonsynchronous trading and bid-ask bounce, induced serial correlation may be present in the return series even if it is independent and serial uncorrelated, see Fisher (1966), Roll (1984) and Lo and MacKinlay (1990). We also remember from Section 4 that we do indeed see this serial correlation for all investigated assets. In order to capture these effects, we follow the approach of Ghysels and Jasiak (1998), Grammig and Wellner (2002) and Dionne et al. (2009), and use an ARMA(p,q) model specification for the return series.

The ARMA(p,q) model combines an autoregressive representation with a moving average process for the errors:

$$r_i = c + e_i + \sum_{j=1}^{p} \phi_j r_{i-j} + \sum_{j=1}^{q} \theta_j e_{i-j} \qquad (5.13)$$

The error term e_i from Eq. (5.13) satisfies

$$e_i = z_i \sigma$$

where $z = \{z_i, i \in \mathbb{Z}\}$ denotes a strong white noise process. The process z is thus composed of i.i.d. random variables, with zero mean and finite variance. Since z_i is a standard normal variable, the expectation of z_i is given by $E(|z_i|) = \sqrt{2/\pi}$. The conditional variance for transaction i is denoted σ_i, i.e $\sigma_i = V_{i-1}(r_i|x_1, \breve{x}_{i-1}, \breve{r}_{i-1})$.

To model the time-varying volatility in the returns found in Section 4.5, we use a modified Exponential GARCH(P,Q) model. The EGARCH formulates the conditional variance equation in terms of the log of the variance, and thus always imposes a positive conditional variance. This removes the need to put constraints on the coefficients in the model estimation, and will always output positive conditional variance during the later simulation, which gives a big advantage regarding computational simplicity. Our EGARCH model is given by:

$$\log(\sigma_i^2) = \gamma \log(x_i) + \kappa + \sum_{j=1}^{P} \tilde{\beta}_j \left\{ \log(\sigma_{i-j}^2) - \gamma \log(x_{i-j}) \right\}$$

$$+ \sum_{j=1}^{Q} [\tilde{\alpha}_j \left\{ |z_{i-j}| - E(|z_{i-j}|) \right\} + \tilde{\xi}_j z_{i-j}] \qquad (5.14)$$

In order to capture the effect the irregular time between trades has on the volatility, we do as Dionne et al. (2009) and add a duration term to our EGARCH equation. The parameter γ specifies the duration weighting for the volatility of an asset. We note that when $\gamma = 0$, which is when the irregular spacing of returns have no effect on volatility, Eq. (5.14) denotes a standard EGARCH model, as first proposed by Nelson (1991). Note that the conditional variance is, as the returns and durations, specified per transaction, i.e. not as a function of time.

With this specification for the normal EGARCH-model, and given the observed duration series $x = (x_1, ..., x_n)$ and return series $r = (r_1, ..., r_n)$, the log-likelihood function is given by:

$$\mathscr{L}(\Theta_r | x, r) = -\frac{1}{2} \sum_{i=1}^{N} \left(\log(\sigma_i^2) + \left(\frac{e_i}{\sigma_i} \right)^2 \right) \tag{5.15}$$

6 Estimation of model parameters

After having specified the log-ACD-ARMA-EGARCH model in Section 5, we used Matlab 8 with the optimization toolbox to estimate the model parameters for each asset. For simplicity and runtime of the estimation, we estimated the model parameters separately for the log-ACD model, the ARMA model and the EGARCH model.

In order to estimate the parameters of the log-ACD model, we maximized the log-likelihood function in Eq. (5.12). We then estimated the parameters of the ARMA model, by using an iterative Least Sum of Squares approach. The e-vector from this estimation was then used in the estimation of the parameters of the EGARCH model, which was done by again maximizing the log-likelihood, this time of Eq. (5.15). This procedure was done for all assets.

There was one drawback with this approach though, and that was that the unconstrained Matlab solver fminsearch, which we used, tends to get stuck in local optima, i.e. it might converge to solutions that are not the global optimum. To try to account for this problem, we utilized a simple multi-start approach, where we ran the optimization algorithm numerous times for different random start values. To calculate the standard errors we applied a numerical approximation approach inspired by Perlin (2012), and given the presence of heteroscedasticity presented in Section 4, we are using Huber-White standard errors. The results of this estimation is presented in Table 6. The Ljung-Box(15) test done on the standardized residuals is presented at the bottom at Table 6. Non-significant parameters are not shown.

When determining the order of the model for each asset, we tried to find a compromise between the principles of Occam's Razor[32] and model performance. To measure model performance we follow Hautsch (2002), Dionne et al. (2009) and Wongsaart et al. (2009) among others, and use the result of the Ljung-Box(15) test statistic on the standardized residuals from the estimation as performance measurement. We tried models of both higher and lower order than the ones presented in Table 6, and kept the one with the lowest order which performed acceptably.

---

[32]As few model parameters as possible.

It is interesting to note that incorporating two lags seems to be the best compromise for adequate performance in most models, the only exception being an ARMA(3,3) for the mean equation of Gold and an EGARCH(1,1) for the conditional variance of Bitstamp. Also worth mentioning is that the parameter c from Eq. (5.13) is non-significant for all assets except Euro. When a parameter was deemed insignificant, we ran the model estimation again without the parameter in question.

If we take a look at the log-ACD parameters, we see that $\beta_1 + \beta_2$ is close to 1 for all assets, indicating a high persistence in durations. All parameters are significant for all assets. The Ljung-Box test statistic for 15 lags is still quite far above the critical value at 24.99, which indicates that there is still autocorrelation present in the data. However, it is an improvement in the order of 100 from the result in Table 4. For instance, the asset with the least improvement is Euro, where the test statistic is reduced from 49 184 to 342.

From the value of the gamma parameters, h and k, we can derive the hazard rate, $\lambda(t)$ as described in Section 5.3. For all assets $hk - 1 < 0$ and $h < 1$, which corresponds to decreasing $\lambda(t)$. In other words, it holds true for all assets that the longer the last trade dates back, the lower the probability for observing the next trade. In Gerhard and Hautsch (2000) a decreasing hazard rate is explained by market participants tending to trade in "a technical trading scheme". Not only the amount of information, but also the speed of information arrival seem to be relevant. This suggest that the market is dominated by traders that tend to learn from past sequences of information in the market, and not only adjust their positions based on exogenous criteria[33].

Figure 11 plots the kernel density estimates[34] of the probability density function for the diurnally adjusted durations. We take notice of the similarity in the distribution for Bitstamp and Btc-e, and it is therefore natural that the estimated scale parameters in the generalized gamma distribution, h and k, are estimated to be alike for the two Bitcoin exchanges. Figure 11 also illustrates the overdispersion, positive skewness and excess kurtosis found in all assets.

The same tendencies as for the log-ACD estimation can be seen in the Ljung-Box(15) test statistic for the residuals from the ARMA and EGARCH estimation. A reduction in the order of 100 is seen for all assets, but the only model to actually pass the adequacy test is the ARMA(2,2) for Euro. Engle (2000) finds similar results and as Dionne et al. (2009) states: "Passing standard tests of model ade-

---

[33]In Easley and O'Hara (1992) no-trade intervals is associated with lack of information, and this result suggest that this hypothesis hold true.

[34]The gaussian kernel is used.

quacy seems to remain an issue when using irregular high-frequency data due to the extremely high number of observations". Keeping this argument in mind[35], together with the fact that the model actually decreases the autocorrelation in the data dramatically and that we mainly want to use it for illustrative risk purposes, we count the test results as satisfactory.

Looking at the parameters for the ARMA models, we note that the sum of the moving average coefficients, $\sum \phi_i$, is positive for all assets. The sum of autoregressive coefficients, $\sum \theta_i$, is negative. According to Lee (2008) this can be explained by the fact that informed traders tend to split large orders into smaller ones, in order to camouflage their moves (not share their information) and avoid making the price move too much. This will typically make the prices move in the same direction, and thus induce serial correlation in the returns.

As for the EGARCH parameters, we note that the sum of beta coefficients, $\sum \tilde{\beta}_1$, is close to 1 for all assets, indicating high persistence in the conditional variance. We can also see that the duration term γ in the volatility equation is statistically significant for all assets. Even though the size of γ can be quite small – Btc-e for instance has a γ coefficient of 0.0345, this still indicates that our model gets a higher information content by including durations in the volatility specification. It is also worth mentioning that for Euro the α-parameters were insignificant, which means there is evidence for a leverage effect[36] in all assets except Euro.

---

[35]We actually have more than 3 times as many observations as Dionne et al. and Engle.

[36]The impact on volatility is different for positive and negative returns of the same absolute value.

Table 6: This table contains the estimated parameters for all assets. Estimates with t-stats lower than the critical value are omitted. $Q_{acd}(15), Q_{arma}(15)$ and Q_{egarch} are the test stats from the Ljung-Box test with 15 lags for the residuals of the ACD, ARMA and EGARCH parameter estimation respectively.

Parameters	Bitstamp		Btc-e		Euro		Gold	
	Estimate	t-stats	Estimate	t-stats	Estimate	t-stats	Estimate	t-stats
ω	-0.0053	-4.8090			-0.0248	-44.5865	-0.0313	-35.0408
α_1	0.1103	42.0552	0.1112	7.2033	0.0889	61.0692	0.0722	64.0140
α_2	-0.1053	-52.2752	-0.1047	-16.3708	-0.0665	-47.5844	-0.0431	-31.9526
β_1	1.8515	108.0023	1.8476	15.9128	1.5626	187.8759	1.4776	110.8217
β_2	-0.8528	-50.6045	-0.8494	-7.5028	-0.5768	-70.6624	-0.4957	-38.5685
h	0.2940	84.1581	0.3299	135.3063	0.6453	506.7600	0.5853	178.0794
k	9.0647	45.4737	9.2992	69.7297	6.0825	451.4829	6.0478	98.8628
c					-0.0020	-2.4125		
ϕ_1	0.8240	17.3520	0.3545	16.7106	0.2306	4.8327	0.8069	43.6643
ϕ_2	-0.1103	-4.9778	-0.0015	-1.9143	0.0201	3.4986	0.1707	16.1475
ϕ_3							-0.0257	-6.4858
θ_1	-1.1768	-24.9906	-0.8630	-39.8790	0.1239	2.5963	-0.4259	-23.4230
θ_2	0.2930	8.0914	0.0814	5.7669	-0.0616	-5.3248	-0.4514	-63.5156
θ_3							-0.0623	-8.7809
κ	-0.0335	-83.6962	-0.0001	-6.294	0.0020	55.4802	-0.0021	-59.2836
$\tilde{\alpha}_1$	0.0081	11.6400	0.00320	3.6541			-0.0162	-24.2356
$\tilde{\alpha}_2$			-0.0039	-4.5187			0.0183	26.1227
$\tilde{\xi}_1$	0.3787	494.5530	0.3858	355.4009	0.3860	212.1686	0.3962	413.1224
$\tilde{\xi}_2$			-0.3670	-356.4270	-0.3475	-196.9791	-0.3373	-380.3510
$\tilde{\beta}_1$	0.9091	6630.1650	1.7547	1808.9800	1.3778	321.9991	1.3877	602.7121
$\tilde{\beta}_2$			-0.7578	-794.1990	-0.3819	-90.2397	-0.4019	-181.1820
γ	-0.1052	-99.5225	0.0345	38.1001	0.0970	4.4208	0.0775	78.3477
$Q_{\text{acd}}(15)$	334		342		96		104	
$Q_{\text{arma}}(15)$	47		25		12		28	
$Q_{\text{egarch}}(15)$	25		287		220		99	

7 Simulation

In this section we describe the strategy used for simulating returns based on our log-ACD-ARMA-EGARCH model specified in Section 5.

