NTNU
Norwegian University of
Science and Technology

The Determinants of Bitcoin Liquidity

Sondre Bergløff
Jacob Emil Tønnesen
Markus Øverli

Industrial Economics and Technology Management
Submission date: May 2018
Supervisor: Peter Molnar, IØT

Norwegian University of Science and Technology
Department of Industrial Economics and Technology Management

Problem Description

The purpose of this thesis is to study the liquidity in the bitcoin markets. As there is no central authority regulating bitcoin, part of the problem is to gather and structure quality data. We investigate both the determinants and the predictors of liquidity on bitcoin. We study whether the determinants and predictors are the same around the world, in different time-zones and trading in different currencies, or whether they differ.

Preface

This thesis is written as the concluding part of the Industrial Economics and Technology Management Master's programme at the Norwegian University of Science and Technology (NTNU) with a specialization in Financial Engineering.

We would like to thank our supervisor, Peter Molnár, for the discussions throughout the semester. His feedback and guidance have been very important to our work.

Trondheim, May 24, 2018.

Jacob Emil Tønnesen, Sondre Bergløff and Markus Øverli

Abstract

Liquidity is one of the most important characteristics of an asset. While it has been studied extensively in conventional markets, little research has been done on the liquidity of bitcoin markets. We investigate determinants of liquidity in the bitcoin markets, both on an hourly and a daily basis. As a measure of liquidity, we use the bid-ask spread, calculated from high-frequency data from four different exchanges located around the world. We find that contemporaneous traded volume and volatility are positively related with the bid-ask spread. We also find that high absolute returns predict high bid-ask spread in the next period. Our findings indicate that bitcoin market makers tend to increase the bid-ask spread in more uncertain times and that higher traded volume can be interpreted as new information arriving in the market.

Sammendrag

Likviditet er av stor betydning, og har blitt inngående studert i konvensjonelle markeder. For bitcoinmarkedene har forskningen på likviditet hittil vært lite omfattende. Ved å bruke forskjellen på kjøps- og salgskurser som mål på likviditet, undersøker vi determinanter for likviditet, både på times- og dagsbasis. Tidsseriene våre baserer seg på flere år med høyfrekvent data fra bitcoinbørser lokalisert i ulike deler av verden. Vi finner at handelsvolum og volatilitet har negativ sammenheng med samtidig likviditet, og at høy absolutt avkastning predikerer lav likviditet i neste periode. Funnene våre indikerer at likviditetstilbydere (market makers) øker differansen mellom kjøps- og salgsordre i perioder med høy usikkerhet, og at stort handelsvolum kan tolkes som at informasjon ankommer markedene.

Contents

1 Introduction

Bitcoin, a decentralized, open-source electronic cash system, is a rapidly developing market with characteristics that are not well understood. However, its popularity is growing tremendously. In late 2013, the total market capitalization surpassed 13 billion USD, having gone from zero at the outset in 2008. In November 2017, the market capitalization hit 300 billion USD, driven by 24-hour volumes of more than 20 billion USD. At the same time, two of the worlds largest futures exchanges (CBOE and CME) competed to be the first to offer bitcoin futures. A few months later, in February 2018, around two-thirds of the all-time high market capitalization had vanished.

Despite increased interest in bitcoin, many advocate caution going forward, dismissing the asset as a speculative bubble and a fraud (Coggan 2017; B. Harris 2018). Proponents of the currency, on the other hand, argue that the full potential is yet to be revealed and that the currency may lead to a paradigm shift in finance and trade (Patel 2017; Rees 2014; Johnson 2018). Either way, understanding the financial drivers and mechanisms of bitcoin will be crucial going forward, as the growth in popularity may attract considerable money from novice traders, private investors, and eventually, larger pension and hedge funds.

Liquidity is an important topic for research as liquidity characteristics severely influence how markets behave. During the financial crisis of 2007-2008, for example, liquidity played a paramount role. Liquid markets are seen to allocate resources more efficiently by enhancing the flow of information through low-cost transfer (Lybek and Sarr 2002). Furthermore, investors perceive liquid assets as more favorable, due to both lower transaction costs and discretion in trading. Amihud and Mendelson (1986) document a highly significant and positive cross-sectional effect of the bid-ask spread on returns in NYSE stocks during the period

1961-1980, indicating a liquidity premium. Brennan and Subrahmanyam (1996) conclude similarly by analyzing intraday cross-sectional data from NYSE/AMEX in the years of 1984 and 1988. Amihud (2002) documents the same effect in a time-series analysis. Acharya and Pedersen (2005) develop an equilibrium model with liquidity risk, showing that an asset's required return depends on its expected liquidity. Chordia et al. (2000) further finds that liquidity in single assets is influenced by market-wide liquidity, thus indicating commonality in liquidity across assets. Following this insight, Pástor and Stambaugh (2003) observe that an asset's sensitivity to market-wide liquidity is strongly related to its expected return.

This paper investigates the bid-ask spread of bitcoin. The bid-ask spread is a measure of transaction cost. In his seminal paper on transaction costs, Demsetz (1968) defines bid-ask spread as *"the cost of using [an exchange] to accomplish a quick exchange of stock for money"*. Since Demsetz (1968), several studies have examined liquidity in relation to other variables, both empirically and theoretically. Demsetz (1968) proposes a theory of the economics behind bid-ask spread by analyzing immediacy in the setting of supply and demand. In his theory, which lays the foundation of much subsequent research, certain actors stand ready to transact in the marketplace. This provision of immediacy incurs a waiting cost borne by the supplier; the supplier must assume risk and allocate funds. Put differently, the supplier of immediacy must hold inventory, which is costly. More recent literature refers to this as *inventory cost*. The bid-ask spread is the markup required to cover this cost. In an actively traded asset, the frequency of transactions will be high, yielding on average shorter waiting times. Based on these insights, Demsetz argues that the *time rate of transactions* should be inversely related to bid-ask spread, all else equal. He also proposes competition as an influencing force on the bid-ask spread. The argument, in short, states that intra- and intermarket competition maintains the bid-ask spread close to actual inventory cost; in less competitive markets, the bid-ask spread might surpass the inventory cost. Demsetz (1968) concludes with an empirical investigation of cross-sectional data on the costs of transacting on NYSE in 1965, and finds evidence of a significant negative effect from trading activity on bid-ask spread, supporting his theory. A large body of literature confirms this finding using cross-sectional data, in addition to supporting the theory that competition is related to bid-ask spread (Tinic 1972; Tinic and West 1972; Hamilton 1976). Benston and Hagerman (1974) and Branch and Freed (1977) extend the research by including measures of risk in a cross-sectional analysis of bid-ask spread determinants. In addition to inverse relationships between bid-ask spread and trading activity, they find significant positive relationships between bid-ask spread and risk, supporting the inventory cost-theory.

In Stoll (1978), *information costs* and *order costs* are discussed as possible constituents of the bid-ask spread, in addition to the aforementioned inventory cost. Order costs are made up of specific costs incurred when transacting (labor, equipment, record keeping, and so on). These may be subject to economies of scale. Information costs arise when the market makers trade against investors equipped with superior information. Under the presence of asymmetric information, the market maker widens the spread to protect against losses. For example, an informed trader may sell at the bid, knowing that the price will decrease in the subsequent period. The market maker, having bought at a higher price, will not be able to recuperate the value after the price decrease unless the bid-ask spread was sufficiently wide. While Stoll's research is purely theoretical, empirical research agrees that the information cost component of the bid-ask spread is positive (Glosten and L. Harris 1988; Hasbrouck 1988).

Most of the early literature on transaction costs analyzed cross-sectional data over relatively short time periods, typically due to limited data availability and data processing capabilities. Advances in these areas have enabled more extensive research into the time-series features and dependencies of transaction costs. The findings from the cross-sectional research are not necessarily valid on a time-series basis, as explicit costs, inventory costs, and information costs may impact liquidity differently over time than across assets. Lee et al. (1993), for example, suggest exactly this. By analyzing 1988 intraday data of 230 NYSE firms, they show that bid-ask spread is positively related to trading volume, concluding that the "*... liquidity providers are sensitive to changes in information asymmetry risk and actively manage this risk by using both spreads and depths*". Other studies also find a positive relationship between trading activity and bid-ask spreads (Daníelsson and Payne 2012; Narayan et al. 2015). Contrary to these findings, Ding (1999), McInish and Wood (1992) and Chordia et al. (2001) find that the bid-ask spread is significantly negatively related to trading activity. A positive relationship between volatility and bid-ask spreads is found in the time-series context, in line with the inventory cost theory (McInish and Wood 1992; Ding 1999; Galati 2000). Still, others find a negative relation between volatility and bid-ask spreads (Narayan et al. 2015).

Collectively, previous literature on transaction costs have shown, both theoretically and empirically, the following:

- The bid-ask spread is made up of inventory costs, information costs, and order costs. Inventory costs are related to the cost incurred by the provision of immediacy and should be impacted by the level of trading activity and volatility. Information costs arise due to asymmetric information, where a market maker incurs losses when trading against informed investors. Order costs are explicit costs related to the provision of immediacy.
- The cross-sectional studies generally agree that trading activity and bid-ask spreads are negatively related, whereas volatility and bid-ask spreads are

positively related. These findings are best explained in the framework of inventory costs.

- The time-series literature is inconclusive concerning the relationship between bid-ask spread and trading activity. Results that indicate a positive relationship between trading activity and bid-ask spread are best explained by information costs, where increased trading activity is a signal of information coming to the market.
- In the time-series literature, the bulk of the evidence indicates a positive relationship between volatility and bid-ask spreads, with some exceptions. This is predicted in the inventory cost framework.

We undertake an empirical investigation of the time-series determinants of liquidity in bitcoin using the bid-ask spread as a measure of liquidity. As an estimate of the bid-ask spread, we apply the estimator proposed in Roll (1984). Goyenko et al. (2009) study widely used proxies of liquidity in financial literature, and conclude that common liquidity proxies indeed measure liquidity, including Roll's measure.

When analyzing the bitcoin markets in light of previous literature on conventional markets, it is important to bear in mind some distinguishing features of the bitcoin markets. For starters, bitcoin has no obvious fundamental value. There is for instance no quarterly reporting enabling investors to assess the value in a meaningful way. As a result, the most important piece of information for investors to act on is trading activity. Another important distinction of the bitcoin markets is that there are no formal market makers such as in traditional dealer markets. Still, liquidity traders, hoping to profit from price movements, act as informal market makers. As the mechanisms behind this form of liquidity provision are opaque, it is hard to analyze the bid-ask spread through the framework of inventory- and order costs.

We find that the bid-ask spread is positively contemporaneously related to traded volume and volatility. From the perspective of information cost theory, increased volume conveys to the informal market that the perceived market value of bitcoin is shifting. Thus, the market maker, not knowing the fundamental value, increases the bid-ask spread to protect against losses. The same logic applies to volatility and the bid-ask spread: In a more volatile market, the market maker charges larger bid-ask spreads to due larger expected fluctuations in price. It can also be explained by the inventory cost theory, where the market maker charges larger bid-ask spread as compensation for holding riskier inventory. Additionally, we find that absolute return is a predictor of the bid-ask spread. This means that market makers increase the bid-ask spread particularly after larger price changes, in line with the information cost theory.

Our work extends the existing literature on the financial aspects of bitcoin by addressing liquidity using high-frequency data over an extended period from several exchanges. For an overview of the economics of bitcoin, see Böhme et al. (2015). Balcilar et al. (2016) studies the relation between volume, returns, and volatility. Urquhart (2016) indicates inefficiency in the market. Several papers have investigated the hedging properties of bitcoin (Dyhrberg 2016; Bouri et al. 2017). Brandvold et al. (2015) examines the price discovery among different exchanges. Regarding liquidity, little research has been done. Donier and Bouchaud (2015) study how measures of liquidity predict market crashes. Dimpfl (2017) gives an overview of liquidity for a selection of major bitcoin exchanges, but with a limited dataset. Donier and Bonart (2017) study the market impact of meta-orders on the Mt. Gox exchange. Kim (2017) compares the transaction cost of changing currencies via bitcoin with that of traditional foreign exchange markets. Easley et al. (2017) conduct both theoretical and empirical investigations into the role of transaction fees on the blockchain itself.

The remainder of this paper is organized as follows. In Section 2, we introduce the data used and any transformations conducted. We also explain the metrics applied. In Section 3 we explain the methods used in the analyses. In Section 4 we present the results and discuss the implications. Finally, in Section 5, we summarize our results and main findings.

2 Data

In order to exchange conventional currencies for bitcoin, one can trade on any of the numerous trading platforms, hereafter referred to as *exchanges*. These exchanges are intermediaries between conventional money and bitcoin but have no direct connection to the bitcoin system. We argue that aggregated high-frequency data from several exchanges is improper to use in a liquidity analysis. Temporary price differences between exchanges occur from time to time, and the bid-ask spread at two exchanges could at a given time be the same, but the mean price could be different[1]. Aggregating the data could lead to incorrect bid-ask spread estimates, depending on how the data is aggregated. For example, if a transaction at the ask price on the high mean exchange is followed by a transaction at the bid price on the low mean exchange, the spread would be overestimated by the size of the price difference of the two exchanges. Therefore, we analyze the different exchanges separately.

[1]This effect is enhanced by strict currency restrictions in Korea and China, limiting arbitrage trading between the exchanges.

2.1 Data source

We obtain data from the website www.bitcoincharts.com[2]. We compare data from other sources with our data and find no deviation in reported price or volume.

The dataset consists of opening and closing price for each minute, as well as traded volume, for six bitcoin exchanges, namely Bitstamp, Coinbase, BTCN, Korbit, Coincheck and Kraken[3]. We only conduct analyses on four out of these six exchanges. See Table 1 for an overview of the included exchanges. The choice of exchanges is driven by data quality and by a desire for having time-zones and currency-pairs from around the world represented. The rationale for this is that we wish to eliminate local anomalies as explanations for observed patterns and phenomena. Specifically, Kraken is not included as a European exchange because the data span a shorter time period than Bitstamp, which is also European. The Japanese exchange Coincheck is excluded because it has worse data quality than Korbit, which is located in the same time-zone. Bitstamp and Coinbase, located in Slovenia and the United States, respectively, are referred to as the *western* exchanges, while BTCN and Korbit, located in China and Korea are referred to as the *eastern* exchanges.