First we use our log-ACD model to simulate the time between trades for each asset. This will then be input to the ARMA-EGARCH model, which simulates the return series for the corresponding asset.

The detailed approach for the whole estimation and simulation process is as follows:

1. We separate the sample period into two parts – one estimation period and one validation period.

2. By using the estimation sample, we estimate the parameters for the log-ACD-ARMA-EGARCH model described in Section 5.

3. Now we are ready to simulate. First we generate random numbers from the standard normal distribution and the generalized gamma distribution[37], for the returns and durations respectively. These will serve as the innovations in the simulated paths.

4. By using the last observations in the estimation period as starting values, we use our log-ACD model to simulate the duration series. This is done in an iterative manner until the number of transactions in the validation period has been reached.

5. We have now simulated the transaction timeline. By employing the specification for our ARMA-EGARCH model we are able to simulate the corresponding returns to all the transactions. The combination of the duration and return time series will be known as one path.

6. We now repeat step 3-5 5000 times in order to obtain 5000 possible paths for each asset.

7. These paths now contain irregular spaced trades. In order to transform them into regular intervals we use the strategy described in Section 5.1. We do this for several different interval lengths.

8. With 5000 possible values for each regular interval, we are able to use the specification of the IVaR model in Section 5.1 to obtain Value at Risk time series for different α-quantiles.

9. In order to test that our model specification correctly models the data, we employ the backtest methodology described in Section 7.1.

10. To get comparable and understandable results, we reintroduce seasonality to the simulated returns. We then repeat step 8 in order to get seasonalized Value at Risk time series for different α-quantiles.

7.1   Backtesting

In order to assess the validity of our model, we follow the backtesting methodology outlined by Alexander (2008), and then apply the coverage tests proposed by

---

[37]In order to draw the innovations from the generalized gamma distribution, we used the WAFO toolbox for Matlab. However, this was a rather slow approach, and if we were to do the same procedure again, we would have invested some time in finding a procedure that requires less computational power.

Kupiec (1995) and Christoffersen (1998). We do both tests for several different interval lengths and confidence levels.

In order for a VaR model to be considered accurate and precise, the number of exceedances[38] should be close to the value estimated by the α-quantile we are trying to model. The exceedances should also be randomly distributed over the sample - if not there will be clustering (periods when exceedances happen more often than average), and our model is bound to overestimate in some periods and underestimate in other periods.

Our backtesting methodology is as follows: We first convert the irregular return series of the validation period into a regular interval time series by the method described in Section 5.1. We then employ a rolling window strategy, where we use all the data available at the end of each window to estimate the VaR for the next window. Each window will here correspond to a regular interval in the validation return series. This VaR estimate is then compared to the actual validation period return for this interval in the manner outlined in the next subsections. The window is then rolled forward, using the information gained in this window to estimate the VaR for the next. The procedure is repeated until a VaR estimate for all the regular intervals in the validation period is calculated. Now one can test how good the specification of the VaR model actually is.

7.1.1   The unconditional test of Kupiec

The Kupiec test checks whether the number of exceedances are equal to the pre-defined VaR level. To do this, a likelihood ratio test is applied to reveal whether the model provides the correct unconditional coverage.

Let H_t be an indicator sequence where H_t takes the value 1 if the observed return, Y_t, is below the predicted VaR quantile, Q_t at time t:

$$H_t = \begin{cases} 1 & \text{if } Y_t \ < \ Q_t \\ 0 & \text{if } Y_t \ \geq \ Q_t \end{cases} \tag{7.1}$$

The null hypothesis is that the unconditional coverage is correct. The test statistic will then be

---

[38]An exceedance is here an observations from the validation sample that are less the estimated VaR.

$$-2ln(LR_{\mathrm{uc}}) = -2[n_0\log(1-\pi_{\mathrm{exp}}) + n_1\log(\pi_{\mathrm{exp}})$$
$$-n_0\log(1-\pi_{\mathrm{obs}}) - n_1\log(\pi_{\mathrm{obs}})] \sim \chi_1^2 \qquad (7.2)$$

Here n_1 and n_0 are the number violations and non-violations respectively, π_{exp} is the expected proportion of exceedances and $\pi_{\mathrm{obs}} = \frac{n_1}{n_0+n_1}$ is the observed proportion of exceedances.

However, in this test it is only the total number of exceedances that counts. How the exceedances are distributed over the sample is not tested. The Kupiec test will thus not provide information about whether there is clustering among these exceedances.

7.1.2   The conditional coverage test of Christoffersen

In order improve the Kupiec test, Christoffersen developed a joint test, both to provide correct coverage, but also to detect whether the exceedances are randomly distributed or happen in clusters. The Christoffersen test statistic is defined as follows:

$$-2\log(LR_{cc}) = -2[n_0\log(1-\pi_{\mathrm{exp}}) + n_1\log(\pi_{\mathrm{exp}}) - n_{00}\log(1-\pi_{01})$$
$$-n_{01}\log(1-\pi_{01}) - n_{10}\log(1-\pi_{11}) - n_{11}\log(\pi_{11})] \sim \chi_2^2 \quad (7.3)$$

n_{ij} is here defined as the number of times an observation with value i is followed by an observation with value j. Further $\pi_{01} = \frac{n_{01}}{n_{00}+n_{01}}$ and $\pi_{11} = \frac{n_{11}}{n_{11}+n_{10}}$.

We note that the Christoffersen test only measures the number of times an exceedance is followed by another exceedance, and that all other relationships between these are thus not captured by the test.

7.2   Backtesting results

The results of the backtest are presented in Table 7. Bold entries denote a failure of the test at the 1% level. First of all we note that the backtests are done on deseasonalized data. The model is estimated at the deseasonalized level, so this seemed like the natural choice. This means that the intervals here do not

denote calendar time units, so they are not directly comparable[39]. However, an approximation can be done quite easily by comparing the number of intervals and interval length to the length of the validation period. Given that we want to perform a comparison of calendar time intervals in the next section, we have chosen the intervals such that the shortest corresponding calendar time is approximately 2 minutes, and the longest interval corresponds to around 2 hours. In order to better perceive how we capture the whole tail distribution, we have chosen 4 different α-quantiles to perform the tests on: 0.5%, 1%, 5% and 10%.

There seems to be a pattern here: The models perform better on longer intervals and on smaller quantiles, especially for the conditional test. The fact that they perform better on longer horizons is quite natural – with more observations in each interval the distribution will be more stable, less determined by chance and thus easier to predict. Dionne et al. (2009) has similar findings. In order to explain the second point, remember from Section 4.5 that the distribution did not seem to fully resemble the conditionally normal distribution. This is probably what is reflected in the test statistics also – the tails of the distribution can be assumed conditionally normal with satisfactory results, but as we get closer to the middle of the distribution, this assumption grows weaker, depending on the asset. Bitstamp for instance performs well even for the 10% quantile. We include this finding for completeness of the analysis, and in order to indicate that one might be careful to use the model with these assumptions on quantiles bigger than 5%.

Overall we find that the models perform well for most quantiles and interval lengths except the ones mentioned above. We therefore conclude that the models are able to capture the characteristics of the data in a satisfactory way, although not perfect.

---

[39]While 100 as interval length in the table corresponds to approximately 23 minutes for Bitstamp, an interval length of 100 corresponds to just over 3 minutes for Euro.

Table 7: Backtesting test statistics for the unconditional and conditional coverage tests. Bold entries denote a failure of the test at 1% level. Critical values of the Kupiec and Christoffersen test at the 1% level are 6.635 and 9.210 respectively.

Interval length	Kupiec's test 0.50 %	1 %	5 %	10 %	Christoffersen's test 0.50 %	1 %	5 %	10 %	# Intervals	Approx. calendar time
Bitstamp										
10	**7.599**	1.941	1.218	**10.762**	9.096	4.886	**20.701**	**23.392**	17249	2 min 20 sec
25	**19.009**	**26.872**	0.780	0.013	**20.189**	**28.626**	**28.942**	**23.921**	6899	5 min 48 sec
50	**14.140**	**22.956**	6.135	0.123	**27.037**	**32.932**	8.193	2.575	3449	11 min 37 sec
100	1.190	2.394	2.234	0.866	1.373	3.101	4.079	4.350	1724	23 min 15 sec
200	2.541	5.114	3.761	0.000	2.710	5.758	4.476	0.248	862	46 min 30 sec
300	0.402	1.606	0.016	0.104	0.472	1.926	0.399	0.325	574	69 min 50 sec
400	0.301	0.444	1.065	1.755	0.357	0.500	2.549	3.757	431	93 min 1 sec
500	0.353	0.708	0.648	2.415	0.364	0.744	1.927	2.579	344	116 min 32 sec
Btc-e										
15	**13.952**	**12.406**	0.564	0.773	**15.367**	**14.203**	**27.418**	1.208	21196	1 min 52 sec
35	0.446	0.008	0.155	0.854	0.827	6.038	**30.508**	**29.460**	9084	4 min 22 sec
100	0.563	1.577	0.056	2.386	0.678	3.290	2.513	4.199	3179	12 min 29 sec
150	1.390	3.701	0.352	5.925	1.443	3.290	**11.735**	**9.439**	2119	18 min 43 sec
250	2.210	1.212	3.770	**14.750**	2.229	1.355	5.817	**28.527**	1271	31 min 12 sec
400	3.186	1.259	4.762	**8.389**	3.191	**6.652**	4.838	**9.611**	794	49 min 57 sec
500	0.500	2.208	5.202	6.535	0.519	2.246	7.340	6.944	635	62 min 28 sec
1000	0.246	0.495	4.912	**10.262**	0.259	0.533	6.728	**18.220**	317	125 min 7 sec
Euro										
75	**10.993**	2.323	**26.389**	**99.132**	**18.696**	**27.156**	**57.056**	**135.052**	4812	2 min 30 sec
150	**10.715**	4.415	**17.750**	**66.599**	**15.669**	7.065	**35.211**	**76.466**	2406	4 min 59 sec
300	6.121	1.847	2.297	**23.620**	8.464	7.213	4.104	**29.829**	1203	9 min 58 sec
400	2.212	1.570	4.408	**20.796**	2.373	3.425	6.806	**22.187**	902	13 min 18 sec
1000	0.676	0.104	2.339	**11.921**	0.743	0.171	3.097	**12.034**	360	33 min 20 sec
1500	0.034	1.046	0.354	5.274	0.051	1.062	1.087	5.435	240	50 min 0 sec
2000	1.013	0.687	1.152	4.537	1.081	**9.305**	8.462	**10.837**	180	66 min 40 sec
4000	0.513	0.013	4.053	5.700	0.559	0.059	4.099	5.980	90	133 min 20 sec
Gold										
50	**9.171**	4.717	**30.367**	**92.171**	**22.812**	**25.929**	**48.827**	**109.958**	10691	1 min 58 sec
150	0.921	0.523	4.567	**19.404**	1.207	1.455	**12.194**	**22.037**	3563	5 min 53 sec
300	0.001	0.190	3.647	**15.719**	0.103	0.499	4.126	**17.240**	1781	11 min 45 sec
500	0.022	0.045	4.492	**18.746**	0.079	0.253	4.695	**19.269**	1069	19 min 35 sec
750	2.583	0.706	5.301	**10.186**	2.588	0.791	5.452	**10.593**	712	29 min 24 sec
1000	0.183	2.760	6.324	**9.805**	0.205	2.783	7.252	**9.934**	534	39 min 12 sec
1500	3.559	7.136	8.812	**7.598**	3.559	7.136	9.135	7.778	356	58 min 48 sec
3000	1.774	3.558	5.410	**11.257**	1.774	3.558	5.549	**11.750**	178	117 min 36 sec

8   Empirical results and implications

8.1   Results

The backtesting results show that the estimated models perform adequately well, but to get an economical interpretation of the results we need to reintroduce seasonality in the simulated durations and returns. Remember from Section 4.5 that the initial deseasonalization depend both on the day of the week, and on the time of day. Reintroducing the seasonality is then just a matter of incorporating the seasonal factors by reversing the procedure. The time to the next transaction is predicted by the log–ACD model on a deseasonalized scale, so by reintroducing seasonality duration by duration we are able to keep track of the time of the next transaction in calendar time.