Though the bitcoin price behaves similarly on all these exchanges, this is not necessarily the case for the bid-ask spread and volume. Table 2 and Table 3 show the correlation of hourly bid-ask spread and volume respectively, between the

[2]There are several websites reporting prices of cryptocurrencies, such as www.coinmarketcap.com, www.blockchain.info, and www.bitcoincharts.com. These sources typically aggregate data from multiple cryptocurrency exchanges, thus reflecting a consensus of the price and volume.

[3]For an overview of bitcoin exchanges, see www.bitcoin.org/en/exchanges

Table 1: Overview of exchanges used in analyses

Exchange name	Country	Time-zone	Currency	First day of data	Last day of data
Bitstamp	Slovenia	UTC+1	USD	01.01.2012	31.12.2017
Coinbase	US	UTC−5	USD	02.12.2014	31.12.2017
BTCN	China	UTC+8	CNY	01.01.2012	29.09.2017
Korbit	Korea	UTC+9	KRW	01.09.2013	31.12.2017

exchanges[4]. The correlation is low for the bid-ask spread, and at times negative for volume.

The focus of this paper is not to compare the magnitude of bid-ask spreads of different exchanges. Hence it is not essential for the data to span the same period. Instead, we choose to use as much of the available data as practically feasible.

We recognize that the exchanges do not necessarily adhere to any regulatory regime, and as a consequence, the data may be compromised with, for example, fake volumes (Anonymous 2017; Popper 2017). Also, periods of no transaction cost, such as in China before January 2017, may have lead to abnormally large trading activity (Rizzo 2017)

[4] These correlations are calculated on a global time basis, meaning that the observations are simultaneous. The negative correlation between volume on western and eastern exchanges is undoubtedly explained, in part, by different trading volume during the day and during the night.

Table 2: Correlation of hourly bid-ask spread between exchanges. Only hours with data for all exchanges are included.

	Bitstamp	Coinbase	BTCN
Coinbase	24.1%	–	–
BTCN	11.1%	8.5%	–
Korbit	24.2%	23.6%	24.1%

Table 3: Correlation of hourly volumes between exchanges. Only hours with data for all exchanges are included.

	Bitstamp	Coinbase	BTCN
Coinbase	32.9%	–	–
BTCN	−2.4%	−6.5%	–
Korbit	13.2%	37.0%	−3.4%

2.2 Variables

In the following, we introduce all variables used in the analyses. We create both hourly and daily time-series of these variables from one-minute data.

Return

Throughout the thesis, we use the absolute value of the difference in the natural logarithm of prices as the return measure, denoted $|r|$:

$$|r_t| = |\ln(P_{t+1}) - \ln(P_t)| \tag{1}$$

Here, t is the time index, and P_t is the closing price of time period t.

Volume

Traded volume V, denominated in bitcoins (BTC), is non-stationary. To achieve better statistical properties, we transform traded volume into a normalized, stationary series by taking the natural logarithm of the ratio between the volume at a given time interval and the moving average volume over the previous year:

$$v_t = \ln(V_t) - \ln\left(\frac{\sum_{i=(t-N)}^{t-1} V_i}{N}\right) \tag{2}$$

This means that some of the available data, typically one year, is solely used as a basis of standardization, and is not used in the actual analysis. The standardized volume, denoted v, can be interpreted as the deviation from the mean volume. Equation 2 shows the volume transformation used in the analyses. N is the number of time intervals in the moving window on the basis of which the normalization is conducted (i.e. 365 for daily volume and 365×24 for hourly volume).

In addition to the exchange-specific traded volume (i.e., the traded volume on a single exchange), it is interesting to examine the total traded volume globally (i.e., the volume from all existing exchanges). *Daily* global volumes are available from www.bitcoinity.org[5]. Because this data is unavailable on a minute or hour frequency, we use the total volume from five of the six exchanges in our dataset. We exclude BTCN from this index due to abnormally high volumes during 2016, which skewed the index and led to diminished correlation with actual global volume. The included exchanges account for approximately 19% of global volume in the period 2014 to 2017[6]. See Table 4 for a volume breakdown of the global volume index. To ascertain whether this index is a robust proxy for true global volume, we calculate the correlation between the index and the actual values on a daily basis and find it to be 88.1%. Figure 1 shows daily global volume (1a) and the index (1b). The global volume index is standardized in the same manner as the local volume (see Equation 2), and is denoted gv.

Realized volatility

We use the sum of squared returns to measure realized volatility:

$$RV_t = \sqrt{\sum_{i=1}^{M} r_{it}^2} \tag{3}$$

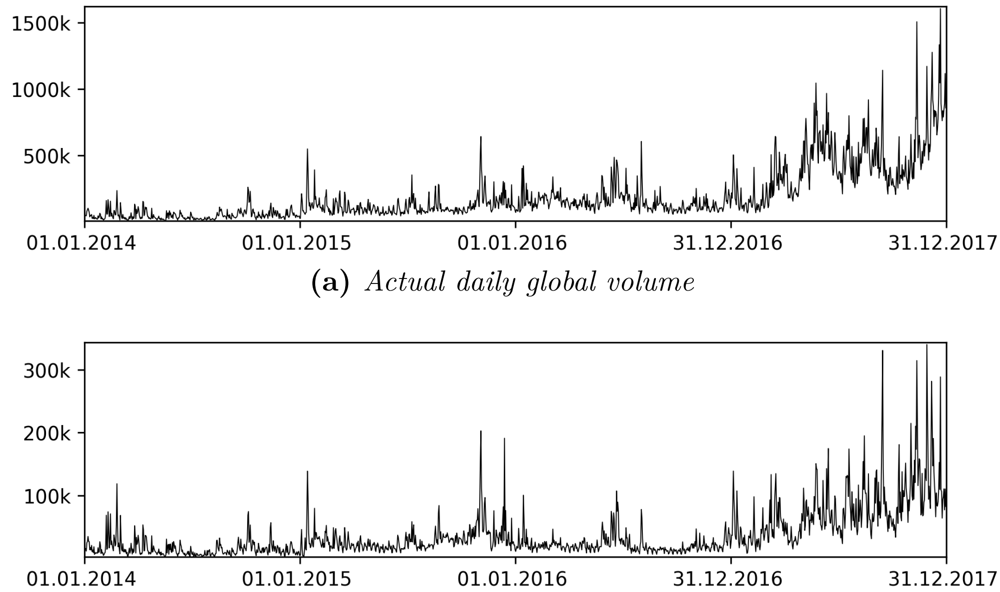
The accuracy of the measure directly depends on the sampling frequency, and one has to trade off information against microstructure noise (Andersen and Bollerslev

[5]bitcoinity.org is a website providing data aggregated from several exchanges

[6]The global bitcoin trading volume is dispersed among various exchanges. Although the exchanges included in the index are among the most significant exchanges, they still represent only a minority of total global volume.

Table 4: Relative size of exchanges included in the global volume index

Exchange	Share of index volume
Bitstamp	39.3%
Coincheck	23.2%
Coinbase	19.0%
Kraken	13.5%
Korbit	5.0%



(b) *Daily global volume index (proxy)*

Figure 1: Comparison of actual global volumes and global volume index

1998). In Equation 3, M is the number of equally spaced time intervals. In Figure 2 we plot median daily realized volatility against sampling intervals. We see that micro noise stabilizes around the 15-min sampling interval. Hence, we use a 15-minute sampling interval in our analyses. To achieve better statistical properties, we use the natural logarithm of realized volatility and denote it rv.

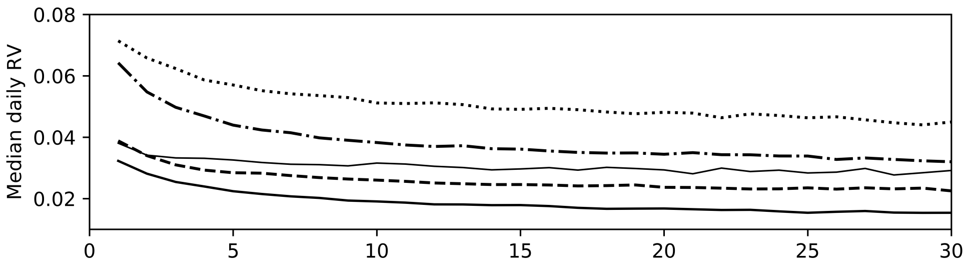


Figure 2: Bitstamp median daily realized volatility for different sampling frequencies, calculated separately for different time periods. The different graphs represent, from top to bottom, the following time periods: 2013, 2014, Jan-June 2017, 2015, 2016.

Bid-ask spread

To measure liquidity, we use Roll's estimator for the bid-ask spread, introduced in Roll (1984):

$$BAS_t = 2\sqrt{\frac{-\sum_{i=1}^{T} \Delta P_{i,t} \Delta P_{i-1,t}}{T-1}} \tag{4}$$

Here, ΔP is the price difference between two consecutive trades, and T is the number of minute pairs during the period in question. The method relies on the assumption of an efficient market where the price of an asset fluctuates between the bid and the ask, given there is no new information. The calculation utilizes the serial covariance of minute prices to calculate an implied bid-ask spread. Roll's estimator depends on negative covariances. We set the spread to zero if the covariance is positive, as in Corwin and Schultz (2012). To calculate relative spreads, which we denote bas, we divide the nominal spread by the closing price of the time interval it is calculated for (i.e., the hour for hourly calculations and the day for daily). We use the relative spread throughout the paper, as the nominal spread would be greatly affected by the price level.

Though Roll's estimator for the bid-ask spread is widely used, its accuracy is, of course, inferior to what could be achieved by studying the actual order book. Unfortunately, we were unable to acquire such data.

2.3 Description of data

From the inception of bitcoin, there have been major changes and events causing extreme observations in price and volume. We exclude such outliers from our analyses. To allow for logarithmic transformations, we remove all data corresponding to a time interval when either volume, bid-ask spread or realized volatility is zero. Table 5 and 7 present properties before transformations of the hourly and daily data, respectively. Table 6 and 8 show the same properties of hourly and daily data, but after transformations have been undertaken. The bid-ask spread displays different properties across the exchanges. The mean bid-ask spread, for example, is more than twice as high on Bitstamp than on Coinbase. The distributions are also dissimilar in terms of skewness, standard deviation, and kurtosis. The volumes and volatilities are no less different. Part of the explanation may be the different sizes of the exchanges or their different geographical location. Undoubtedly, the fact that they span different time-periods plays an important role as well. During the years 2013 and 2014, when Coinbase was yet to open, the bid-ask spread in the bitcoin markets were, on average, higher than in the following years. See Figure 3 for a comparison of the bid-ask spread over time for the four exchanges. Nevertheless, as is evident from the figure, Bitstamp has had a higher bid-ask spread than the three other exchanges almost the entire period. We do not pursue an explanation as to why this is the case.

Substantial autocorrelation is present for volume, volatility and bid-ask spread for all the four exchanges. In Figure 4 and Figure 5 we include scatter plots of the daily bid-ask spread against absolute returns, volume, realized volatility and lagged values of spread for all four exchanges. Positive correlation between spread and volatility is visible, and so is the positive autocorrelation.

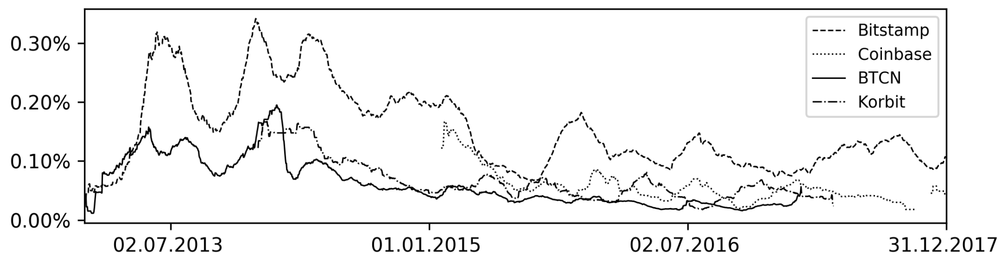


Figure 3: Comparison of daily bid-ask spread of the four exchanges over time. Coinbase and Korbit start operations in the course of the displayed time-period.

Figure 6 shows the development in trading volume for each exchange. We observe that the volume on the western exchanges stays fairly stable over our sample period, while the volume on the eastern exchanges increases significantly in trading volume from late 2016 to the end of the dataset (excluding extreme observations). Figure 7 shows the development of bid-ask spread for each individual exchange (before removal of extreme values).

Table 5: Summary statistics of variables before transformations on an hourly basis. r^H, V^H, RV^H and bas^H are the hourly returns, volume, realized volatility and bid-ask spread, respectively. μ, σ, AC_1 and AC_{10} are the mean, standard deviation, order 1 and order 10 autocorrelation coefficients. Note that data for exchanges span different time-periods.

	Variable	Min	Max	μ	σ	Kurtosis	Skewness	AC_1	AC_{10}	n
Bitstamp	r^H	-9.66%	9.05%	0.01%	0.80%	14.66	-0.513	-0.026	0.005	35 251
	V^H	1.016	9 575	426.1	518.6	29.05	4.000	0.479	0.221	35 251
	RV^H	0.11%	199.86%	54.17%	39.03%	1.387	1.277	0.537	0.412	35 251
	bas^H	0.00%	0.98%	0.15%	0.12%	5.645	1.950	0.544	0.402	35 251
Coinbase	r^H	-7.49%	6.74%	0.01%	0.60%	20.16	-0.328	0.004	0.023	16 779
	V^H	8.349	5 536	310.6	273.2	41.48	4.531	0.628	0.306	16 779
	RV^H	0.02%	199.86%	30.65%	29.35%	6.173	2.205	0.598	0.452	16 779
	bas^H	0.00%	1.35%	0.06%	0.06%	39.86	4.202	0.505	0.338	16 779
BTCN	r^H	-7.26%	7.35%	0.02%	0.72%	15.30	-0.335	-0.038	0.008	25 852
	V^H	0.143	845 724	5 892	24 530	190.4	10.84	0.780	0.461	25 852
	RV^H	0.00%	199.60%	43.03%	39.31%	2.114	1.542	0.608	0.503	25 852
	bas^H	0.00%	2.24%	0.12%	0.16%	10.05	2.737	0.730	0.619	25 852
Korbit	r^H	-7.21%	5.87%	0.01%	0.72%	12.20	-0.489	0.006	-0.033	10 805
	V^H	0.060	1 346	116.3	124.8	7.186	2.179	0.716	0.510	10 805
	RV^H	0.00%	199.99%	44.04%	36.69%	2.704	1.625	0.469	0.324	10 805
	bas^H	0.00%	1.06%	0.08%	0.08%	20.63	3.374	0.325	0.222	10 805

Table 6: Summary statistics of transformed variables on an hourly basis. $|r^H|$, v^H, rv^H and bas^H are the hourly absolute returns, transformed volume, log-realized volatility and bid-ask spread, respectively. μ, σ, AC_1 and AC_{10} are the mean, standard deviation, order 1 and order 10 autocorrelation coefficients. Note that data for exchanges span different time-periods.

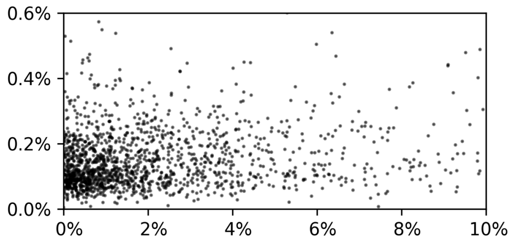
	Variable	Min	Max	μ	σ	Kurtosis	Skewness	AC_1	AC_{10}	n
Bitstamp	$\lvert r^H \rvert$	0.00%	9.66%	0.48%	0.64%	25.05	3.864	0.295	0.220	35 251
	v^H	-5.534	2.951	-0.389	0.949	0.348	-0.255	0.511	0.144	35 251
	rv^H	-6.794	0.692	-0.890	0.816	3.083	-0.967	0.499	0.377	35 251
	bas^H	0.00%	0.98%	0.15%	0.12%	5.645	1.950	0.544	0.402	35 251
Coinbase	$\lvert r^H \rvert$	0.00%	7.49%	0.34%	0.50%	30.90	4.401	0.313	0.223	16 779
	v^H	-3.007	3.449	-0.149	0.632	0.657	-0.042	0.633	0.124	16 779
	rv^H	-8.735	0.692	-1.582	0.945	1.277	-0.494	0.588	0.446	16 779
	bas^H	0.00%	1.35%	0.06%	0.06%	39.86	4.202	0.505	0.338	16 779
BTCN	$\lvert r^H \rvert$	0.00%	7.35%	0.41%	0.59%	22.95	3.865	0.314	0.219	25 852
	v^H	-7.728	4.190	-0.365	1.084	2.801	-1.166	0.713	0.352	25 852
	rv^H	-10.362	0.691	-1.286	1.082	7.500	-1.497	0.634	0.533	25 852
	bas^H	0.00%	2.24%	0.12%	0.16%	10.05	2.737	0.730	0.619	25 852
Korbit	$\lvert r^H \rvert$	0.00%	7.21%	0.42%	0.58%	18.96	3.508	0.295	0.166	10 805
	v^H	-6.660	2.361	-0.261	0.867	2.583	-0.848	0.369	0.123	10 805
	rv^H	-22.396	0.693	-1.168	0.943	30.35	-2.226	0.387	0.253	10 805
	bas^H	0.00%	1.06%	0.08%	0.08%	20.63	3.374	0.325	0.222	10 805

Table 7: Summary statistics of variables before transformations on a daily basis. r^D, V^D, RV^D and bas^D are the daily returns, volume, realized volatility and bid-ask spread, respectively. μ, σ, AC_1 and AC_{10} are the mean, standard deviation, order 1 and order 10 autocorrelation coefficients. Note that data for exchanges span different time-periods.

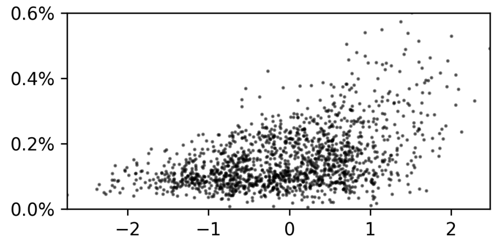
	Variable	Min	Max	μ	σ	Kurtosis	Skewness	AC_1	AC_{10}	n
bitstamp	r^D	-9.93%	9.86%	0.41%	3.00%	1.580	-0.004	0.032	0.081	1 545
	V^D	644.0	51 735	10 573	7 819	2.706	1.515	0.611	0.359	1 545
	RV^D	14.67%	199.59%	66.11%	37.05%	1.345	1.242	0.657	0.345	1 545
	bas^D	0.00%	0.67%	0.15%	0.09%	3.170	1.515	0.731	0.544	1 545
coinbase	r^D	-9.81%	9.91%	0.35%	2.50%	2.754	-0.103	0.020	0.082	677
	V^D	1 622	33 670	7 421	4 063	9.176	2.437	0.583	0.223	677
	RV^D	7.01%	180.70%	39.55%	27.72%	4.444	1.845	0.634	0.212	677
	bas^D	0.00%	0.33%	0.06%	0.04%	6.202	2.087	0.545	0.287	677
btcn	r^D	-9.65%	9.89%	0.31%	2.82%	1.864	0.112	0.085	0.053	1 026
	V^D	126.8	5 126 824	156 246	498 181	33.03	5.374	0.844	0.660	1 026
	RV^D	5.76%	195.11%	50.20%	36.71%	2.071	1.482	0.699	0.516	1 026
	bas^D	0.00%	0.52%	0.06%	0.05%	11.13	2.686	0.548	0.416	1 026
korbit	r^D	-9.89%	9.55%	0.23%	2.66%	2.365	-0.076	0.039	0.032	815
	V^D	26.10	13 095	1 624	2 110	5.154	2.246	0.859	0.740	815
	RV^D	8.17%	198.23%	55.60%	36.94%	1.961	1.407	0.556	0.312	815
	bas^D	0.00%	0.41%	0.05%	0.05%	6.152	2.061	0.502	0.262	815

Table 8: Summary statistics of transformed variables on a daily basis. $|r^D|$, v^D, rv^D and bas^D are the daily absolute returns, transformed volume, log-realized volatility and bid-ask spread, respectively. μ, σ, AC_1 and AC_{10} are the mean, standard deviation, order 1 and order 10 autocorrelation coefficients. Note that data for exchanges span different time-periods.

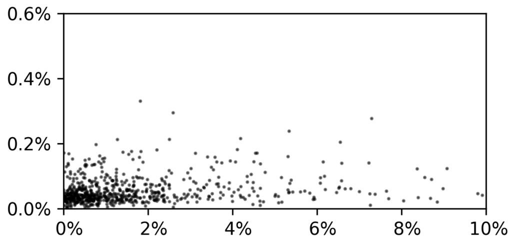
	Variable	Min	Max	μ	σ	Kurtosis	Skewness	AC_1	AC_{10}	n
bitstamp	$\lvert r\rvert^D$	0.00%	9.93%	2.12%	2.17%	1.776	1.499	0.185	0.133	1 545
	v^D	-2.747	2.479	-0.037	0.857	-0.357	-0.102	0.748	0.556	1 545
	rv^D	-1.919	0.691	-0.558	0.537	-0.588	0.088	0.739	0.437	1 545
	bas^D	0.00%	0.67%	0.15%	0.09%	3.170	1.515	0.731	0.544	1 545
coinbase	$\lvert r\rvert^D$	0.00%	9.91%	1.70%	1.86%	3.400	1.851	0.214	0.147	677
	v^D	-1.362	4.456	0.356	0.960	2.676	1.470	0.920	0.831	677
	rv^D	-2.657	0.592	-1.135	0.640	-0.421	0.123	0.727	0.335	677
	bas^D	0.00%	0.33%	0.06%	0.04%	6.202	2.087	0.545	0.287	677
btcn	$\lvert r\rvert^D$	0.00%	9.89%	1.93%	2.08%	2.013	1.592	0.221	0.064	1 026
	v^D	-2.186	3.833	0.352	1.228	-0.830	0.094	0.921	0.789	1 026
	rv^D	-2.855	0.668	-0.934	0.708	-0.497	-0.020	0.797	0.559	1 026
	bas^D	0.00%	0.52%	0.06%	0.05%	11.13	2.686	0.548	0.416	1 026
korbit	$\lvert r\rvert^D$	0.00%	9.89%	1.81%	1.97%	2.859	1.749	0.206	0.062	815
	v^D	-2.216	2.136	0.369	0.663	0.355	-0.221	0.616	0.439	815
	rv^D	-2.504	0.684	-0.790	0.644	-0.495	-0.012	0.707	0.432	815
	bas^D	0.00%	0.41%	0.05%	0.05%	6.152	2.061	0.502	0.262	815



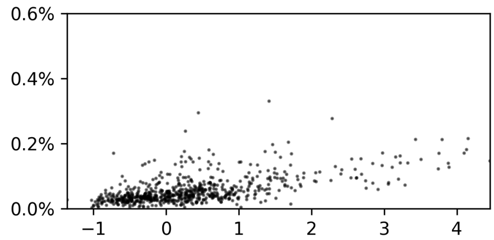
(a) *Bitstamp, spread and absolute return*



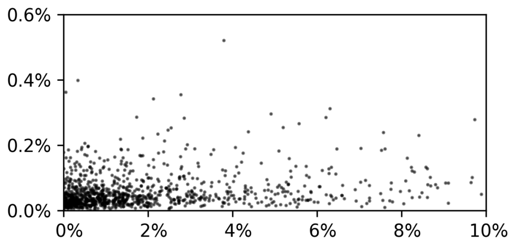
(b) *Bitstamp, spread and volume*



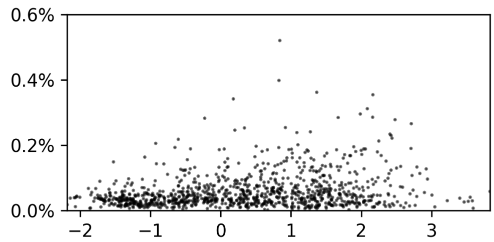
(c) *Coinbase, spread and absolute return*



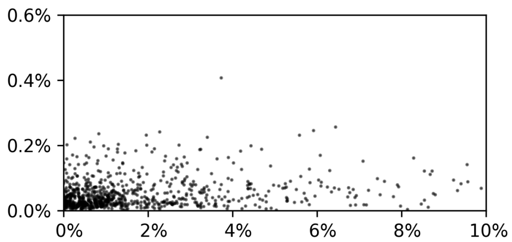
(d) *Coinbase, spread and volume*



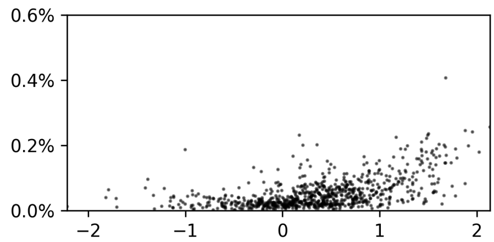
(e) *BTCN, spread and absolute return*



(f) *BTCN, spread and volume*

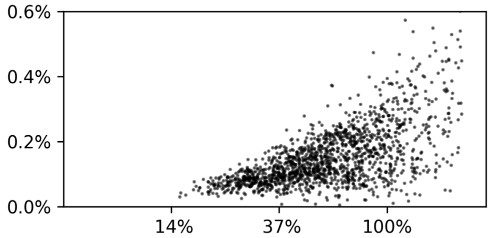


(g) *Korbit, spread and absolute return*

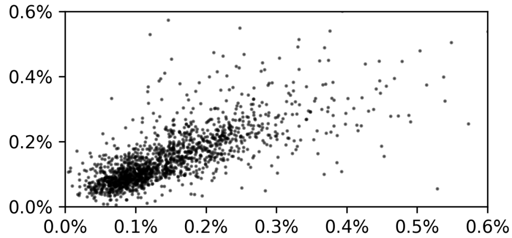


(h) *Korbit, spread and volume*

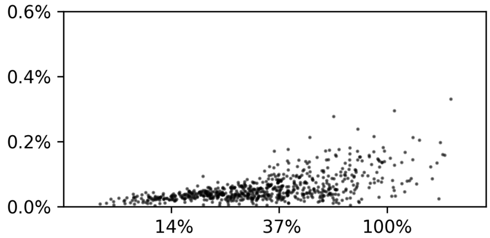
Figure 4: Daily bid-ask spread (y-axis) plotted against absolute return and standardized volaume (x-axis).