We have calculated the IVaR for several intraday time horizons, ranging from 2 minutes to 2 hours as can be seen in Table 8 and Figure 4. We considered these intervals as in our view they represent the shortest interval that would be any purpose to model, and the approximate worst case scenario for confirming a Bitcoin transaction. To compute these estimates, we follow the procedure described in Section 5.1 for the different interval lengths. The mean and standard deviation of the resulting IVaR time series are reported in Table 8, while Figure 4 presents how the mean changes given different IVaR levels in a more illustrative way.

First of all we remark that the estimated level for the Bitcoin exchanges, Bitstamp and Btc-e, is much larger than the level for Euro and Gold. This is not surprising, given the relatively high difference in the standard deviation of the returns seen in Table 3. The standard deviation for the IVaR estimates is also quite much higher for the Bitcoin exchanges, however we note that Gold has a standard deviation closer to that of Bitstamp and Btc-e than Euro. It is also interesting to note that there seem to be some differences between Bitstamp and Btc-e – Bitstamp has a slightly higher estimated mean for all IVaR levels and interval lengths. It can maybe be explained by the fact that Bitstamp was a bit more variable in the estimation sample, and is not necessary a feature that will be present in other periods.

When looking closer at the numbers, we see more clearly how much more risky Bitcoin really is compared to traditional assets. For instance, we are on average 99% sure that the Bitcoin exchange rate at Bitstamp will not lose more than 1.03% of its value compared to USD in the next 2 minutes. If it is the Euro/USD exchange rate we are worried about however, we can be 99% sure that it will not go down more than 0.37% in the next 2 *hours.* 2 hours as the interval length for

Table 8: Comparison of IVaR estimates. Interval lengths are given in seconds.

Interval length	Mean				Standard Deviation			
	0.5 %	1 %	5 %	10 %	0.5 %	1 %	5 %	10 %
Bitstamp								
120	1.24 %	1.03 %	0.60 %	0.42 %	0.06 %	0.04 %	0.02 %	0.01 %
300	1.61 %	1.36 %	0.80 %	0.57 %	0.08 %	0.05 %	0.02 %	0.02 %
600	2.06 %	1.74 %	1.04 %	0.75 %	0.09 %	0.06 %	0.03 %	0.03 %
900	2.42 %	2.06 %	1.24 %	0.90 %	0.11 %	0.08 %	0.04 %	0.03 %
1800	3.24 %	2.77 %	1.70 %	1.25 %	0.14 %	0.10 %	0.05 %	0.04 %
2700	3.89 %	3.32 %	2.06 %	1.52 %	0.17 %	0.11 %	0.06 %	0.05 %
3600	4.43 %	3.79 %	2.37 %	1.75 %	0.18 %	0.13 %	0.07 %	0.05 %
7200	6.09 %	5.26 %	3.31 %	2.46 %	0.23 %	0.17 %	0.09 %	0.07 %
Btc-e								
120	0.96 %	0.82 %	0.49 %	0.36 %	0.05 %	0.03 %	0.02 %	0.02 %
300	1.32 %	1.12 %	0.68 %	0.49 %	0.06 %	0.05 %	0.03 %	0.02 %
600	1.74 %	1.49 %	0.91 %	0.66 %	0.08 %	0.06 %	0.04 %	0.03 %
900	2.08 %	1.78 %	1.09 %	0.79 %	0.10 %	0.07 %	0.04 %	0.03 %
1800	2.84 %	2.44 %	1.51 %	1.10 %	0.12 %	0.10 %	0.06 %	0.05 %
2700	3.42 %	2.94 %	1.83 %	1.35 %	0.15 %	0.12 %	0.07 %	0.06 %
3600	3.91 %	3.38 %	2.11 %	1.55 %	0.17 %	0.14 %	0.08 %	0.07 %
7200	5.44 %	4.71 %	2.96 %	2.18 %	0.22 %	0.18 %	0.12 %	0.09 %
Euro								
120	0.06 %	0.05 %	0.03 %	0.02 %	0.01 %	0.01 %	0.00 %	0.00 %
300	0.09 %	0.08 %	0.05 %	0.04 %	0.01 %	0.01 %	0.00 %	0.00 %
600	0.12 %	0.11 %	0.07 %	0.05 %	0.02 %	0.01 %	0.01 %	0.00 %
900	0.15 %	0.13 %	0.09 %	0.06 %	0.02 %	0.02 %	0.01 %	0.01 %
1800	0.21 %	0.19 %	0.12 %	0.09 %	0.03 %	0.02 %	0.01 %	0.01 %
2700	0.26 %	0.23 %	0.15 %	0.12 %	0.03 %	0.03 %	0.01 %	0.01 %
3600	0.30 %	0.26 %	0.18 %	0.14 %	0.04 %	0.03 %	0.02 %	0.01 %
7200	0.42 %	0.37 %	0.25 %	0.20 %	0.05 %	0.04 %	0.02 %	0.02 %
Gold								
120	0.10 %	0.08 %	0.05 %	0.04 %	0.02 %	0.02 %	0.01 %	0.01 %
300	0.15 %	0.13 %	0.08 %	0.06 %	0.03 %	0.03 %	0.02 %	0.01 %
600	0.20 %	0.18 %	0.11 %	0.08 %	0.04 %	0.04 %	0.02 %	0.02 %
900	0.25 %	0.21 %	0.14 %	0.10 %	0.05 %	0.04 %	0.03 %	0.02 %
1800	0.34 %	0.30 %	0.19 %	0.14 %	0.07 %	0.06 %	0.04 %	0.03 %
2700	0.42 %	0.37 %	0.24 %	0.18 %	0.09 %	0.08 %	0.05 %	0.04 %
3600	0.48 %	0.42 %	0.27 %	0.20 %	0.10 %	0.09 %	0.06 %	0.04 %
7200	0.68 %	0.60 %	0.39 %	0.29 %	0.14 %	0.12 %	0.08 %	0.06 %

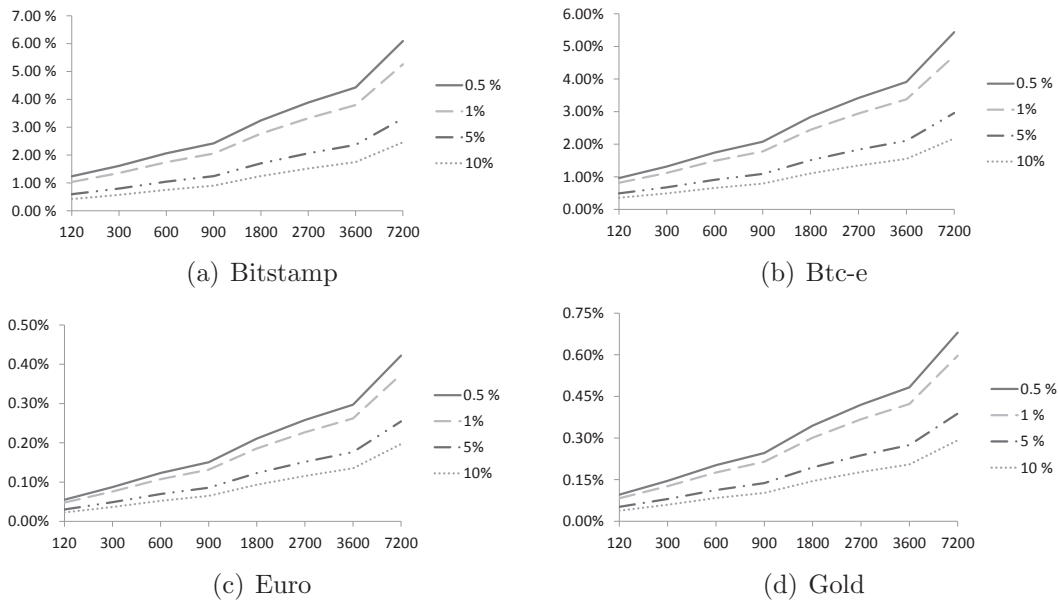(a) Bitstamp

(b) Btc-e

(c) Euro

(d) Gold

Figure 4: Graphical presentation of how the IVaR level changes with different interval lengths.

Bitstamp gives us a IVaR estimate of 5.26%, almost 15 times bigger than the Euro estimate! It is clear from this that even if you only intend to hold bitcoins for a short time, there will be a considerable risk in doing so. The IVaR estimates for Gold generally lies around 50% over the Euro estimates, reflecting quite well that the price of Gold varies quite much more than the Euro/USD exchange rate.

Figure 4 illustrates the relationship between interval length and confidence level more clearly. It is no surprise that when the interval length gets longer, the potential loss for a given confidence level is bigger. We also remark how the relationship between the Bitcoin exchanges and the other assets seems to change as the confidence level and interval lengths change. The highest difference in the IVaR estimate is present for short intervals and high confidence levels (low α-quantiles), whereas a reduction in confidence level or an increase in interval length will lead to less differences in the IVaR estimates.

Figure 5 shows how the IVaR evolves within the first common[40] trading day with interval length equal to 30 minutes and at the 1% quantile. While Euro and Gold show a clear pattern, Bitstamp and Btc-e behave less systematically[41]. Note also the difference in absolute value of the IVaR. While the highest IVaR is almost

---

[40]The first day of the validation sample is a Saturday, the first day with trading for all assets is Monday, February 3rd.