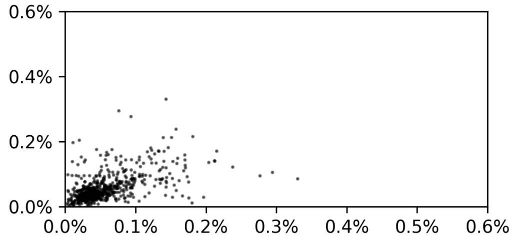
(a) *Bitstamp, spread and volatility*



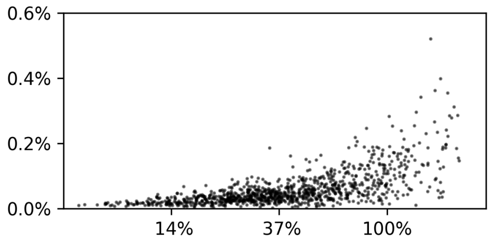
(b) *Bitstamp, spread and lagged spread*



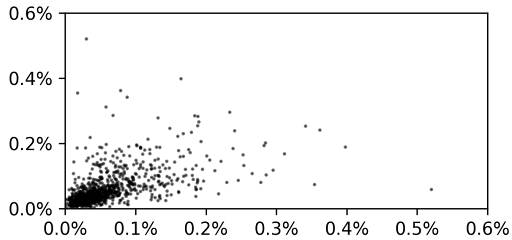
(c) *Coinbase, spread and volatility*



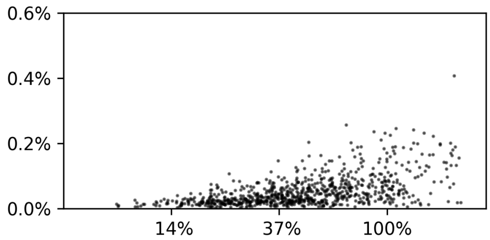
(d) *Coinbase, spread and lagged spread*



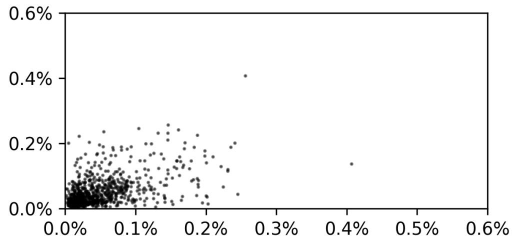
(e) *BTCN, spread and volatility*



(f) *BTCN, spread and lagged spread*

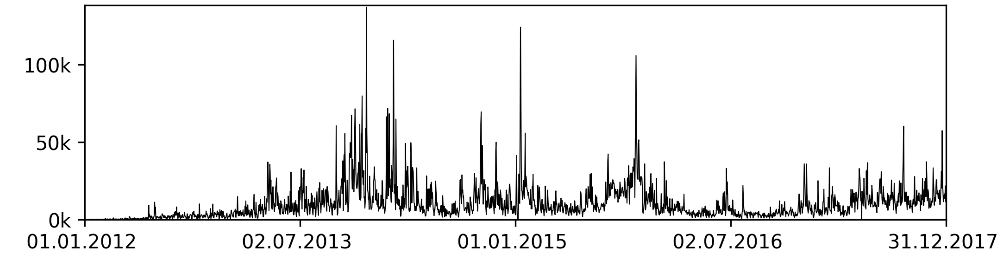


(g) *Korbit, spread and volatility*

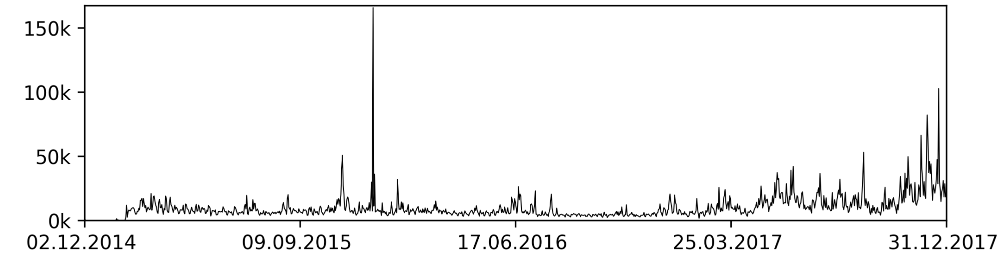


(h) *Korbit, spread and lagged spread*

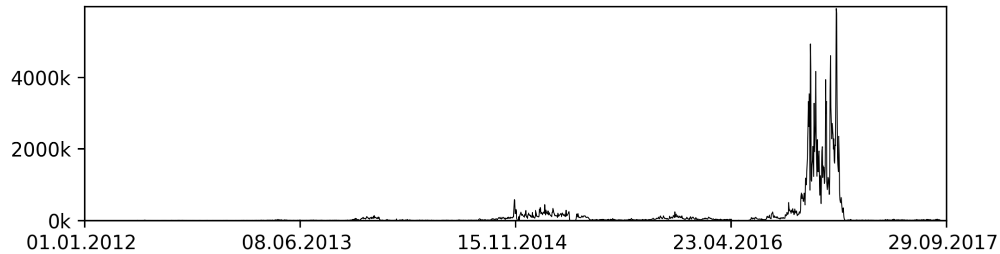
Figure 5: Daily bid-ask spread (y-axis) plotted against volatility and lagged spread (x-axis). Realized volatility shown on logarithmic scale



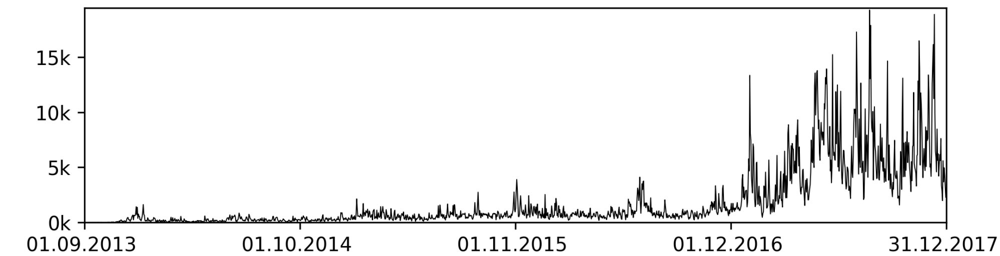
(a) *Bitstamp*



(b) *Coinbase*

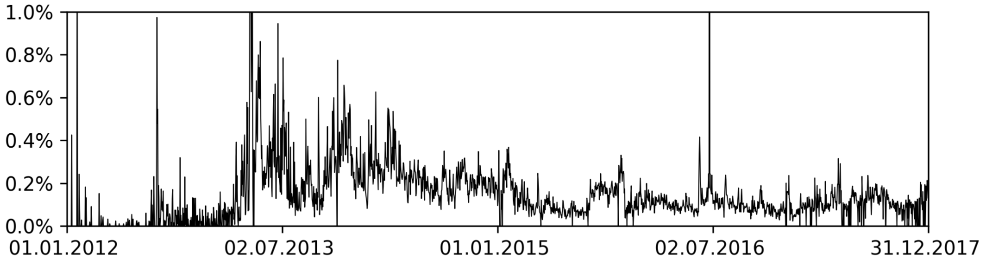


(c) *BTCN*

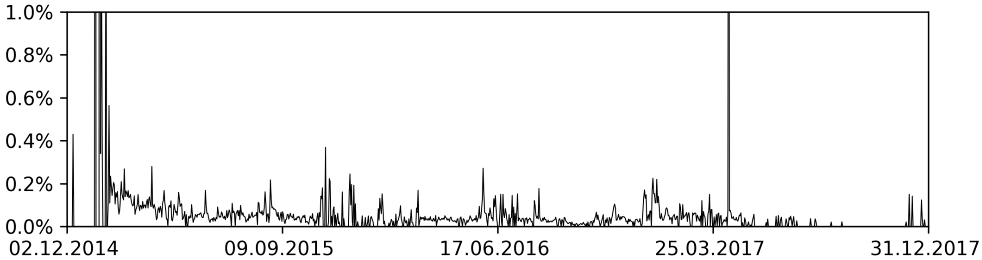


(d) *Korbit*

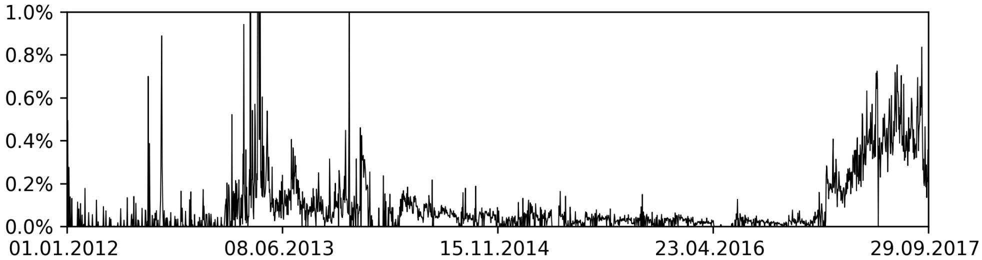
Figure 6: Development of daily traded volume over time for exchanges used in analysis, denominated in bitcoin. Extreme observations shown here are not included in further analysis.



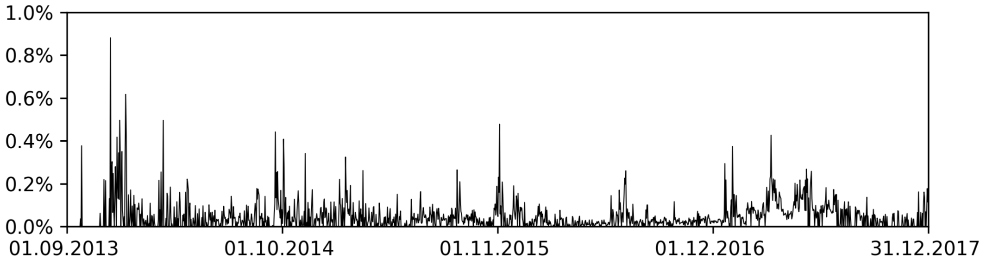
(a) *Bitstamp*



(b) *Coinbase*



(c) *BTCN*



(d) *Korbit*

Figure 7: Development of daily bid-ask spread over time for exchanges used in analysis. Extreme observations shown here are not included in further analysis.

3 Methodology

We study which variables are contemporaneously related to bid-ask spread, and which are predictive of the bid-ask spread. These multiple linear regression analyses are referred to as *determinants* analyses. Absolute return, exchange-specific volume, global volume, realized volatility and time-based dummy variables are included as explanatory variables. The regressions conducted are on the form:

$$bas_t = \beta_1 |r_t| + \beta_2 v_t + \beta_3 gv_t + \beta_4 rv_t + controls + \epsilon_t \tag{5}$$

$$bas_t = \beta_1 |r_{t-1}| + \beta_2 v_{t-1} + \beta_3 gv_{t-1} + \beta_4 rv_{t-1} + controls + \epsilon_t \tag{6}$$

Equation 5 and 6 summarize the contemporaneous and predictive models, respectively. To take into account the autocorrelation and possible seasonality of the bid-ask spread, *controls* are included in the regression models. These controls consist of dummy variables for time and autoregressive terms of the dependent variable. For the analyses on the hourly bid-ask spread, t refers to the present hour. For the analyses on daily bid-ask spread, t refers to the present day. The equations are estimated independently for each exchange. To make the regression coefficients more comparable, all coefficients in the regression tables are standardized[7].

The analyses are conducted both on an hourly and daily basis. The reason for doing this is to test the robustness of the conclusions and to see whether the patterns and correlations that exist across hours exist across days as well.

Dummy variables for time allow us to investigate whether seasonality is a significant determinant of bitcoin liquidity. For the analyses that are on an hourly basis (referred to as *intraday seasonality*), these dummy variables could represent

[7] They are standardized in such a way that a β coefficient in the regression means that a one standard deviation increase in the explanatory variable is associated with a β standard deviations increase in the bid-ask spread.

either one or more hours each. In order not to overparameterize the model, we let each dummy variable represent three hours[8].

For the analyses on daily variables, the most likely seasonal pattern is over the course of a week (referred to as *intraweek seasonality*). There can either be dummy variables for days of the week, or dummies indicating whether it is a weekend or not. We use day of the week dummies to be able to capture differences between for example Mondays and Wednesdays.

4 Results

In the following section, we present the results of our analyses. First, we establish a benchmark model, against which the determinants analyses can be compared, before we investigate the explanatory power of seasonality in the bid-ask spread. We then present the results from regressions of absolute returns, local volume, global volume, and realized volatility on the bid-ask spread.

4.1 Benchmark model

The determinants analyses require a benchmark model to see whether including a new variable adds explanatory value. Hence, we do not simply look for correlation among variables, but whether contemporaneous relations are significant beyond what the benchmark can explain. As discussed in Section 2.3 there is significant autocorrelation in the bid-ask spread. Hence, we construct the benchmark using autoregressive terms. These autoregressive terms can either be a standard Auto-Regressive (AR) model or a Heterogeneous Auto-Regressive (HAR) model.

[8] McInish and Wood (1992) use 30-minute intervals in their intraday analysis of bid-ask spread in equity markets. As NYSE is only open 6.5 hours per day, they require a much smaller number of dummy variables than what would be the case for the uninterrupted bitcoin markets.

The choice of which terms to include is based both upon explanatory power and brevity/simplicity. Through testing different configurations for these analyses, the HAR model is found superior, as it has both better explanatory power and fewer terms than the AR model. The flexibility of the HAR model enables us to capture longer-term variations with few variables[9]. To be able to compare the results from different exchanges, it is advantageous to use the same benchmark model for all the exchanges. It turns out that the benchmark model we decided upon is suitable for all the exchanges.

Table 9 and 10 summarize how the HAR terms are chosen for the analyses on hourly and daily bid-ask spread, respectively. The tables only show results for the Bitstamp exchange. As is evident from Table 9, when adding the 24-hour lagged bid-ask spread in model (2), the Akaike Information Criterion (AIC) improves. Its coefficient is, however, insignificant when adding the daily average in model (3). We thus remove it in model (4). The removal improves the AIC. Model (6) includes the previous hour value, daily average, and weekly average. It is the best-suited model for our purposes. However, for some practical purposes, like forecasting bid-ask spread with a limited data set, model (4) would be the most suitable benchmark since this very simple model performs almost equally to model (6). We have a substantial amount of data, spanning long time periods. AIC penalizes new terms relatively lightly when there are many observations. This would not be the case for a shorter data set. Note that results remain the same whether we use model (4) or (6) in our further analysis.

[9]Some hours and days are not included in the regressions, either due to flawed data or due to removed zero-values, as discussed in Section 2.3. For these instances, a choice must be made of how to treat the AR-terms. If a data point is missing, it will propagate into the following period's lagged terms. We choose to exclude all these data points from the analyses, rather than giving them a zero-value or moving an additional step backward. This causes a larger loss of data but avoids skewing the analyses.

The benchmark for the analysis on a daily basis is chosen similarly, as displayed in Table 10.

HAR benchmark Based on the findings in Table 9 and 10, we use the following models for regressions on hourly and daily bid-ask spread, respectively:

$$bas_t^H = \beta_1 bas_{t-1}^H + \beta_2 bas_{t-24,t-1}^H + \beta_3 bas_{t-7*24,t-1}^H \tag{7}$$

$$bas_t^D = \beta_1 bas_{t-1}^D + \beta_2 bas_{t-7,t-1}^D + \beta_3 bas_{t-60,t-1}^D \tag{8}$$

4.2 Seasonality as a determinant

Table 11 shows that when dummy variables for time-of-day are used as explanatory variables in the regression, they only explain a limited part of the variation of the bid-ask spread. This is not the case for the benchmark model (described in Section 4.1). It is evident that the benchmark model explains far more of the variation in the bid-ask spread while using fewer variables.

When the dummy variables are used in combination with the benchmark model, the R-squared only improves marginally. Therefore, in order not to overparameterize the model, we exclude the dummy variables from the main regressions.

We have shown that the explanatory power of seasonality on the bid-ask spread is low. However, there are distinct seasonal patterns in several variables. Figure 8 displays the intraday seasonality of volume, bid-ask spread, and volatility for each of the four exchanges. The plots show the average value for each variable over the course of the day, in the local time of each exchange. The values are averaged into three-hour periods. Some observations are apparent: volumes peak during typical working hours for all exchanges; the volumes on the western exchanges tend to decrease earlier than on the eastern exchanges; the bid-ask spread peaks in the middle of the work-day on the western exchanges, and in the middle of the night on the eastern exchanges.

Table 9: Choice of HAR benchmark for hourly bid-ask spread. bas^H_{t-1}, bas^H_{t-24}, $bas^H_{t-24,t-1}$, $bas^H_{t-7*24,t-1}$ are the bid-ask spread for the previous hour, the value 24 hours prior, the average over the last day and the average over the last week, respectively. Reported results are for Bitstamp. Results for other exchanges are very similar.

	Dependent variable: Hourly bid-ask spread					
	(1)	(2)	(3)	(4)	(5)	(6)
bas^H_{t-1}	0.535**	0.453**	0.260**	0.270**	0.277**	0.267**
	(0.008)	(0.009)	(0.010)	(0.011)	(0.011)	(0.011)
bas^H_{t-24}		0.212**	0.015		−0.001	
		(0.008)	(0.009)		(0.009)	
$bas^H_{t-24,t-1}$			0.620**	0.636**	0.477**	0.267**
			(0.016)	(0.014)	(0.024)	(0.022)
$bas^H_{t-7*24,t-1}$					0.214**	0.214**
					(0.022)	(0.022)
# Obs.	22 686	22 686	22 686	22 686	22 686	22 686
Adj. R^2	0.293	0.333	0.404	0.412	0.416	0.416
AIC	62 190	60 650	57 710	48 380	48 210	48 200

(Normalized standard errors in parentheses)
** $p < 0.01$; * $p < 0.05$

Table 10: Choice of HAR benchmark for daily bid-ask spread. bas^D_{t-1}, $bas^D_{t-7,t-1}$, $bas^D_{t-30,t-1}$ and $bas^D_{t-60,t-1}$ are the bid-ask spread for the previous day, the average over the last week, month and 2 months, respectively. Reported results are for Bitstamp. Results for other exchanges are very similar.

	Dependent variable: Daily bid-ask spread				
	(1)	(2)	(3)	(4)	(5)
bas^D_{t-1}	0.753**	0.405**	0.412**	0.409**	0.409**
	(0.021)	(0.046)	(0.047)	(0.047)	(0.047)
$bas^D_{t-7,t-1}$		0.456**	0.302**	0.321**	0.325**
		(0.055)	(0.062)	(0.068)	(0.063)
$bas^D_{t-30,t-1}$			0.181**	0.012	
			(0.052)	(0.092)	
$bas^D_{t-60,t-1}$				0.172*	0.180**
				(0.081)	(0.045)
# Obs.	1 392	1 392	1 392	1 392	1 392
Adj. R^2	0.590	0.643	0.648	0.651	0.651
AIC	2 437	2 245	2 228	2 221	2 219

(Normalized standard errors in parentheses)
** $p < 0.01$; * $p < 0.05$

Table 11: Comparison of explanatory power, adjusted R-squared, of different model specifications. Regressions for bid-ask spread on an hourly basis. Adding dummy variables does not contribute significantly to the explanatory power of the HAR benchmarks.

	Dependent variable: Hourly bid-ask spread			
Model	Bitstamp	Coinbase	BTCN	Korbit
Time-of-day dummies	0.004	0.004	0.001	0.013
Benchmark model	0.416	0.392	0.702	0.286
Dummies & Benchmark	0.418	0.395	0.703	0.293

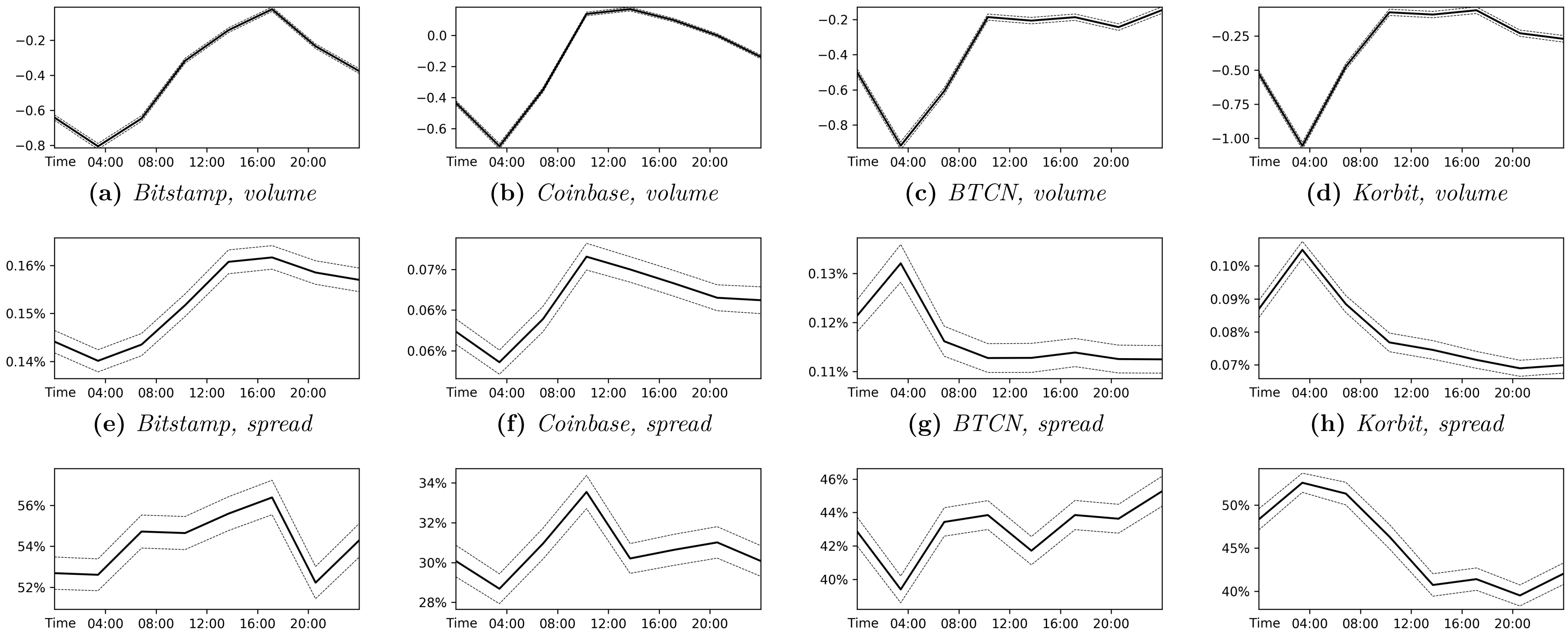
The fact that volumes are higher when most people are awake, and vice versa, is expected. The differences between the exchanges may be due to western traders being active also on eastern exchanges, resulting in a wider plateau of high trading volume, and higher bid-ask spreads in the night.

The patterns of the average volume and average bid-ask spread are similar for Bitstamp and Coinbase, but opposite for BTCN and Korbit. Even though average intraday patterns are opposite, volume and bid-ask spread can be positively related. Figure 9 shows how modest[10] the intraday seasonality of the bid-ask spread is when compared to the actual variation across three randomly chosen days.

Table 12 displays the results of a regression on bid-ask spread with time-of-day dummy variables, for each of the four exchanges. The coefficients are significant and convey the same pattern as was visible in Figure 8. The model, however, only explains a negligible share of the variation in the bid-ask spread. This is precisely what Figure 9 illustrates: though there is a pattern, it is not consequential.

The intraday pattern of the global volume index, shown in Figure 10, does not covary visibly with any of the other variables.

[10]An analogous analysis of the intraday seasonality in for example electricity prices would yield an entirely different picture. Here, one would observe very similar time-series each day, with minor random deviations from the pattern (Do et al. 2016).



(a) *Bitstamp, volume* **(b)** *Coinbase, volume* **(c)** *BTCN, volume* **(d)** *Korbit, volume*

(e) *Bitstamp, spread* **(f)** *Coinbase, spread* **(g)** *BTCN, spread* **(h)** *Korbit, spread*

(i) *Bitstamp, realized volatility* **(j)** *Coinbase, realized volatility* **(k)** *BTCN, realized volatility* **(l)** *Korbit, realized volatility*

Figure 8: Intraday seasonality of volume, bid-ask spread and realized volatility. Averages over three hour intervals. Local time of exchange. Dotted lines indicate 95% confidence interval.

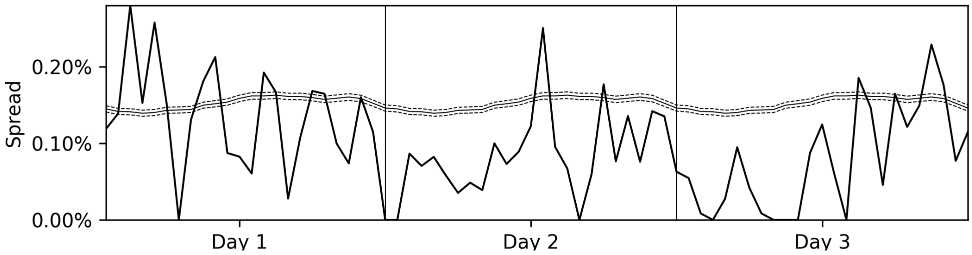


Figure 9: Hourly bid-ask spread on the Bitstamp exchange 01.06.2015-03.06.2015. Plotted with the average daily pattern throughout the day and its 95% confidence interval. The dates are chosen arbitrarily.

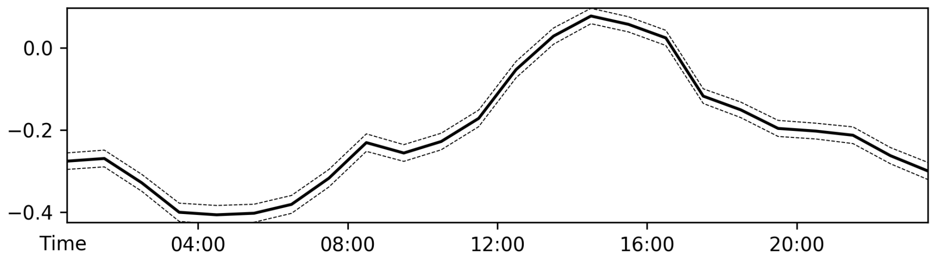


Figure 10: Intraday seasonality of global volume index. Universal coordinated time

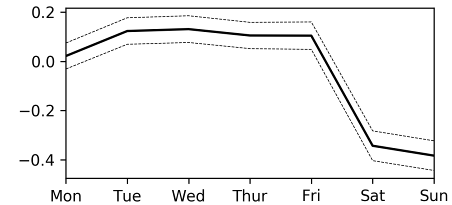
The intraweek seasonality is analyzed in a similar fashion to intraday seasonality. Figure 11 shows the seasonality plots, and Table 13 shows the results of a regression on bid-ask spread. These analyses offer less clear results compared with the intraday analyses: The coefficients in Table 13 are mostly insignificant; In Figure 11, volumes tend to be higher during working days than during the weekend; Apart from volume, there are no obvious patterns in any variables.

Table 12: Regression for bid-ask spread using dummy variables for 8 three-hour periods. Coefficients signify deviation from mean, i.e. a positive coefficient represents a greater value in the corresponding time frame than the average during the day as a whole. Local time for each exchange.