[41]Similar patterns can be found for the other day of the week, they are however not shown.

the double of the lowest for Euro and more than double for Gold, the difference between the highest and lowest IVaR for Bitstamp and Btc-e is relatively much smaller. The lack of clear intraday seasonality in the IVaR in Bitcoin can be largely explained by the fact that the traders have no opening and closing hours to relate to, and by the global nature of the trading[42].
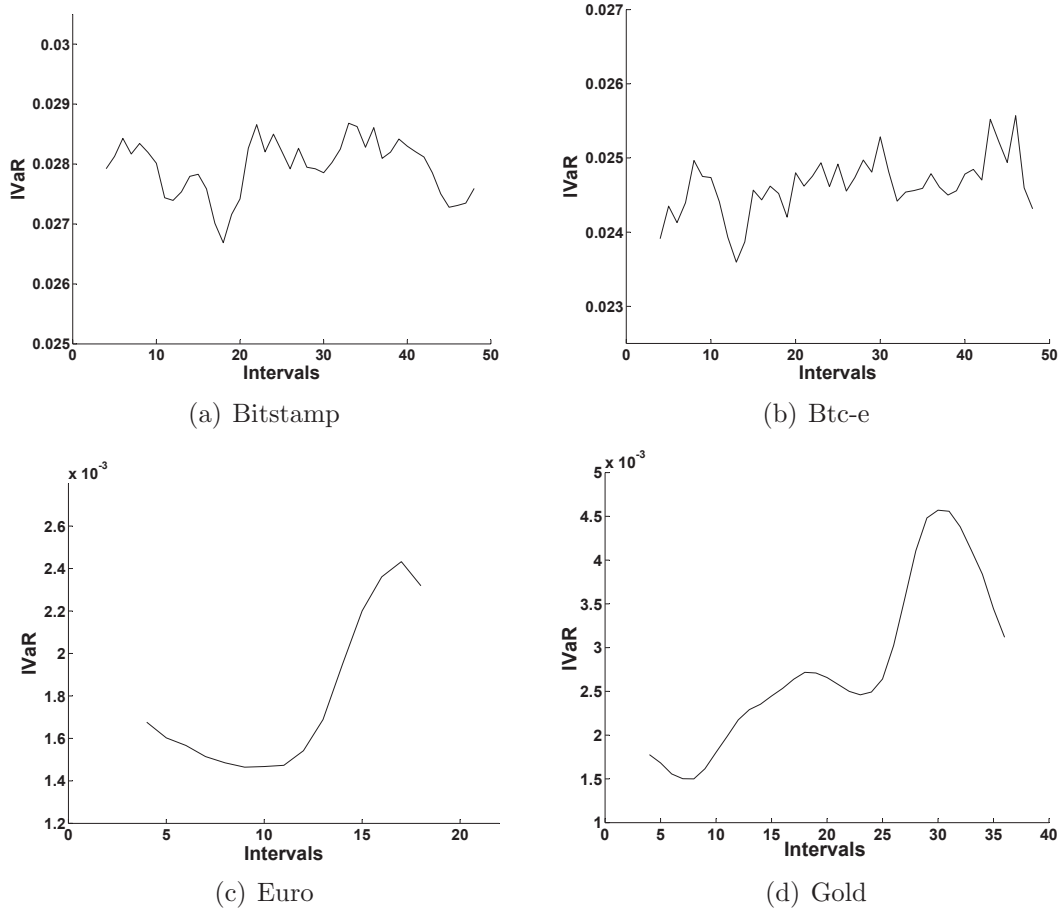

(a) Bitstamp

(b) Btc-e

(c) Euro

(d) Gold

Figure 5: Intraday patterns in the IVaR estimates, here for a the first Monday in the validation sample.

8.2   Relationship between durations and volatility

Since we have already seen that durations affect volatility, we thought it would be interesting to investigate this relationship a bit further. Following the literature

---

[42]We still note that durations and squared returns exhibit interday seasonality as explained earlier.

outlined in Section 3.5.1, we have studied whether price and return durations can give some additional insight on the intraday risk in Bitcoin.

Table 11 reports the mean return duration of different return thresholds, R. The attractiveness of this measure is that it is straightforward to use it in a comparison. We see that for the Bitcoin exchanges it takes around 10 minutes on average to move the price by 1% and somewhat over an hour to move the price by 2%. However, it takes significantly longer to move the price by 5%, indicating that the price tend to move and revert back frequently on short intraday horizons. As the trade durations in Table 2 and 4, the return durations are overdispersed[43], revealing a great deal of variability.

Euro and Gold has a lot longer mean return durations, and for Euro the price does not even change 2% from the initial price in the sample. Both have very few observations when R is 0.5% or higher. In short, return durations that happen intradaily in Bitcoin need several trading days to happen in Euro and Gold.

As explained in Section 3.5.1, price durations is an interesting concept to look at because of its relationship with volatility. Table 12 reports the average time it takes to move the price by C dollars for Bitstamp and Btc-e. Again we notice shorter durations on Btc-e than on Bitstamp except when $C = 20$ dollars. Depending on the exchange, it takes between half an hour and 40 minutes to move the price by 10 dollars, meaning that by the time it takes to confirm a transaction, the price might have changed quite significantly.

The hypotheses discussed in Section 3.5.1 are tested by turning back to our log-ACD model[44], this time modeling price durations with a threshold of $C = 3$ dollars[45]. The estimated parameters are summarized in Table 13.

The coefficient for the lagged intensity of trading, y_1, is found to be negative for Bitstamp and Btc-e. This indicates that after a period of high trading intensity, the expected price durations become shorter, and thus the volatility becomes higher. The clustering in the arrival times of transactions seems to influence the price process in a statistically significant way. This is in accordance with the hypothesis put forward by Easley and O'Hara (1992). For average volume, the coefficients are again found to be negative, which is in accordance with the hypothesis that higher transaction volume shortens expected price durations and raises volatility.

Using Eq. (3.8) we are now able to model volatility utilizing the extra information

---

[43]The standard deviation is very high compared to the mean.

[44]We use the log-EACD model, since there are unfortunately no simple expression for the conditional hazard function if the generalized gamma distribution is used.

[45]This greatly reduces the number of observations in the sample, for Bitstamp it reduces from 872 992 down to 35 812 observations.

in the included variables. Figure 12 plots this volatility for January and February 2014 for Bitstamp. We notice significantly higher volatility in February than January. The time of increased volatility corresponds well to the dates of the Mt.Gox insolvency debacle that shocked the market in late February. The weakness of this measure is that one cannot interpret the value in a straightforward way, however it is still an appealing model because one can predict future volatility by using information in past price durations. This may very well work as input in a Monte Carlo simulation, replacing the EGARCH specification in our model in Section 5. However we end this discussion here, by leaving the investigation between price durations and volatility in Bitcoin for future research.

8.3 Discussion

If we try to analyze the results obtained in Section 8.1 and 8.2, the obvious result is that for intradaily time horizons it is more risky to own bitcoins than Euro and Gold: For a given interval the potential price impact is much bigger at the Bitcoin exchanges, and the average time needed for a given change to happen is much shorter. In order to put in into a more practical perspective, we will now shortly discuss how this risk can impact the decision of which transaction medium to use.

While Gold is an excellent store of value and has been for quite some time, it is not easily used as a medium of transaction – the amount of merchants accepting Gold has dwindled the last few thousand years. When comparing it to Bitcoin it therefore might be natural to compare it as a long term investment. Given that we are focused rather on the short term fluctuations of the prices, we do not further elaborate on the fact that on the intradaily level, Bitcoin is far more risky than Gold. Fiat currencies on the other hand (as depicted by the Euro/USD exchange rate) are the most common medium of exchange in the world today, and will therefore be well suited for such an analysis.

If we look at uncertainty of value (exchange rate risk) as a cost of using a specific transaction medium, we get a clearer picture as to why more and more merchants are accepting Bitcoin as payment. When looking at the estimates for Bitstamp in Table 8, we see that for the next 2 hours[46] we are 90% sure that the Bitcoin price will not have lost more than 2.46% of its value. Given that transaction fee is voluntary, the total cost of a Bitcoin transaction primarily consists of the exchange rate risk[47].

---

[46]And on average confirmation will take considerable less time.

[47]Here we assume that one already owns the bitcoins, as there usually are fees related to the purchase of bitcoins.

For fiat currencies the picture is quite opposite. The estimated IVaR for Euro at the 10% quantile with an interval length of 2 hours is only 0.20%, less than a tenth of Bitstamp's level. It is when it comes to transaction costs that fees for using traditional systems really start adding up[48] – credit cards for instance typically have a transaction fee between 2-3% (Shy and Wang, 2011), while remittances like discussed in Section 2.3 have above 8% fees on average.

The cost of added risk is not directly comparable to transaction cost since it is not something one explicitly has to pay, but one can rather try to figure out which alternative is the most desirable. Using Bitcoin, one is exposed to a changing exchange rate, which means that some of the time you will have to pay more or receive less bitcoins than what you thought at time of purchase, at the confidence levels seen in Table 8 and Figure 4. What is interesting though is that we are actually fairly sure that the exchange rate will drop less than the fiat transaction cost, which you have to pay *every* transaction. Something that is important to keep in mind though is that there will of course be situations where the value of a bitcoin will drop more than the levels in Figure 8. For transactions sensitive to exchange rate risk, it therefore seems smart to somehow hedge against this risk[49] or just use an alternative transaction medium.

However, the Bitcoin exchange rate also has the possibility of not changing during the given interval length, or even rise in value. Given that investing in Bitcoin would actually have been one of 2013's best investments with a whopping return of over 4000%, it is no wonder that the number of people using or considering Bitcoin as an alternative to traditional payment systems is quickly rising. While the long term investment horizon is not the topic of this paper, we remark that if you think the price of bitcoins is going to rise in the future, the average transaction discrepancy[50] should be in your favor.

As long as merchants charge the same prices for goods independent of the payment method, sticking with traditional system of fiat currencies and credit cards seems to be a safer alternative. Especially as long as holding bitcoins even for only a couple of hours will expose you to considerable risk. However, if these factors change, moving away from the obvious safety issues of leaving credit card information online (see Section 2.3) will probably be tempting for more than just the Bitcoin enthusiasts.

---

[48]Unless one uses cash, which is impossible for e-commerce.
[49]Like going short in bitcoins
[50]Difference between exchange rate at time of transfer and at time of confirmation.

9 Conclusion

In this paper we have modeled the intraday risk in Bitcoin by using a log-ACD-ARMA-EGARCH model. Based on this model we have then run Monte Carlo simulations, in order to get an intraday Value at Risk estimate for the exchange rate risk. We have then done the same analysis for the Euro/USD exchange rate as well as the price of Gold, in order to be able to do a comparison between these.

The IVaR estimates clearly show that Bitcoin is by far the riskiest asset of the three investigated when regarding exchange rate risk. Both Bitstamp and Btc-e consistently have IVaR levels around 10 times or more above those of Euro and Gold. The results of the price and return durations analysis give the same conclusion – it takes considerably less time to move the Bitcoin price than the price of Gold or the Euro/USD exchange rate.

However, as discussed in the previous section, the potential downside Bitcoin users face is not discouraging. We are actually quite confident that the Bitcoin exchange rate will drop less than the incurred costs of using traditional methods of payment and money transferring. Users of the Bitcoin system should anyhow be cognizant of who will be assuming the exchange rate risk when doing transactions, especially when performing transactions that are sensitive to this risk. And while it might create barriers for mainstream adoption, in our opinion the benefits that Bitcoin promises should be enough to provide a viable alternative to the existing financial system.

9.1 Key model characteristics and future work

In order to sum up, we shortly note some of the key characteristics of the model, and also some areas that we think could be really interesting to explore if we were to expand the analysis further.

Given that there is a relationship between durations and volatility, by explicitly including the duration in the volatility specification we are better able to model the return series. Also by making the modeling and simulating the whole trade sequence, the IVaR estimate will have a higher information content than a model which only uses observations at regular intervals. We also point out the flexibility of the model: Once the model parameters have been estimated, IVaR estimates can be estimated for any time horizon. From Figure 12 and Table 7 we clearly see that even though the volatility in January and February were quite different, the model still performs well during backtesting.

The fact that the returns are assumed conditionally normal seems to give satisfactory results during the backtesting, especially for the Bitcoin exchanges. However, it would have been interesting to do the analysis again with something else than the normal or Student t distribution, perhaps a multivariate distribution or an extreme value distribution. What is also worth noting is that by converting the irregular time series back to a regular spaced interval time series some information will necessarily be lost. Making a Value at Risk estimate or some other risk measurement that directly uses the irregular time series could therefore be an interesting topic. Lastly we think that further elaborating on the instantaneous volatility model in Section 3.5.1 would be really interesting. One could for instance use this model or a modified version as the volatility specification when modeling the returns, and thus use more of the information generated by the durations.

References

Anat R. Admati and Paul Pfleiderer. A Theory of Intraday Patterns: Volume and Price Variability. *The Review of Financial Studies*, Vol. 1(No. 1):pp. 3–40, 1988.

Carol Alexander. *Market Risk Analysis*. Wiley, 2008.

Stanislav Anatolyev and Dmitry Shakin. Trade intensity in the Russian stock market: dynamics, distribution and determinants. *Applied Financial Economics*, Vol. 17:pp. 87–104, 2007.

Marc Andreessen. Why Bitcoin Matters – New York Times. `http://mobile.nytimes.com/blogs/dealbook/2014/01/21/why-bitcoin-matters/`, Jan. 2014. Accessed: 15-May-2014.

Luc Bauwens. Econometric Analysis of Intra-daily Trading Activity on Tokyo stock Exchange. *Monetary and Ecomonic Studies, Institute for Monetary and Economic Studies, Bank of Japan*, Vol. 24:pp. 1–23, 2006.

Luc Bauwens and Pierre Giot. The logarithmic ACD model: An Application to the Bid-Ask Quote Process of Three NYSe Stocks. *Annales D'Economic Et De Statistique*, Vol. 60:pp. 117–149, 2000.

Luc Bauwens and David Veredas. The stochastic conditional duration model: a latent variable model for the analysis of financial durations. *Journal of Econometrics*, Vol. 119:pp. 381–412, 2004.

Marcus Bensasson and Georgios Georgiou. Cyprus Sets Levy on Bank of Cyprus Uninsured Depositors at 47.5% – Bloomberg. `http://www.bloomberg.com/news/2013-07-30/cyprus-sets-levy-on-bank-of-cyprus-uninsured-depositors-at-47-5-.html`, July 2013. Accessed: 28-May-2014.

Bitpay.com. Press Release: BitPay Exceeds 100,000,000 dollars in Bitcoin Transactions Processed. `http://blog.bitpay.com/2013/12/11/bitpay-exceeds-100-000-000-in-bitcoin-transactions-processed.html`, Dec. 2013. Accessed: 05-May-2014.

Tim Bollerslev. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, Vol. 31:pp. 307–327, 1986.

Tim Bollerslev, Ray Y. Chou, and Kenneth F. Kroner. ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics*, Vol. 52:pp.5–59, 1992.

Danny Bradbury. Additional Sources Confirm Chinas Payment Processor Ban, Bitcoin Price Falls 200 dollars – Coindesk. `http://www.coindesk.com/sources-confirm-chinas-payment-processor-ban-bitcoin-price-falls/`, Dec. 2013. Accessed: 27-April-2014.

Vit Bubak and Filip Zikes. Trading Intensity and Intraday Volatility on the Praque Stock Exchange: Evidence from an Autoregressive Conditional Duration Model. *Czezh Journal of Economics and Finance*, Vol. 56:pp. 223–243, 2006.

David Chaum, Amos Fiat, and Moni Naor. Untraceable electronic cash. *Advances in Cryptology*, Vol. 88:pp. 319–327, 1990.

Peter F. Christoffersen. Evaluating Interval Forecasts. *International Economic Review*, Vol. 39(No. 4):pp. 841–862, 1998.

Gilbert Colletaz, Christophe Hurlin, and Sessi Tokpavi. Irregularly Spaced Intraday Value at Risk (ISIVaR) Models. Forecasting and Predictive Abilities. *LEO, University of Orleans, discussion paper*, 2007.

Aswath Damodaran. Value At Risk (VAR). `people.stern.nyu.edu/adamodar/pdfiles/papers/VAR.pdf`, 2014. Accessed: 10-May-2014.

James D'Angelo. Why Bitcoin is Not a Ponzi Scheme. `http://youtu.be/7u4F8cpzqao`, Nov. 2013. Accessed: 12-May-2014.

Serge Darolles, Christian Gourieroux, and Gaelle Le Fol. Intraday Transaction Price Dynamics. *Annales D'Economie Et De Statstique*, Vol. 60, 2000.

Douglas W. Diamond and Robert E. Verrecchia. Constraints on short-selling and asset price adjustment to private information . *Journal of Financial Economics*, Vol. 18(No. 2):pp. 277 – 311, 1987.

Georges Dionne, Pierre Duchesne, and Maria Pacurar. Intraday Value at Risk (IVaR) using tick-by-tick data with application to the Toronto Stock Exchange. *Journal of Empirical Finance*, Vol. 16:pp. 777–792, 2009.

Alfonso Dufour and Robert F. Engle. Time and the Price Impact of a Trade. *The Journal of Finance*, Vol. 6, 2000.

Alfonso Dufour and Robert F. Engle. The ACD Model: Predictability of the Time Between Consecutive Trades. *Discussion papers in finance, The University of Reading*, 2000-05.

David Easley and Maureen O'Hara. Time and the process of security price adjustment. *Journal of Finance*, Vol. 47:pp. 577–606, 1992.

Robert F. Engle. The Econometrics of Ultra-High-Frequency Data. *Econometrica*, Vol. 68(No. 1):pp. 1–22, 2000.

Robert F. Engle and Jeffrey R. Russell. Forecasting the Frequency of Changes in Quoted Foreign Exchange Prices with the Autoregressive Conditional Duration Model. *Journal of Empirical Finance*, Vol. 4:pp. 187–212, 1997.

Robert F. Engle and Jeffrey R. Russell. Autoregressive Conditional Duration: a New Model for Irregularly Spaced Transaction Data. *Econometrica*, Vol. 66:pp. 1127–1162, 1998.

E.J. Fagan. Bitcoin and international crime – The Baltimore Sun. `http://www.baltimoresun.com/news/opinion/oped/bs-ed-bitcoin-20131125,0,3265347.story`, Nov. 2013. Accessed: 22-May-2014.

Maureen Farrell. Bitcoin prices surge post-Cyprus bailout – CNN. `http://money.cnn.com/2013/03/28/investing/bitcoin-cyprus/`, March. 2013. Accessed: 11-April-2014.

Dingan Feng, George J. Jiang, and Peter Song. Stochastic Conditional Duration Models with Leverage Effect for Financial Transaction Data. *Journal of Financial Econometrics*, Vol. 2:pp. 390–421, 2004.

Marcelo Fernandes and Joachim Grammig. A family of Autoregressive Conditional Duration Models. *Journal of Econometrics*, Vol. 130:pp. 1–23, 2006.

Lawrence Fisher. Some New Stock Market Indices. *The Journal of Business*, Vol. 39(No. 1):pp. 191–225, 1966.

Glenn Fleishmann. On the Matter of Why Bitcoin Matters. `https://medium.com/the-magazine/23e551c67a6`, Jan. 2014. Accessed: 12-May-2014.

Forbes. 2013: Year Of The Bitcoin. `http://www.forbes.com/sites/kitconews/2013/12/10/2013-year-of-the-bitcoin/`, Dec. 2013. Accessed: 6-May-2014.

Frank Gerhard and Nikolaus Hautsch. Determinants of inter-trade durations and hazard rates using proportional hazard arma models. *Center of Finance and Econometrics, University of Konstanz in its series CoFE Discussion Paper with number 00-20.*, 2000.

Eric Ghysels and Joanna Jasiak. GARCH for Irregularly Spaced Financial Data: The ACD-GARCH Model. *Studies in Nonlinear Dynamics and Econometrics*, Vol. 2:pp. 133–149, 1998.

Eric Ghysels, Christian Gourieroux, and Joanna Jasiak. Stochastic volatility duration models. *Journal of Econometrics*, Vol. 119:pp. 413–433, 2004.

Pierre Giot. Market Risk Models for Intraday Data. *The European Journal of Finance*, Vol. 11:pp. 309–324, 2005.

Pierre Giot and Joachim Grammig. How large is liquidity risk in an automated auction marked? *Empirical Economics*, Vol. 30:pp. 867–887, 2006.

Ronald E. Glaser. Bathtub and Related Failure Rate Characterizations. *Journal of the American Statistical Association*, Vol. 75:pp. 667–672, 1980.

A.Z Gorski, S. Drozdz, and J.Speth. Financial multifractality and its subtleties: an example of DAX. *Physica A*, Vol. 316:pp. 496–510, 2002.

Joachim Grammig and Kai-Oliver Maurer. Non-Monotonic Hazard Functions and the Autoregressive Conditional Duration Model. *Econometrics Journal*, Vol. 3: pp. 16–38, 2000.

Joachim Grammig and Marc Wellner. Modeling the interdependence of volatility and inter-transaction duration processes. *Journal of Econometrics*, Vol. 106(No. 2):pp. 369–400, 2002.

Sven Ove Hansson. Risk. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Spring 2014 edition, 2014.

J.D. Harrison. Bitcoin presents promising (but risky) payment alternative for small businesses – The Washington Post. `http://www.washingtonpost.com/business/on-small-business/bitcoin-presents-promising-but-risky-payment-alternative-for-small-businesses/2014/04/02/6637aad8-ba7b-11e3-9a05-c739f29ccb08_story.html`, April 2014. Accessed: 28-May-2014.

Nikolaus Hautsch. Modelling intraday trading activity using Box-Cox ACD models. *Working Paper, Center of Finance and Econometrics, University of Konstanz*, 2002.

Nikolaus Hautsch. *Econometrics of Financial High-Frequency Data*. Springer, 2nd edition, 2012.

Alex Hern. Missing: hard drive containing Bitcoins worth 4m in Newport landfill site – The Guardian. `http://www.theguardian.com/technology/2013/nov/27/hard-drive-bitcoin-landfill-site`, Nov. 2013. Accessed: 21-May-2014.

Nick Jones. Fiat Currency: Using the Past to See into the Future. `http://dailyreckoning.com/fiat-currency/`, 2014. Accessed: 28-May-2014.