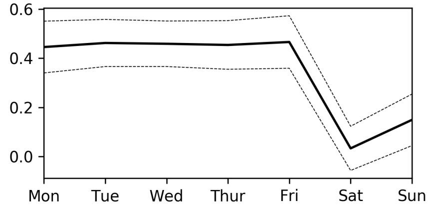
	Dependent variable: Hourly bid-ask spread			
	Bitstamp	Coinbase	BTCN	Korbit
00:00-02:59	−0.0081**	−0.0036**	0.0046	0.0067*
	(0.002)	(0.001)	(0.003)	(0.003)
03:00-05:59	−0.0120**	−0.0074**	0.0152**	0.0246**
	(0.002)	(0.001)	(0.004)	(0.004)
06:00-08:59	−0.0087**	−0.0020	−0.0007	0.0082**
	(0.002)	(0.001)	(0.003)	(0.003)
09:00-11:59	−0.0005	0.0056**	−0.0041	−0.0034
	(0.002)	(0.001)	(0.003)	(0.002)
12:00-14:59	0.0086**	0.0041**	−0.0038	−0.0057**
	(0.002)	(0.001)	(0.003)	(0.002)
15:00-17:59	0.0095**	0.0024	−0.0027	−0.0087**
	(0.002)	(0.001)	(0.003)	(0.002)
18:00-20:59	0.0063**	0.0006	−0.0043	−0.0113**
	(0.002)	(0.001)	(0.003)	(0.002)
21:00-23:59	0.0048**	0.0003	−0.0043	−0.0104**
	(0.002)	(0.001)	(0.002)	(0.002)
# Obs.	35 251	16 779	25 859	10 805
Adj. R^2	0.004	0.004	0.001	0.013
AIC	−373 404	−201 327	−259 896	−123 164

(Normalized standard errors in parentheses)

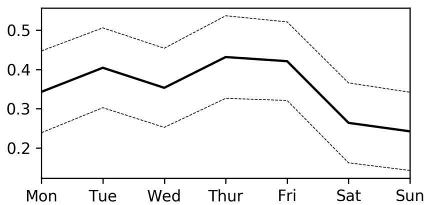
** $p < 0.01$; * $p < 0.05$



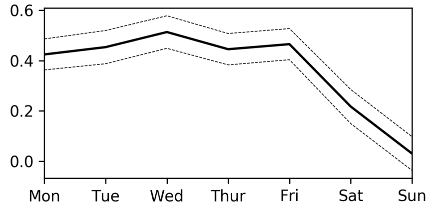
(a) *Bitstamp, volume*



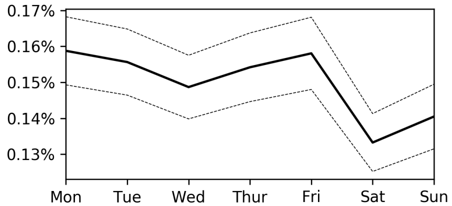
(b) *Coinbase, volume*



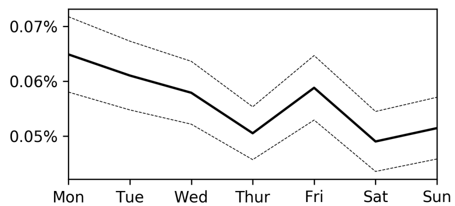
(c) *BTCN, volume*



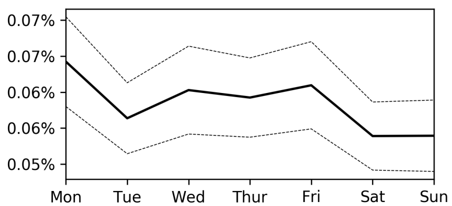
(d) *Korbit, volume*



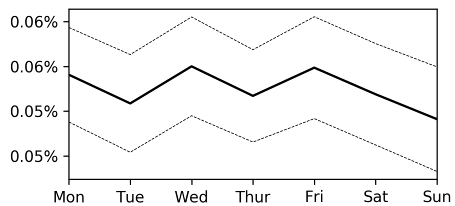
(e) *Bitstamp, spread*



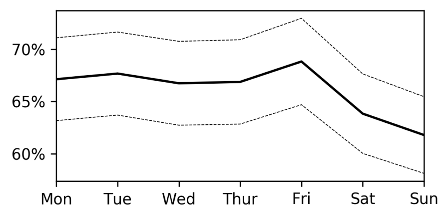
(f) *Coinbase, spread*



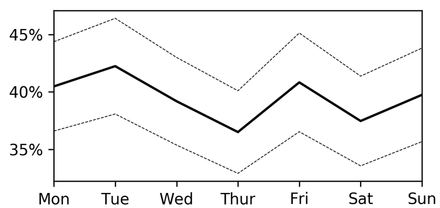
(g) *BTCN, spread*



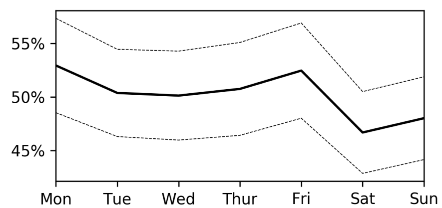
(h) *Korbit, spread*



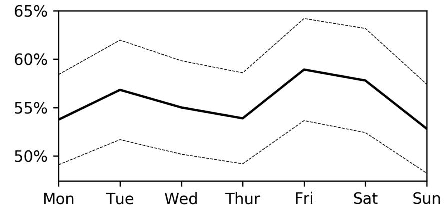
(i) *Bitstamp, realized volatility*



(j) *Coinbase, realized volatility*



(k) *BTCN, realized volatility*



(l) *Korbit, realized volatility*

Figure 11: Intraweek seasonality of volume, bid-ask spread and realized volatility. Dotted lines indicate 95% confidence interval.

Table 13: Regression for bid-ask spread using dummy variables for days of the week. Coefficients signify deviation from mean, i.e. a positive coefficient represents greater than average value in the corresponding day than during the week as a whole.

	Dependent variable: Daily bid-ask spread							
	Bitstamp		Coinbase		BTCN		Korbit	
Mon	0.009 (0.006)		0.009 (0.005)		0.006 (0.005)		0.002 (0.004)	
Tue	0.006 (0.006)		0.005 (0.004)		−0.002 (0.004)		−0.002 (0.004)	
Wed	−0.001 (0.006)		0.002 (0.004)		0.002 (0.005)		0.003 (0.004)	
Thu	0.004 (0.006)		−0.006 (0.003)		0.001 (0.004)		−0.001 (0.004)	
Fri	0.008 (0.007)		0.003 (0.004)		0.003 (0.005)		0.002 (0.004)	
Sat	−0.017** (0.005)		−0.007 (0.005)		−0.005 (0.004)		−0.001 (0.005)	
Sun	−0.009 (0.006)		−0.005 (0.004)		−0.005 (0.004)		−0.003 (0.005)	
Weekday		0.005 (0.003)		0.003 (0.002)		0.002 (0.002)		0.001 (0.002)
Weekend		−0.018* (0.005)		−0.008* (0.004)		−0.006 (0.003)		−0.003 (0.004)
# Obs.	1545	1545	677	677	1026	1026	815	815
Adj. R^2	0.006	0.008	0.008	0.006	-0.001	0.002	−0.006	−0.001
AIC	−17308	−17316	−8594	−8598	−12519	−12527	−10155	−10164

(Normalized standard errors in parentheses)

** $p < 0.01$; * $p < 0.05$

4.3 Return, volume, global volume and realized volatility as determinants

As discussed in Section 4.1, the autoregressive terms in the benchmark explain a substantial part of the variation in the bid-ask spread. Here we investigate whether absolute returns, local volumes, global volumes or realized volatility add explanatory power beyond what the benchmark provides. When dummy variables for seasonality are not included, we are left with the following equations for hourly and daily bid-ask spread, respectively:

$$bas_t^H = \beta_1 bas_{t-1}^H + \beta_2 bas_{t-24,t-1}^H + \beta_3 bas_{t-7*24,t-1}^H + \beta_4 |r_t^H| \\ + \beta_5 v_t^H + \beta_6 gv_t^H + \beta_7 rv_t^H + \epsilon_t \quad (9)$$

$$bas_t^D = \beta_1 bas_{t-1}^D + \beta_2 bas_{t-7,t-1}^D + \beta_3 bas_{t-60,t-1}^D + \beta_4 |r_t^D| \\ + \beta_5 v_t^D + \beta_6 gv_t^D + \beta_7 rv_t^D + \epsilon_t \quad (10)$$

Table 14 shows that on an hourly basis, including absolute returns, local volumes, global volumes and realized volatility adds explanatory power to the model. Local volumes and realized volatility consistently have significant coefficients at the 1% level. The adjusted R-squared of the benchmark model for Bitstamp improves from 0.416 to 0.484 when adding the contemporaneous values of absolute returns, local volumes, global volumes and realized volatility. A one standard deviation increase in local volumes is associated with a 0.181 standard deviations increase in the bid-ask spread. The equivalent number for realized volatility is 0.129.

This positive relationship between bid-ask spread and local volume, and bid-ask spread and realized volatility, is found for all four exchanges. What is noteworthy is that while the explanatory power of the benchmark model varies among the exchanges, the contemporaneous model is consistently better, and the same variables contribute explanatory power to all exchanges. The positive contemporaneous relation between volume and bid-ask spread is consistent with the findings of Daníelsson and Payne (2012) and Narayan et al. (2015), indicating that market makers indeed adjust spreads in line with the theory of information costs. The positive contemporaneous relation between volatility and bid-ask spread is consistent with the findings of McInish and Wood (1992), Ding (1999) and Galati (2000), lending support to the theory of inventory costs, where market makers adjust spreads to compensate for increased risk in their inventories.

The rightmost four columns of Table 14 show the predictive models, where the lagged values of the explanatory variables are included instead of the contemporaneous values. Across all four exchanges, previous hour absolute return is the most important determinant. Absolute returns measure the magnitude of price change. Such movements may indicate information arriving at the markets, and thus, market makers need to adjust their spreads in order to protect against losses, as per the information cost theory. Apart from absolute returns, no lagged variables consistently have significant coefficients in the predictive model.

In Table 15 we show the results of regressions on the daily bid-ask spread. Overall, the conclusions from the regressions on hourly bid-ask spread are found here as well. Realized volatility and local volumes are contemporaneously related to the bid-ask spread, though not as conclusively. High daily absolute returns are predictive of high bid-ask spread, as was the case in hourly data.

The following findings are consistent across all four exchanges, both for hourly and daily bid-ask spread: Local volume and bid-ask spread are positively contemporaneously related; Realized volatility and bid-ask spread are positively contemporaneously related; Absolute return is predictive of the bid-ask spread.

We do not find any persistent relationship between global volumes and bid-ask spread. This indicates that the bid-ask spread responds to the local conditions rather than global market conditions. As a robustness test, we conduct regressions with dummy variables for both hourly and daily bid-ask spread. We do not include the results here, as they convey the same conclusions as without dummy variables.

Table 14: Regression on hourly bid-ask spread for four different exchanges. bas^H_{t-1}, $bas^H_{t-24,t-1}$ and $bas^H_{t-7*24,t-1}$ are the bid-ask spread for the previous hour, the average over the last day and the average over the last week respectively. In addition to the benchmark HAR-model, both contemporaneous and predictive models are displayed. The predictive model has the previous hour values as explanatory variables.

	Benchmark model				Contemporaneous model				Predictive model			
	Bitstamp	*Coinbase*	*BTCN*	*Korbit*	*Bitstamp*	*Coinbase*	*BTCN*	*Korbit*	*Bitstamp*	*Coinbase*	*BTCN*	*Korbit*
bas^H_{t-1}	0.267**	0.256**	0.337**	0.229**	0.214**	0.201**	0.308**	0.185**	0.242**	0.240**	0.332**	0.221**
	(0.011)	(0.042)	(0.017)	(0.052)	(0.010)	(0.040)	(0.017)	(0.048)	(0.011)	(0.042)	(0.017)	(0.048)
$bas^H_{t-24,t-1}$	0.476**	0.550**	0.374**	0.524**	0.344**	0.450**	0.326**	0.302**	0.438**	0.508**	0.366**	0.513**
	(0.022)	(0.045)	(0.040)	(0.075)	(0.022)	(0.042)	(0.039)	(0.074)	(0.023)	(0.045)	(0.040)	(0.073)
$bas^H_{t-7*24,t-1}$	0.214**	0.133**	0.288**	0.116	0.302**	0.123**	0.304**	0.130*	0.278**	0.154**	0.301**	0.123
	(0.022)	(0.034)	(0.038)	(0.071)	(0.022)	(0.036)	(0.037)	(0.066)	(0.023)	(0.036)	(0.039)	(0.074)
$\|r^H\|$					0.010	0.007	−0.021*	0.038	0.046**	0.061*	0.049**	0.212*
					(0.008)	(0.016)	(0.010)	(0.028)	(0.010)	(0.028)	(0.013)	(0.086)
v^H					0.181**	0.161**	0.111**	0.125**	0.010	0.036**	0.018*	−0.094**
					(0.006)	(0.013)	(0.008)	(0.024)	(0.007)	(0.011)	(0.008)	(0.027)
gv^H					0.001	−0.021*	0.003	0.044*	0.063**	0.029**	0.008	0.039
					(0.007)	(0.010)	(0.006)	(0.019)	(0.008)	(0.011)	(0.006)	(0.034)
rv^H					0.129**	0.150**	0.134**	0.270**	0.006	0.015	−0.023*	−0.056
					(0.005)	(0.011)	(0.008)	(0.021)	(0.006)	(0.013)	(0.010)	(0.044)
# Obs.	22 686	8 029	11 878	2 383	22 686	8 029	11 878	2 383	22 686	8 029	11 878	2 383
Adj. R^2	0.416	0.392	0.702	0.286	0.484	0.447	0.720	0.370	0.424	0.400	0.703	0.309
AIC	48 204	17 356	22 903	5 498	45 428	16 608	22 131	5 204	47 906	17 254	22 862	5423

(Normalized standard errors in parentheses)

** $p < 0.01$; * $p < 0.05$

Table 15: Regressions on daily bid-ask spread for four different exchanges. bas^D_{t-1}, $bas^D_{t-7,t-1}$ and $bas^D_{t-60,t-1}$ are the bid-ask spread for the previous day, the average over the last week and the average over the last 2 months, respectively. In addition to the benchmark HAR-model, both contemporaneous and predictive models are displayed. The predictive model has the previous day values as explanatory variables.

	Benchmark model				Contemporaneous model				Predictive model			
	Bitstamp	*Coinbase*	*BTCN*	*Korbit*	*Bitstamp*	*Coinbase*	*BTCN*	*Korbit*	*Bitstamp*	*Coinbase*	*BTCN*	*Korbit*
bas^D_{t-1}	0.409**	0.304**	0.078	0.347**	0.326**	0.198**	0.015	0.284**	0.397**	0.251**	0.053	0.310**
	(0.047)	(0.073)	(0.062)	(0.072)	(0.043)	(0.064)	(0.057)	(0.062)	(0.047)	(0.070)	(0.059)	(0.073)
$bas^D_{t-7,t-1}$	0.325**	0.217	0.566**	0.274**	0.246**	0.035	0.460**	0.091	0.323**	0.136	0.556**	0.327**
	(0.063)	(0.136)	(0.103)	(0.078)	(0.055)	(0.106)	(0.099)	(0.074)	(0.064)	(0.106)	(0.105)	(0.078)
$bas^D_{t-60,t-1}$	0.180**	0.328**	0.260**	0.297**	0.203**	0.088	0.121	0.156*	0.135**	0.460**	0.240**	0.296**
	(0.045)	(0.102)	(0.080)	(0.077)	(0.045)	(0.097)	(0.072)	(0.073)	(0.048)	(0.097)	(0.077)	(0.080)
$\lvert r^D \rvert$					−0.020	0.062	−0.031	0.003	0.099**	0.217**	0.107**	0.190**
					(0.017)	(0.042)	(0.024)	(0.033)	(0.018)	(0.058)	(0.036)	(0.058)
v^D					0.049	0.257**	0.004	0.224**	0.113**	−0.116	−0.022	0.023
					(0.035)	(0.078)	(0.017)	(0.054)	(0.038)	(0.071)	(0.017)	(0.059)
gv^D					0.015	−0.095*	−0.008	0.060	−0.134**	−0.061	−0.027	−0.028
					(0.035)	(0.045)	(0.026)	(0.045)	(0.039)	(0.066)	(0.026)	(0.064)
rv^D					0.215**	0.389**	0.348**	0.242**	−0.010	0.184**	0.052	−0.114*
					(0.028)	(0.063)	(0.039)	(0.036)	(0.028)	(0.065)	(0.038)	(0.045)
# Obs.	1 392	553	844	550	1 392	553	844	550	1 392	553	844	550
Adj. R^2	0.650	0.404	0.504	0.431	0.704	0.527	0.574	0.557	0.661	0.446	0.513	0.451
AIC	2 218	1 272	1 644	1 219	1 987	1 148	1 520	1 086	2 177	1 235	1 632	1 203

(Normalized standard errors in parentheses)

** $p < 0.01$; * $p < 0.05$

4.4 Subsample analysis of determinants

In the following we conduct similar analyses as in Section 4.3, but on subsets of our data. The bitcoin price has gone through periods of rapid growth and decline, and we investigate whether the findings from Section 4.3 hold for a *growth* period and a *decline* period. In the growth period we investigate, 05.10.2015 to 30.12.2016, the price increased by 300% from \$240 to \$961. In the decline period, 01.01.2015 to 20.10.2015, the price dropped 14%, from \$314 to \$269.

Table 16 and 17 show results for regressions on hourly and daily bid-ask spread for the growth period. Table 18 and 19 show results for regressions on hourly and daily bid-ask spread for the decline period. The results are consistent with the findings for the full data set. This indicates that the relationships between the bid-ask spread, absolute return, traded volume and realized volatility exist regardless of whether the bitcoin price is increasing or decreasing. The coefficients are less significant, as would be expected when the regressions are conducted on fewer observations.

Table 16: Regression on hourly bid-ask spread for four different exchanges in a growth period. bas^H_{t-1}, $bas^H_{t-24,t-1}$ and $bas^H_{t-7*24,t-1}$ are the bid-ask spread for the previous hour, the average over the last day and the average over the last week respectively. In addition to the benchmark HAR-model, both contemporaneous and predictive models are displayed. The predictive model has the previous hour values as explanatory variables. The time period calculated for is 05.10.2015 - 30.12.2016.

	Benchmark model				Contemporaneous model				Predictive model			
	Bitstamp	*Coinbase*	*BTCN*	*Korbit*	*Bitstamp*	*Coinbase*	*BTCN*	*Korbit*	*Bitstamp*	*Coinbase*	*BTCN*	*Korbit*
bas^H_{t-1}	0.256**	0.269**	0.200**	0.247	0.200**	0.214**	0.146**	0.186	0.230**	0.251**	0.186**	0.198
	(0.017)	(0.041)	(0.024)	(0.191)	(0.016)	(0.040)	(0.023)	(0.164)	(0.017)	(0.041)	(0.024)	(0.116)
$bas^H_{t-24,t-1}$	0.552**	0.549**	0.513**	0.108	0.403**	0.474**	0.330**	−0.282	0.502**	0.513**	0.474**	0.149
	(0.039)	(0.059)	(0.071)	(0.180)	(0.037)	(0.058)	(0.070)	(0.184)	(0.040)	(0.059)	(0.073)	(0.194)
$bas^H_{t-24*7,t-1}$	0.088*	0.070	0.226**	0.390	0.131**	0.053	0.219**	0.398	0.113**	0.081	0.233**	0.334
	(0.040)	(0.049)	(0.067)	(0.269)	(0.039)	(0.056)	(0.067)	(0.208)	(0.041)	(0.055)	(0.069)	(0.272)
$\|r^H\|$					0.014	0.023	0.050*	0.162**	0.105**	0.041	0.071	0.738*
					(0.015)	(0.023)	(0.021)	(0.062)	(0.022)	(0.034)	(0.040)	(0.363)
v^H					0.195**	0.148**	0.008	0.272**	−0.010	0.039**	−0.014	−0.042
					(0.013)	(0.020)	(0.010)	(0.065)	(0.014)	(0.015)	(0.011)	(0.073)
gv^H					−0.010	−0.027	0.048**	0.070	0.055**	0.022	0.028	−0.164
					(0.014)	(0.014)	(0.013)	(0.049)	(0.016)	(0.016)	(0.014)	(0.123)
rv^H					0.195**	0.126**	0.199**	0.398**	0.034**	0.026	0.023	−0.310
					(0.011)	(0.015)	(0.019)	(0.061)	(0.012)	(0.016)	(0.019)	(0.167)
# Obs.	6 755	3 933	3 316	438	6 755	3 933	3 316	438	6 755	3 933	3 316	438
R^2	0.319	0.402	0.420	0.124	0.405	0.457	0.473	0.309	0.333	0.410	0.427	0.285
AIC	15 761	7 567	5 913	1 297	14 856	7 193	5 601	1 198	15 631	7 516	5 879	1 212

(Normalized standard errors in parentheses)

** $p < 0.01$; * $p < 0.05$

Table 17: Regressions on daily bid-ask spread for four different exchanges in a growth period. bas^D_{t-1}, $bas^D_{t-7,t-1}$ and $bas^D_{t-60,t-1}$ are the bid-ask spread for the previous day, the average over the last week and the average over the last 2 months, respectively. In addition to the benchmark HAR-model, both contemporaneous and predictive models are displayed. The predictive model has the previous day values as explanatory variables. The time period calculated for is 05.10.2015 - 30.12.2016.

	Benchmark model				Contemporaneous model				Predictive model			
	Bitstamp	*Coinbase*	*BTCN*	*Korbit*	*Bitstamp*	*Coinbase*	*BTCN*	*Korbit*	*Bitstamp*	*Coinbase*	*BTCN*	*Korbit*
bas^D_{t-1}	0.611**	0.301*	0.311**	0.502**	0.498**	0.153	0.238*	0.368**	0.614**	0.230*	0.311**	0.418**
	(0.089)	(0.124)	(0.108)	(0.134)	(0.080)	(0.097)	(0.105)	(0.117)	(0.095)	(0.114)	(0.113)	(0.128)
$bas^D_{t-7,t-1}$	0.113	0.219	0.425**	0.032	0.052	−0.054	0.197	−0.327*	0.105	0.162	0.425**	0.014
	(0.091)	(0.220)	(0.116)	(0.130)	(0.083)	(0.161)	(0.114)	(0.143)	(0.085)	(0.166)	(0.122)	(0.102)
$bas^D_{t-60,t-1}$	0.078	0.154	−0.072	0.265	−0.001	0.023	−0.252	0.022	0.072	0.090	−0.230	0.182
	(0.088)	(0.177)	(0.144)	(0.168)	(0.081)	(0.151)	(0.151)	(0.138)	(0.084)	(0.176)	(0.171)	(0.188)
$\lvert r^D \rvert$					−0.085*	0.041	0.057	−0.077	0.066	0.304**	0.147	0.336*
					(0.034)	(0.051)	(0.053)	(0.056)	(0.046)	(0.111)	(0.075)	(0.166)
v^D					0.069	0.082	−0.092	0.190**	0.157*	0.029	−0.114	−0.064
					(0.076)	(0.122)	(0.054)	(0.065)	(0.072)	(0.122)	(0.064)	(0.073)
gv^D					−0.001	0.046	−0.027	0.113	−0.172**	−0.052	−0.042	0.137
					(0.075)	(0.104)	(0.058)	(0.078)	(0.066)	(0.146)	(0.066)	(0.101)
rv^D					0.304**	0.484**	0.344**	0.437**	−0.009	0.132	0.019	−0.092
					(0.049)	(0.093)	(0.065)	(0.075)	(0.052)	(0.079)	(0.080)	(0.092)
# Obs.	398	263	276	223	398	263	276	223	398	263	276	223
R^2	0.569	0.222	0.262	0.291	0.649	0.440	0.328	0.542	0.573	0.320	0.273	0.375
AIC	730	661	639	560	652	579	617	466	730	630	639	536

(Normalized standard errors in parentheses)

** $p < 0.01$; * $p < 0.05$

Table 18: Regression on hourly bid-ask spread for four different exchanges in a decline period. bas^H_{t-1}, $bas^H_{t-24,t-1}$ and $bas^H_{t-7*24,t-1}$ are the bid-ask spread for the previous hour, the average over the last day and the average over the last week respectively. In addition to the benchmark HAR-model, both contemporaneous and predictive models are displayed. The predictive model has the previous hour values as explanatory variables. The time period calculated for is 01.01.2015 - 20.10.2015.

	Benchmark model				Contemporaneous model				Predictive model			
	Bitstamp	*Coinbase*	*BTCN*	*Korbit*	*Bitstamp*	*Coinbase*	*BTCN*	*Korbit*	*Bitstamp*	*Coinbase*	*BTCN*	*Korbit*
bas^H_{t-1}	0.241**	0.253**	0.097*	0.164	0.159**	0.192**	0.066	0.104	0.207**	0.232**	0.093	0.104
	(0.022)	(0.071)	(0.049)	(0.191)	(0.020)	(0.066)	(0.046)	(0.160)	(0.024)	(0.071)	(0.048)	(0.146)
$bas^H_{t-24,t-1}$	0.427**	0.537**	0.522**	0.175	0.236**	0.397**	0.364**	0.027	0.382**	0.482**	0.470**	0.221
	(0.055)	(0.062)	(0.089)	(0.302)	(0.050)	(0.056)	(0.092)	(0.264)	(0.056)	(0.061)	(0.086)	(0.277)
$bas^H_{t-24*7,t-1}$	0.260**	0.160**	0.312**	−0.097	0.465**	0.157**	0.336**	−0.427	0.327**	0.191**	0.298**	0.081
	(0.058)	(0.050)	(0.094)	(0.469)	(0.058)	(0.050)	(0.092)	(0.432)	(0.061)	(0.052)	(0.099)	(0.537)
$\|r^H\|$					−0.000	0.000	0.046	0.207*	0.034	0.080	0.085	0.266
					(0.016)	(0.018)	(0.034)	(0.095)	(0.026)	(0.048)	(0.045)	(0.162)
v^H					0.336**	0.192**	0.050**	0.136	0.013	0.032	0.034*	−0.172
					(0.019)	(0.019)	(0.015)	(0.091)	(0.020)	(0.018)	(0.017)	(0.154)
gv^H					−0.013	−0.014	0.064**	−0.112	0.067**	0.038*	−0.017	0.006
					(0.017)	(0.016)	(0.022)	(0.088)	(0.022)	(0.017)	(0.017)	(0.126)
rv^H					0.142**	0.179**	0.124**	0.191*	0.019	0.013	0.027	−0.141
					(0.014)	(0.018)	(0.019)	(0.086)	(0.018)	(0.019)	(0.025)	(0.162)
# Obs.	3 873	3 111	1 817	65	3 873	3 111	1 817	65	3 873	3 111	1 817	65
R^2	0.315	0.345	0.266	0.028	0.459	0.417	0.299	0.168	0.322	0.354	0.274	0.018
AIC	9 403	7 158	3 999	116	8 492	6 799	3 921	110	9 369	7 118	3 984	121

(Normalized standard errors in parentheses)

** $p < 0.01$; * $p < 0.05$

Table 19: Regressions on daily bid-ask spread for four different exchanges in a decline period. bas^D_{t-1}, $bas^D_{t-7,t-1}$ and $bas^D_{t-60,t-1}$ are the bid-ask spread for the previous day, the average over the last week and the average over the last 2 months, respectively. In addition to the benchmark HAR-model, both contemporaneous and predictive models are displayed. The predictive model has the previous day values as explanatory variables. The time period calculated for is 01.01.2015 - 20.10.2015.

	Benchmark model				Contemporaneous model				Predictive model			
	Bitstamp	*Coinbase*	*BTCN*	*Korbit*	*Bitstamp*	*Coinbase*	*BTCN*	*Korbit*	*Bitstamp*	*Coinbase*	*BTCN*	*Korbit*
bas^D_{t-1}	0.368**	0.260*	−0.028	0.185	0.236**	0.177*	−0.105	0.219	0.329**	0.202	−0.030	0.136
	(0.074)	(0.101)	(0.086)	(0.150)	(0.065)	(0.082)	(0.085)	(0.125)	(0.061)	(0.109)	(0.088)	(0.147)
$bas^D_{t-7,t-1}$	0.534**	0.365**	0.504**	0.143	0.439**	0.049	0.379**	−0.357**	0.385**	0.288*	0.384**	0.194
	(0.108)	(0.137)	(0.144)	(0.181)	(0.102)	(0.120)	(0.145)	(0.132)	(0.108)	(0.129)	(0.144)	(0.185)
$bas^D_{t-60,t-1}$	−0.032	0.250**	0.186	−0.031	0.033	0.049	−0.279	−0.635	0.132	0.370**	0.295	0.020
	(0.067)	(0.096)	(0.186)	(0.354)	(0.080)	(0.124)	(0.263)	(0.398)	(0.078)	(0.126)	(0.270)	(0.349)
$\|r^D\|$					0.033	0.137	0.014	−0.015	0.213**	0.139*	0.193	0.198*
					(0.036)	(0.078)	(0.073)	(0.083)	(0.044)	(0.068)	(0.125)	(0.088)
v^D					0.273**	0.360**	0.169	0.054	0.390**	−0.119	−0.156	−0.025
					(0.088)	(0.117)	(0.130)	(0.111)	(0.097)	(0.120)	(0.129)	(0.109)
gv^D					−0.033	−0.001	0.047	0.180	−0.435**	−0.102	−0.066	−0.116
					(0.101)	(0.062)	(0.063)	(0.109)	(0.107)	(0.082)	(0.065)	(0.109)
rv^D					0.155**	0.273**	0.293**	0.535**	0.103	0.192	0.236*	−0.031
					(0.052)	(0.084)	(0.090)	(0.110)	(0.070)	(0.126)	(0.099)	(0.095)
# Obs.	261	212	185	124	261	212	185	124	261	212	185	124
R^2	0.638	0.466	0.122	0.033	0.727	0.604	0.206	0.343	0.701	0.481	0.151	0.036
AIC	458	470	488	336	388	410	473	292	412	468	486	340

(Normalized standard errors in parentheses)

** $p < 0.01$; * $p < 0.05$

5 Conclusions

This paper analyzes the liquidity of bitcoin. In particular, we study which variables are contemporaneously related to the liquidity of bitcoin, which variables can predict it, and whether liquidity exhibits seasonal patterns. We measure liquidity as bid-ask spread, which is inversely related to liquidity. Though the bitcoin markets have been recently analyzed in terms of returns and volatility, little research has been done on liquidity. We use one-minute price and volume data for four different exchanges, spanning several years.