Philippe Jorion. *Value at Risk: The new benchmark for managing financial risk.* McGraw-Hill, 3rd edition, 2007.

Paul H. Kupiec. Techniques for verifying the accuracy of risk management models. *The Journal of Derivatives*, Vol. 3(No. 2):pp.73–84, 1995.

Katrina Lamb. Measuring and Managing Investment Risk. `http://www.investopedia.com/articles/08/risk.asp`, Aug. 2011. Accessed: 15-April-2014.

Leslie Lamport, Robert Shostak, and Marshall Pease. The Byzantine Generals Problem. *ACM Transactions on Programming Languages and Systems*, Vol. 4 (No. 3):pp. 382–401, 1982.

Eunjung Lee. Why Stock Investors Split Their Orders? An Analysis of Strategic Trading Behaviors in the Korea Stock Exchange. *Asia-Pacific Journal of Financial Studies*, Vol. 37(No. 3):pp. 391–424, 2008.

LexisNexis Group. 2013 LexisNexis®True Cost of Fraudˆ SM Study. `http://www.lexisnexis.com/risk/downloads/assets/true-cost-fraud-2013.pdf`, Sep. 2013. Accessed: 28-May-2014.

Sharon Xaiowen Lin and Michael N. Tamvakis. Effects of NYMEX Trading on IPE Brent Crude Futures Market: A Duration Analysis. *Energy Policy*, Vol. 32: pp. 77–82, 2004.

Thomas J. Linsmeier and Neil D. Pearson. Value at Risk. *Financial Analyst Journal*, Vol. 56(No. 2):pp. 47–67, 2000.

Andrew W. Lo and A. Craig MacKinlay. An Econometric Analysis of Nonsynchronous Trading. *Journal of Econometrics*, Vol. 45:pp. 181–211, 1990.

Asger Lunde. A Generalized Gamma Autoregressive Conditional Duration Model. *Discussion Paper, Aalborg University*, 1999.

Chris Mack. Is This Time Different for the Dollar? `http://www.resourceinvestor.com/2011/01/24/is-this-time-different-for-the-dollar`, Jan. 2011. Accessed: 28-May-2014.

Eric Mack. The Bitcoin Pizza Purchase That's Worth 7 Million Today – Forbes. `http://www.forbes.com/sites/ericmack/2013/12/23/the-bitcoin-pizza-purchase-thats-worth-7-million-today/`, Dec. 2013. Accessed: 10-April-2014.

Sarah Meiklejohn, Marjory Pomarole, Grant Jordan, Kirill Levchenko, Damon
   McCoy, Geoffrey M.Voelker, and Stefan Savage. A Fistful of Bitcoins: Charac-
   terizing Payments Among Men with No Names. 2013.

Andrew Miller and Joseph J. LaViola. Anonymous Byzantine Consensus from
   Moderately-Hard Puzzles: A Model for Bitcoin. *Working Paper*, 2014. URL
   `https://socrates1024.s3.amazonaws.com/consensus.pdf`.

Edward Moyer. PayPal shuts out WikiLeaks. `http://www.cnet.com/news/`
   `paypal-shuts-out-wikileaks/`, Dec. 2010. Accessed: 28.05.2014.

Satoshi Nakamoto. Bitcoin: A Peer-to-Peer Electronic Cash System. 2008.

Daniel B. Nelson. Conditional Heterskedasticity in Asset Return: A New Ap-
   proach. *Econometrica*, Vol. 59(No. 2):pp. 347–370, 1991.

Marcelo Perlin. Estimation and Simulation of ACD mod-
   els in Matlab. `https://sites.google.com/site/marceloperlin/`
   `matlab-code/estimation-and-simulation-of-acd-models-in-matlab`, 2012.
   Accessed: 14-March-2014.

Alanna Petroff. Bitcoin reshaping future of money – CNN. `http://money.cnn.com/`
   `2013/12/10/technology/bitcoin-currency-fred-wilson/`, Dec. 2013. Accessed:
   19-May-2014.

Michael Riley, Ben Elgin, Dune Lawrence, and Carol Matlack. Missed
   Alarms and 40 Million Stolen Credit Card Numbers: How Target Blew
   It – Bloomberg Businessweek. `http://www.businessweek.com/articles/`
   `2014-03-13/target-missed-alarms-in-epic-hack-of-credit-card-data`,
   March 2014. Accessed: 28-May-2014.

Richard Roll. A Simple Implicit Measure of the Effective Bid-Ask Spread in an
   Efficient Market. *The Journal of Finance*, Vol. 39(No. 4):pp. 1127–1139, 1984.

Ozy Shy and Zhu Wang. Why Do Payment Card Networks Charge Proportional
   Fees? *American Economic Review*, Vol. 101(No.4):pp. 1575–90, 2011.

Jon Southurst. Mt. Gox Files for Bankruptcy, Claims 63.6m Debt – Coindesk.
   `http://www.coindesk.com/mt-gox-files-bankruptcy-claims-63-6m-debt/`,
   Feb 2014. Accessed: 8-May-2014.

Tim Swanson. Learning from Bitcoin's past to improve its future. `http://www.`
   `ofnumbers.com/wp-content/uploads/2014/04/Learning-from-Bitcoins-past.`
   `pdf`, Apr. 2014. Accessed: 22-May-2014.

The Bitcoin Wiki. Deflationary spiral. `https://en.bitcoin.it/wiki/Deflationary_spiral`, 2014a. Accessed: 28-May-2014.

The Bitcoin Wiki. Buying Bitcoins (the newbie version). `https://en.bitcoin.it/wiki/Buying_Bitcoins_(the_newbie_version)`, 2014b. Accessed: 10-May-2014.

The World Bank. Migration and Remittances: Recent Developments and Outlook. `http://siteresources.worldbank.org/INTPROSPECTS/Resources/334934-1288990760745/MigrationandDevelopmentBrief22.pdf`, April 2014. Accessed: 28-May-2014.

Jonathon Trugman. Welcome to 21st-century Ponzi scheme: Bitcoin – New York Post. `http://nypost.com/2014/02/15/welcome-to-21st-century-ponzi-scheme-bitcoin/`, Feb. 2014. Accessed: 14-May-2014.

Yiu-Kuen Tse and Shouwei Liu. Intraday value-at-risk: An asymmetric autoregressive conditional duration approach. *Working Paper*, 2013.

David Veredas, Juan M. Rodriquez-Poo, and Antoni Espasa. On the (intradaily) seasonality and dynamics of a financial point process: A semiparametric approach. *Working Paper*, 2001.

Washington State Department of Financial Institutions. Consumer Alert: Bitcoin and Virtual Currencies. `http://www.dfi.wa.gov/consumers/alerts/bitcoin.htm`, 2014. Accessed: 16-May-2014.

Joe Weisenthal. Here's The Answer To Paul Krugman's Difficult Question About Bitcoin – BusinessInsider. `http://www.businessinsider.com/why-bitcoin-has-value-2013-12`, Dec. 2013. Accessed: 3-May-2014.

Pipat Wongsaart, Jiti Gao, and David E. Allen. A semiparametric approach to nonlinear ACD model. *Conference paper, Proceedings of the 38th Australian Conference of Economists*, 2009.

Michael Yuanjie Zhang, Jeffrey R. Russell, and Ruey S. Tsay. A nonlinear autoregressive conditional duration model with applications to financial transaction data. *Journal of Econometrics*, Vol. 104:pp. 179–207, 2001.

Eric Zivot. Analysis of High Frequency Financial Data: Models, Methods and Software, 2005.

A Appendix

A.1 More details on the workings of Bitcoin

A.1.1 The block chain and the mining process

A Bitcoin address is an identifier of 27-34 alphanumeric characters, that represents a possible destination for a Bitcoin transaction[51]. The block chain is a big list of all the transactions ever made, and thus by closer inspection one can find which of these addresses that hold what balances. This ledger is publicly available, and every computer connected directly[52] to the Bitcoin network has a full copy of it. Transactions are pooled into blocks that are ordered sequentially. Hence the term, block chain, a chain of blocks. The network can then use the block chain to distinguish between those transactions that are legitimate, and those trying to use bitcoins already spent elsewhere.

In order to add a new block to the block chain, users compete to solve a hard cryptographic problem[53]. The solution is known as a proof-of-work, and has to be included in the block to ensure its validity. Once per block, a special transaction is allowed: The transaction that creates new bitcoins[54]. These new bitcoins are awarded to the user who find it. At the time of writing the reward is 25 bitcoins, but this will halve every 210 000 blocks. In order to try to keep the rate at which new blocks are found relatively constant at 10 minutes, the difficulty level of the problem is increased when the computational power of the network increases[55]. This makes the creation of new bitcoins predictable in both time and amount. The process of trying to solve these cryptographic problems is popularly know as mining, and the users who compete are thus known as miners.

The process that creates new blocks requires a lot of computational resources. Today one can buy customized hardware for this process, and at the time of writing the Bitcoin network has more than 800 times the hash power of the World's 500 most powerful supercomputers combined (Fleishmann, 2014). The reward for creating new blocks actually serves two purposes: In addition to increase the total number of bitcoins in existence, it also subsidizes the maintenance of the block

---

[51]Further details about Bitcoin addresses is presented in A.1.3

[52]Running the full Bitcoin client.

[53]This is constructed in such a way that it is really difficult to find the solution, but once a solution is know, verifying that it is valid is trivial.

[54]Remember from before that "a bitcoin" does not actually exist. The creation of a new bitcoin is thus just a special transaction.

[55]Otherwise the expected time until a solution is proposed would decrease.

chain. However since the "production rate" declines over time, and there is a limit on how many bitcoins there will ever be, miners need an additional incentive to maintain the block chain. A transaction can therefore contain a voluntarily transaction fee, which will be awarded to the miner who included the transaction in the block chain. However, which transactions to include in the next block is entirely up to each miner.

A.1.2   Bitcoin transactions

After a transaction is broadcast to the Bitcoin network, it may be included in the next block that is published to the network. We say "may" because the next block can already be full, or the creator of the next block might not include it. In practice however, most transactions will be included even without a transaction fee, as it will be beneficial for the miner to do so. In order to understand why, we note that if transactions did not confirm, the Bitcoin network would not have any value, and thus there would be no value in mining. Figure 6 illustrates this point. It takes about an average of 10 minutes to confirm a transaction, which is the same as the time between each block is added to the block chain.
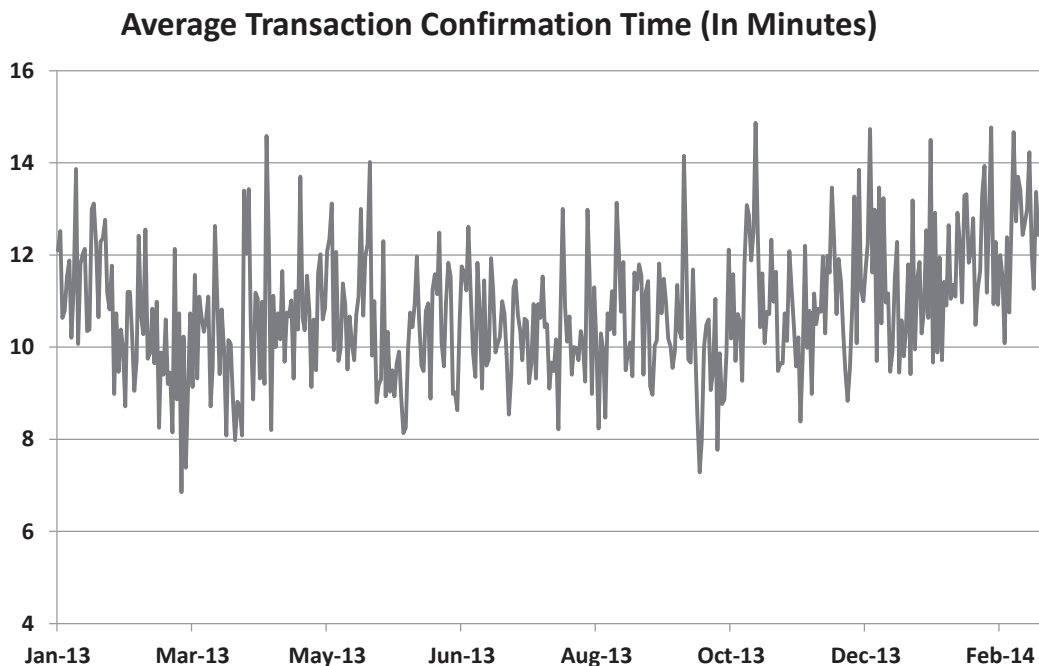

Figure 6: Average transaction time in minutes, per day (Jan 2013 - March 2014), source blockchain.info.

However in practice one confirmation is rarely considered "enough". To be protected against double spending and invalid blocks, more confirmations are usually needed[56]. Whenever a block is created after the transaction was first broadcast, the number of confirmations is increased by one. Merchants, exchanges, individuals and others that accept bitcoins as payment can set their own threshold on how many confirmations are needed before a transaction is considered valid. The standard adopted by most seems to be 6 confirmations ($\sim$1 hour), and the official Bitcoin client also shows transactions as "unconfirmed" until 6 blocks confirm the transaction. As illustrated by Figure 7, a transaction made at time $t = 0$ is unconfirmed until 6 new blocks point back to the block with the transaction.
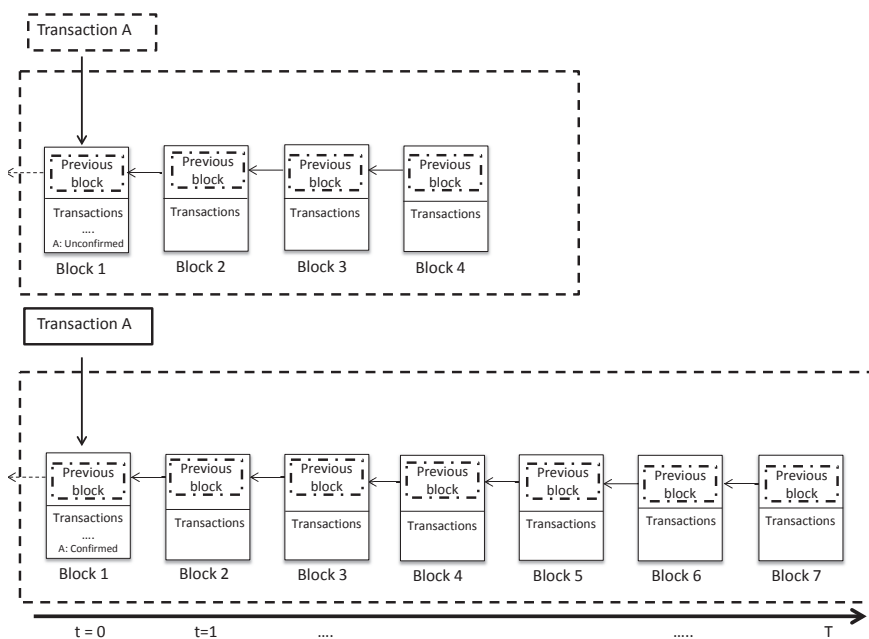

Figure 7: Illustration of transaction in block chain.

A.1.3   Bitcoin addresses

Addresses can be generated at no cost by any user of Bitcoin. Most Bitcoin addresses are 34 characters, consisting of random digits and uppercase and lowercase letters, beginning with the number 1 or 3. The number "0", the uppercase letter "O", uppercase letter "I" and lowercase letter "l" are never used to prevent visual ambiguity. Several of the characters of a Bitcoin address are used as a checksum

---

[56]We refer the curious reader to Appendix A.1.4 for an explanation on these issues.

so that typographical errors can be automatically found and rejected. This is also to ensure that a 33-character, or shorter, address is in fact valid and is not simply an address with a missing character.

Bitcoin uses public key encryption techniques for security. Simply put, it uses a key pair that is mathematically related so that a user who knows one of these keys can perform an action that the "knower" of the other key can verify, but not recreate. This allows the holder of the private key to prove to the public that he/she has it. The public key identifies a Bitcoin address, while the private key allows access to the bitcoins found at that address.

In a Bitcoin wallet, each address is represented by three numbers, the address itself and the key pair. The public key is derived from the private key, and given the computational power of computers today, deriving the private key from the public key is essentially impossible. The address is linked to the key pair by being a hash[57] of the public portion of the key. From the public key it requires three steps to derive the address, first by applying two hash functions, and finally adding the checksum[58]. Simply put, the private key is used for sending bitcoins, the address is used to receive bitcoins and the public key to make the network able to verify the transaction.

Private keys are typically stored in the wallet file, for instance locally on a computer, however paper wallets[59] are also possible. The private key has one function: It is needed to create valid transactions that spend bitcoins on the address. If the private key to an address is lost, any associated bitcoins are effectively lost forever. While this makes for amusing headlines in the news, see Hern (2013), it is now estimated that around 15% of all mined bitcoins have been lost forever this way[60] (Swanson, 2014).

A.1.4   More on the block chain

Each block contains:

- A timestamp

---

[57]A hash is the output of a hash function that maps data of arbitrary length to data of a fixed length. It is used in cryptography because it is trivial to generate hash value from input data and easy to verify that the data matches the hash, but near impossible to "fake" a hash value.

[58]To ensure that mistyping one character does not send your bitcoins into a black hole.

[59]A mechanism for storing bitcoins offline as a physical document.

[60]The broader economic consequence of lost bitcoins is however marginal, one could argue that the entire world–wide economy could operate on a single bitcoin, due to the divisibility of bitcoin.

- The hash of the previous block as a reference

- At least one transaction

- The Merkle Root

- The block's own hash

- Difficulty statement

Every block contains a hash of the previous block, which creates the chain of blocks from the genesis block to the current one. This ensures that each block is guaranteed to come chronologically, because the previous block's hash would otherwise not be known. Each block is also computationally impractical to modify once it has been in the chain for a while because every block after it would also have to be regenerated. These properties are what make double-spending of bitcoins very difficult.

Honest clients only build onto a block if it is the latest block in the longest valid chain. Note that length in this sense is calculated as total combined difficulty of that chain and not the number of blocks. A chain is valid if all of the blocks and transactions within it are valid, and only of it starts with the genesis block. The genesis block is the first block of the block chain.

As many miners compete to find the next block, there will be situations where more than one valid block is discovered in a short time span. This is resolved as one of the new chains progresses to a greater length, at which any client that receives the newest block knows to discard the shorter chain. These discarded blocks are referred to as orphaned blocks. This process also ensures that the blockchain is considered impossible to forge, as long as most clients are honest. When a transaction is submitted to the network, it is passed on peer to peer by all clients.

Orphaned blocks are not used for anything and the chain of orphaned blocks are referred to as an invalid chain. To resolve this, when a client switch to another, longer chain, all valid transactions of the blocks inside the invalid chain are re-added to the pool of queued transactions and will be included in another block. The reward for the blocks on the shorter chain however is lost.

The Merkle Root is what makes it possible to verify Bitcoin transactions without running a full network node[61]. Every transaction has a hash associated with it. In a block, all of the transaction hashes in the block are themselves hashed[62], and the result is the Merkle root. This is included in the block header, and with this

---

[61]Running the Bitcoin client.

[62]Sometimes several times – the exact process is quite complex.

scheme it is possible to securely verify that a transaction has been accepted by
the network by only downloading the tiny block headers[63]. One can not check the
transaction oneself, but confirm that a node in the network has accepted it, and
see that blocks added after it further confirm the transaction is accepted by the
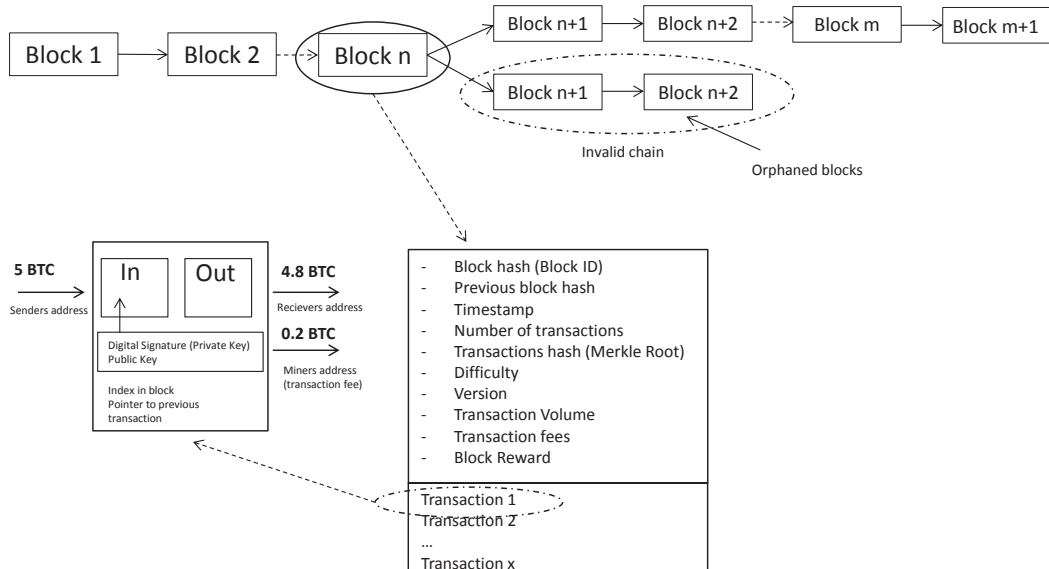network.

Figure 8: Illustration of the block chain, block structure and transaction.