We find that traded volume and realized volatility are negatively related to liquidity, and that high absolute returns predict lower liquidity in the next period. These conclusions are valid for all the four analyzed exchanges, located in four different time-zones on three different continents, trading in three different currencies. The conclusions are true both on an hourly and a daily basis. Even though seasonal patterns in liquidity are statistically significant, they are economically insignificant.

The positive relationship between volume and bid-ask spread may indicate that market makers adjust bid-ask spreads to avoid losing money when trading against informed traders, as volume may signal informed trading. The positive relationship between volatility and bid-ask spread can be explained by the inventory cost theory, where it is expected that market makers adjust bid-ask spreads in response to increased volatility, as the risk of their positions increases.

A possible point of further research would be to compare the liquidity, and determinants thereof, to both traditional foreign exchange markets and to traditional assets like stock indices and commodities like gold. It would be interesting to add data from Google Trends or similar as explanatory variables in the regression analyses, to account for the spike in popularity for bitcoin the last few years. Another interesting approach would be to isolate unexpected large trading volumes in order to capture information coming to the market. Also, other measures of liquidity, such as depth, would probably provide a better understanding. Acquiring an extensive dataset of high-frequency data including quotes and order books would undoubtedly enhance the analyses.

References

Acharya, V. V. and L. H. Pedersen (2005). "Asset pricing with liquidity risk". In: *Journal of financial Economics* 77.2, pp. 375–410.

Amihud, Y. (2002). "Illiquidity and stock returns: cross-section and time-series effects". In: *Journal of Financial Markets* 5.1, pp. 31–56. ISSN: 1386-4181. URL: https://doi.org/10.1016/S1386-4181(01)00024-6.

Amihud, Y. and H. Mendelson (1986). "Liquidity and Stock Returns". In: *Financial Analysts Journal* 42.3, pp. 43–48. ISSN: 0015198X. URL: http://www.jstor.org/stable/4478932.

Andersen, T. G. and T. Bollerslev (1998). "Deutsche Mark-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer Run Dependencies". In: *The Journal of Finance* 53.1, pp. 219–265. ISSN: 00221082. URL: http://doi.wiley.com/10.1111/0022-1082.85732.

Anonymous (2017). *Fake It Till You Make It: When Bitfinex themselves used to spoof their entire orderbook.* URL: https://medium.com/@bitfinexed/fake-it-till-you-make-it-when-bitfinex-themselves-used-to-spoof-their-entire-orderbook-18294585338 (visited on 12/12/2017).

Balcilar, M. et al. (2016). "Can Volume Predict Bitcoin Returns and Volatility? A Nonparametric Causality-in-Quantiles Approach". In: *Economic Modelling* 64, pp. 71–81. URL: https://ssrn.com/abstract=2938555.

Benston, G. J. and R. L. Hagerman (1974). "Determinants of bid-asked spreads in the over-the-counter market". In: *Journal of Financial Economics* 1.4, pp. 353–364. ISSN: 0304-405X. URL: https://www.sciencedirect.com/science/article/pii/0304405X74900142.

Böhme, R. et al. (2015). "Bitcoin: Economics, Technology, and Governance". In: *The Journal of Economic Perspectives* 29, pp. 213–238. URL: https://www.jstor.org/stable/24292130.

Bouri, E. et al. (2017). "On the hedge and safe haven properties of Bitcoin: Is it really more than a diversifier?" In: *Finance Research Letters* 20, pp. 192–198. URL: http://www.sciencedirect.com/science/article/pii/S1544612316301817.

Branch, B. and W. Freed (1977). "Bid-asked Spreads on the AMEX and the Big Board". In: *The Journal of Finance* 32.1, pp. 159–163. ISSN: 00221082. URL: http://doi.wiley.com/10.1111/j.1540-6261.1977.tb03249.x.

Brandvold, M. et al. (2015). "Price discovery on Bitcoin exchanges". In: *Journal of International Financial Markets, Institutions and Money* 36, pp. 18–35. ISSN: 1042-4431. URL: https://www.sciencedirect.com/science/article/pii/S104244311500027X.

Brennan, M. J. and A. Subrahmanyam (1996). "Market microstructure and asset pricing: On the compensation for illiquidity in stock returns". In: *Journal of*

Financial Economics 41.3, pp. 441–464. ISSN: 0304-405X. URL: https://www.sciencedirect.com/science/article/pii/0304405X9500870K.

Chordia, T., R. Roll, and A. Subrahmanyam (2000). "Commonality in liquidity". In: *Journal of Financial Economics* 56.1, pp. 3–28. ISSN: 0304405X. URL: https://www.sciencedirect.com/science/article/pii/S0304405X99000574.

— (2001). "Market Liquidity and Trading Activity". In: *The Journal of Finance* 56.2, pp. 501–530. ISSN: 00221082. URL: http://doi.wiley.com/10.1111/0022-1082.00335.

Coggan, P. (2017). *The bitcoin bubble.* URL: https://www.economist.com/blogs/buttonwood/2017/11/greater-fool-theory-0 (visited on 12/18/2017).

Corwin, S. A. and P. Schultz (2012). "A Simple Way to Estimate Bid-Ask Spreads from Daily High and Low Prices". In: *The Journal of Finance* 67.2, pp. 719–760. ISSN: 1540-6261. URL: http://dx.doi.org/10.1111/j.1540-6261.2012.01729.x.

Daníelsson, J. and R. Payne (2012). "Liquidity determination in an order-driven market". In: *The European Journal of Finance* 18.9, pp. 799–821. ISSN: 1351-847X. URL: http://www.tandfonline.com/doi/abs/10.1080/1351847X.2011.601654.

Demsetz, H. (1968). "The Cost of Transacting". In: *The Quarterly Journal of Economics* 82.1, p. 33. ISSN: 00335533. URL: https://academic.oup.com/qje/article-lookup/doi/10.2307/1882244.

Dimpfl, T. (2017). *Bitcoin Market Microstructure*. University of Tubingen. URL: https://ssrn.com/abstract=2949807.

Ding, D. K. (1999). "The determinants of bid-ask spreads in the foreign exchange futures market: A microstructure analysis". In: *Journal of Futures Markets* 19.3, pp. 307–324. ISSN: 0270-7314. URL: https://doi.org/10.1002/(SICI)1096-9934(199905)19:3%3C307::AID-FUT4%3E3.0.CO;2-5.

Do, L. P. C., K. Lin, and P. Molnár (2016). "Electricity consumption modelling: A case of Germany". In: *Economic Modelling* 55, pp. 92–101. ISSN: 0264-9993.

Donier, J. and J. Bonart (2017). "A Million Metaorder Analysis of Market Impact on the Bitcoin". In: *World Scientific* 1.2, p. 29. URL: https://doi.org/10.1142/S2382626615500082.

Donier, J. and J. Bouchaud (2015). "Why Do Markets Crash? Bitcoin Data Offers Unprecedented Insights". In: *PLOS ONE* 10.10. ISSN: 1932-6203. URL: http://dx.plos.org/10.1371/journal.pone.0139356.

Dyhrberg, A. H. (2016). "Bitcoin, gold and the dollar – A GARCH volatility analysis". In: *Finance Research Letters* 16, pp. 85–92. ISSN: 1544-6123. URL: https://www.sciencedirect.com/science/article/pii/S1544612315001038.

Easley, D., M. O'Hara, and S. Basu (2017). "From Mining to Markets: the Evolution of Bitcoin Transaction Fees". In: *SSRN Electronic Journal*. ISSN: 1556-5068. URL: https://www.ssrn.com/abstract=3055380.

Galati, G. (2000). "Trading volumes, volatility and spreads in foreign exchange markets: Evidence from emerging market countries". URL: https://www.bis.org/publ/work93.htm.

Glosten, L. and L. Harris (1988). "Estimating the components of the bid/ask spread". In: *Journal of Financial Economics* 21.1, pp. 123–142. ISSN: 0304-405X. URL: https://www.sciencedirect.com/science/article/pii/0304405X88900347.

Goyenko, R. Y., C. W. Holden, and C. A. Trzcinka (2009). "Do liquidity measures measure liquidity?" In: *Journal of Financial Economics* 92.2, pp. 153–181. ISSN: 0304-405X. URL: https://doi.org/10.1016/j.jfineco.2008.06.002.

Hamilton, J. L. (1976). "Competition, Scale Economies, and Transaction Cost in the Stock Market". In: *The Journal of Financial and Quantitative Analysis* 11.5, p. 779. ISSN: 00221090. URL: http://www.jstor.org/stable/2330580?origin=crossref.

Harris, B. (2018). *Bitcoin is the greatest scam in history*. URL: https://www.recode.net/2018/4/24/17275202/bitcoin-scam-cryptocurrency-mining-pump-dump-fraud-ico-value (visited on 04/26/2018).

Hasbrouck, J. (1988). "Trades, quotes, inventories, and information". In: *Journal of Financial Economics* 22.2, pp. 229–252. ISSN: 0304-405X. URL: https://www.sciencedirect.com/science/article/pii/0304405X88900700.

Johnson, S. (2018). "Beyond the Bitcoin Bubble". In: *New York Times Magazine*. URL: https://www.nytimes.com/2017/11/21/technology/bitcoin-bitfinex-tether.html (visited on 04/26/2018).

Kim, T. (2017). "On the transaction cost of Bitcoin". In: *Finance Research Letters* 23, pp. 300–305. ISSN: 1544-6123. URL: https://doi.org/10.1016/j.frl.2017.07.014.

Lee, C. M. C., B. Mucklow, and M. J. Ready (1993). "Spreads, Depths, and the Impact of Earnings Information: An Intraday Analysis". In: *The Review of Financial Studies* 6, pp. 345–374. URL: http://www.jstor.org/stable/2962058.

Lybek, T. and A. Sarr (2002). *Measuring Liquidity in Financial Markets*. International Monetary Fund. URL: https://econpapers.repec.org/paper/imfimfwpa/02%7B%5C_%7D2f232.htm.

McInish, T. H. and R. A. Wood (1992). "An Analysis of Intraday Patterns in Bid/Ask Spreads for NYSE Stocks". In: *The Journal of Finance* 47.2, p. 753. ISSN: 00221082. URL: http://www.jstor.org/stable/2329122?origin=crossref.

Narayan, P. K., S. Mishra, and S. Narayan (2015). "New empirical evidence on the bid-ask spread". In: *Applied Economics* 47.42, pp. 4484–4500. ISSN: 0003-6846. URL: http://www.tandfonline.com/doi/full/10.1080/00036846.2015.1031870.

Pástor, L. and R. F. Stambaugh (2003). "Liquidity Risk and Expected Stock Returns". In: *Journal of Political Economy* 111.3, pp. 642–685. ISSN: 0022-3808. URL: http://www.journals.uchicago.edu/doi/10.1086/374184.

Patel, D. (2017). *How The Bitcoin Revolution Will Affect Entrepreneurs*. URL: https://www.forbes.com/sites/deeppatel/2017/06/26/how-the-bitcoin-revolution-will-affect-entrepreneurs/#2b813cde59cb (visited on 12/18/2017).

Popper, N. (2017). "Warning Signs About Another Giant Bitcoin Exchange". In: *New York Times*. URL: https://www.nytimes.com/2017/11/21/technology/bitcoin-bitfinex-tether.html (visited on 12/12/2017).

Rees, M. (2014). *Bitcoin to Earth: Don't Look Now, but your Paradigm is Shifting*. URL: https://bitcoinmagazine.com/articles/bitcoin-earth-dont-look-now-paradigm-shifting-1406062368/ (visited on 04/26/2018).

Rizzo, P. (2017). *Chinas Bitcoin Exchanges End No-Fee Trading in Market Shake-Up*. URL: https://www.coindesk.com/chinas-big-three-bitcoin-exchanges-end-no-fee-policy/ (visited on 12/12/2017).

Roll, R. (1984). "A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market". In: *The Journal of Finance* 39.4, pp. 1127–1139. ISSN: 00221082, 15406261. URL: http://www.jstor.org/stable/2327617.

Stoll, H. R. (1978). "The Supply of Dealer Services in Securities Markets". In: *The Journal of Finance* 33.4, pp. 1133–1151. ISSN: 00221082. URL: http://doi.wiley.com/10.1111/j.1540-6261.1978.tb02053.x.

Tinic, S. M. (1972). "The Economics of Liquidity Services". In: *The Quarterly Journal of Economics* 86.1, p. 79. ISSN: 00335533. URL: https://academic.oup.com/qje/article-lookup/doi/10.2307/1880494.

Tinic, S. M. and R. R. West (1972). "Competition and the Pricing of Dealer Service in the Over-the-Counter Stock Market". In: *The Journal of Financial and Quantitative Analysis* 7.3, p. 1707. ISSN: 00221090. URL: http://www.jstor.org/stable/2329797?origin=crossref.

Urquhart, A. (2016). "The Inefficiency of Bitcoin". In: *Economics Letters* 148.Supplement C, pp. 80–82. ISSN: 0165-1765. URL: http://www.sciencedirect.com/science/article/pii/S0165176516303640.