---

[63]Downloading the entire block chain, which at the time of writing has a size of approximately
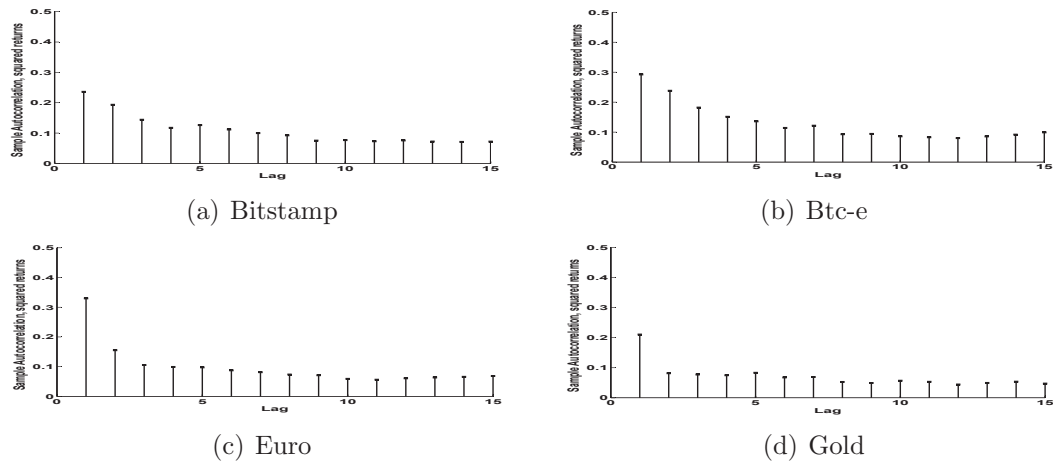18 GB, is thus unnecessary.

A.2   Figure and tables

(a) Bitstamp

(b) Btc-e

(c) Euro

(d) Gold

Figure 9: Sample autocorrelation function for diurnally adjusted squared returns in estimation sample.

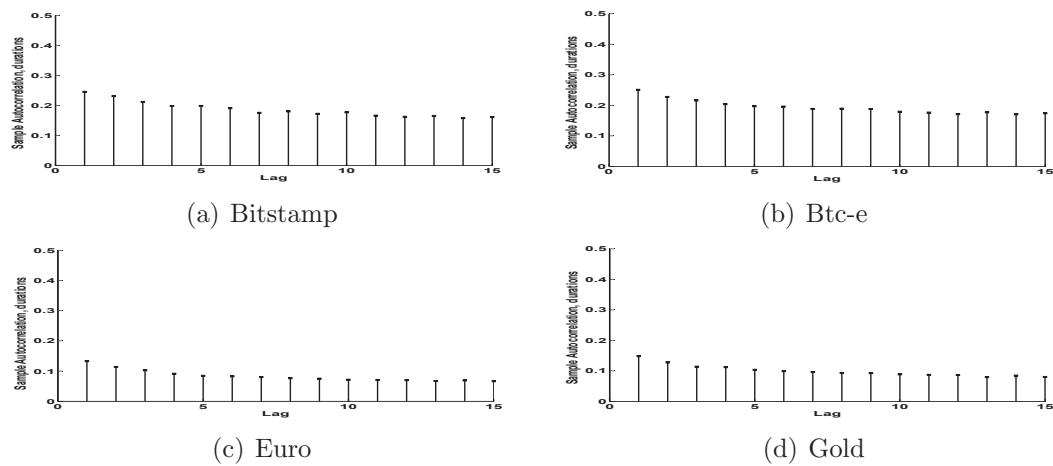(a) Bitstamp

(b) Btc-e

(c) Euro

(d) Gold

Figure 10: Sample autocorrelation function for diurnally adjusted durations in estimation sample.
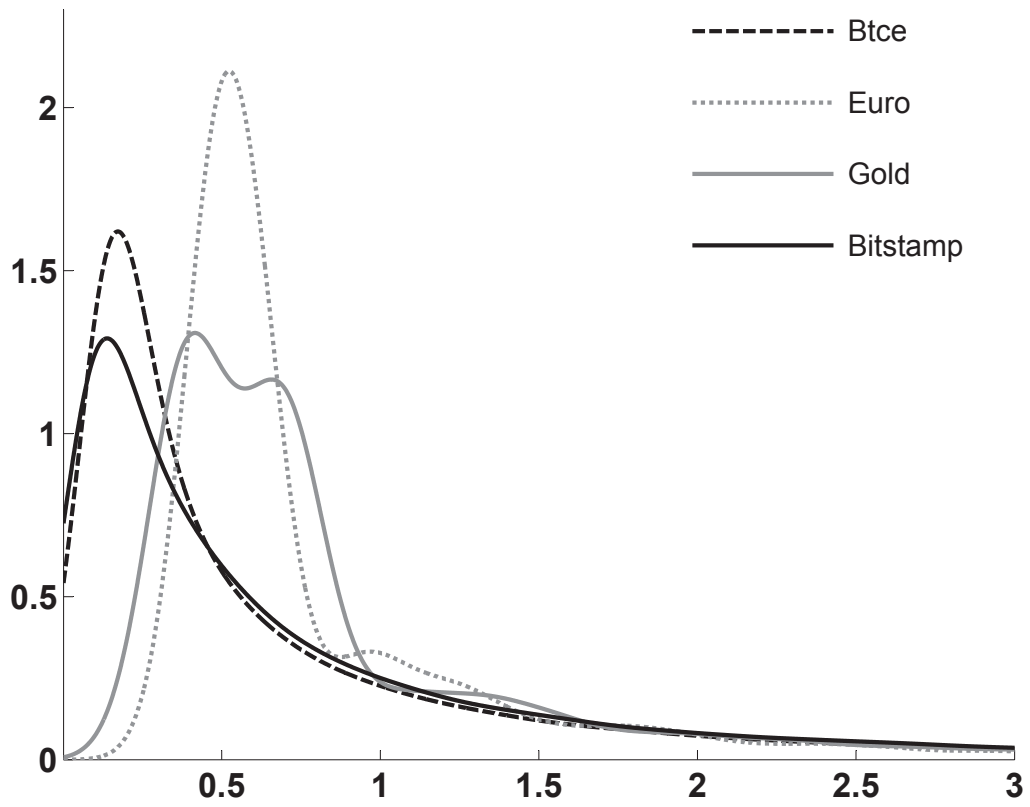
Figure 11: Probability density functions for diurnally adjusted durations.
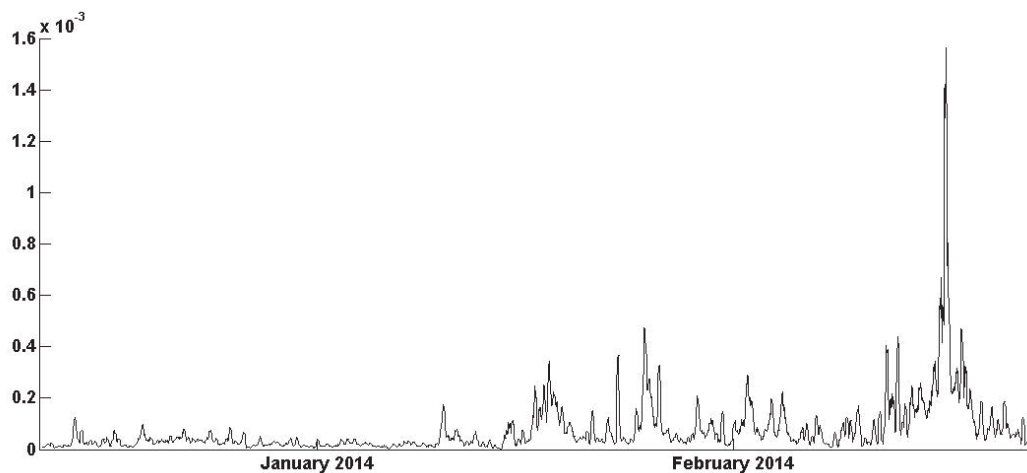
Figure 12: Instantanous volatility of the price at Bitstamp in Jan-Feb 2014.

Table 9: Weekday average for durations and squared returns.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Duration							
Bitstamp	11.77	10.96	16.09	14.49	11.22	18.50	17.98
Btc-e	6.38	6.04	7.75	7.52	5.87	8.54	7.77
Euro	2.21	2.06	2.02	1.86	1.91		
Gold	2.64	2.57	2.52	2.48	2.49		
Square Return							
Bitstamp	4.31E-06	7.53E-06	3.83E-06	2.65E-06	5.84E-06	3.03E-06	2.78E-06
Btc-e	2.55E-06	3.77E-06	2.53E-06	2.18E-06	3.31E-06	2.31E-06	2.41E-06
Euro	2.60E-10	2.67E-10	2.76E-10	3.72E-10	4.10E-10		
Gold	1.97E-09	1.17E-09	1.39E-09	1.35E-09	1.55E-09		

Table 10: Average total trading volume across weekdays.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Average total volume							
Bitstamp	225949	249756	146072	169985	255312	101948	120791
Btc-e	141099	172930	117036	123450	177096	88865	110069

Table 11: Return durations.

R	0.10 %	0.50 %	1 %	2 %	5 %
Bitstamp					
Mean	0 min 24 sec	3 min 12 sec	14 min 34 sec	82 min 56 sec	∼12 hrs
St.Dev	1 min 27 sec	15 min 1 sec	64 min 17 sec	269 min 28 sec	∼27.5 hrs
Obs	208620	26525	5830	1024	118
Number of trades	4.18	33	150	851	7392
Btc-e					
Mean	0 min 11 sec	1 min 50 sec	10 min 11 sec	67 min 3 sec	∼12.75 hrs
St.Dev	1 min 02 sec	11 min 5 sec	51 min	247 min 56 sec	∼35 hrs
Obs	473143	46563	8348	1267	108
Number of trades	4.78	49	270	1782	20453
Euro					
Mean	187 min 12 sec	∼72 hrs	∼283 hrs	—	—
St.Dev	∼9 hrs	∼61 hrs	∼230 hrs	—	—
Obs	441	19	4	—	—
Number of trades	12354	283019	1197352	—	—
Gold					
Mean	24 min 15 sec	∼10 hrs	∼38 hrs	∼145 hrs	∼957 hrs
St.Dev	162 min 28 sec	∼16 hrs	∼48 hrs	∼31 hrs	—
Obs	3421	132	35	9	1
Number of trades	2625	67953	244209	940052	6874321

Table 12: Average price durations.

C	1 USD	2 USD	3 USD	5 USD	10 USD	20 USD
Bitstamp	0 min 30 sec	1 min 12 sec	2 min 22 sec	7 min 9 sec	41 min 20 sec	232 min 13 sec
Btc-e	0 min 14 sec	0 min 34 sec	1 min 19 sec	5 min 0 sec	29 min 41 sec	242 min 54 sec

Table 13: Parameter estimates for log-EACD(m,q) model for price durations with threshold C=3 USD. The model is specified as: $\psi_i = exp\left(\omega + \sum_{j=1}^{m} \alpha_j \left(\frac{x_{i-j}}{\psi_{i-j}}\right) + \sum_{j=1}^{q} \beta_j log(\psi_{i-j}) + y_1\zeta_{i-1} + y_2\upsilon_{i-1}\right)$, where ζ_{i-1} is the trading intensity in the previous price duration, and υ_{i-1} is the average volume per trade.

| | Bitstamp | | Btc-e | |
Parameters	Estimate	t-stats	Estimate	t-stats
ω	-0.0474	-9.72	-0.0136	-5.07
α_1	0.0858	21.84	0.1928	23.52
α_2			-0.1637	19.21
β_1	0.9468	186.56	1.3547	29.89
β_2			-0.3663	-8.33
$y1$	-0.0312	-12.46	-0.0054	-7.88
$y2$	-0.0103	-9.62	-0.0120	-5.31