

Louvain School of Management

Co-movements in market quality of cryptocurrencies

Authors : Hubert ANCIAUX and Nicolas NICAISE
Supervisor : Mikael PETITJEAN
Academic year : 2018-2019

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

Abstract

Louvain School of Management

Master Thesis

Co-movements in market quality of cryptocurrencies

by Hubert ANCIAUX and Nicolas NICAISE

Using intraday data of the transactions and order book of the five cryptocurrencies with the highest market capitalizations in January 2018 from the Bitfinex exchange between August 2017 and July 2018, we extend the market quality and cryptocurrency literatures by investigating the market quality of cryptocurrencies. We replicate the approaches of Marsh and Mazza (2017) and Beaupain, Giot, and Petitjean (2010) in order to assess market quality commonality, liquidity commonality and the impact of volatility regimes on liquidity commonality. We found significant evidence of liquidity commonality (varying in intensity depending on the types of liquidity proxies chosen) in both quiet and volatile regimes for our set of cryptocurrencies with a positive relationship between commonality and volatility. Our results also suggest the presence of commonality in market quality movements of our set of cryptocurrencies. Counter-intuitively, the intensity of the common movements seems to be more important at the end of our period compared to the first three months of the period. Comparisons of the average explanatory power of Chordia, Roll, and Subrahmanyam (2000)'s market model at given periods leads to mixed results. Most conclusive is that for 16 out of 30 of our liquidity proxies across our set of cryptocurrencies, the market model regression assessed on a weekly basis had a significantly higher average explanatory power during the month preceding the end of the Bitcoin bubble as dated by Liu et al. (2019) compared to the month following it. This partially corroborates Liu et al. (2019), who also had stronger explanatory power in their model during the bubble.

Acknowledgements

We would like to thank our supervisor, Mr Mikael Petitjean, who gave us directions and precious guidance to realize this master's thesis. We would also like to express our sincere gratitude to Mr Christophe Desagre, for his availability and his invaluable help all along our project. Finally, we would like to thank our families for their support.

Contents

Abstract	iii
Acknowledgements	v
1 Introduction	1
2 Literature review	5
2.1 Cryptocurrency literature	5
2.1.1 Definitions and key concepts	5
2.1.2 Regulation	6
2.1.3 Pricing bubbles	7
2.1.4 Specificity as an asset class	8
2.1.5 Market efficiency	8
2.1.6 Introduction of futures	11
2.1.7 Other aspects of cryptocurrencies	12
2.2 Market quality	12
Adverse selection	13
Cost of immediacy	13
Liquidity	13
Price efficiency	13
Volatility	13
2.2.1 Commonality	14
Sources and pricing of commonality in market quality	15
2.2.2 Impact of high frequency trading	16

2.2.3	Volatility regimes and liquidity distributions	18
2.3	Market quality of cryptocurrencies	18
3	Methodology	19
3.1	Research questions and hypotheses	19
3.1.1	Commonality in market quality	19
3.1.2	Improvement of market efficiency, introduction of futures on Bitcoin and Bitcoin bubble	20
3.1.3	Influence of volatility regimes	21
3.1.4	Evolution of liquidity commonality	21
3.2	Description of our methodology	22
3.2.1	Assessing commonality in market quality	22
3.2.2	Assessing the impact of volatility regimes on liquidity com- monality	25
3.2.3	Evolution of the commonality in liquidity across periods	29
4	Data	31
4.1	Description of the datasets	31
4.2	Preparatory analysis	32
4.3	Data cleaning	34
5	Empirical analysis	37
5.1	Results : commonality in market quality	37
5.1.1	Analysis of the results for the whole period	38
5.1.2	Comparison between the beginning and the end of the period .	41
5.1.3	Comparisons before and after the introduction of Bitcoin futures	41
5.1.4	Comparison before and after the end of the Bitcoin bubble . . .	41
5.1.5	Results of robustness checks	42
5.2	Results : impacts of volatility regimes on liquidity commonality	42
5.2.1	Summary statistics for the proxies	42
5.2.2	Market-model regressions disregarding volatility regimes . . .	44

5.2.3	Market-model regressions conditioned on volatility regimes . .	46
5.3	Results : evolution of liquidity commonality	46
6	Conclusion	53
A	Proxies used by Marsh and Mazza (2017) to capture market quality	57
A.1	Cost of immediacy	57
A.2	Adverse selection	57
A.3	Liquidity	58
A.4	Price efficiency	59
A.5	Volatility	59
B	Description of the variables of the datasets	61
B.1	Trade datasets	61
B.2	Order Book	61
C	Description of our set of cryptocurrencies	65
C.1	Bitcoin (BTC)	65
C.2	Bitcoin Cash (BCH)	66
C.3	Eos (EOS)	66
C.4	Ethereum (ETH)	68
C.5	Ripple (XRP)	68
D	Most significant time skips in the order book datasets	71
E	Summary statistics of the trade and order books datasets	77
F	Descriptive statistics of the market quality proxies on 30 minutes intervals	81
G	Robustness check : non-standardized proxies	85
H	Robustness check : mean as aggregation rule	89
I	Weight variables - Statistics	93

J	Evolution of daily and weekly Adj. R-squared by periods	101
K	Results of the statistical tests of means	107
	Bibliography	111

List of Figures

5.1 Average correlations between proxies	44
C.1 BTC market capitalization	66
C.2 BCH market capitalization	67
C.3 EOS market capitalization	67
C.4 ETH market capitalization	68
C.5 XRP market capitalization	69
D.1 Significant time skip in BTC order book	71
D.2 Significant time skip in BCH order book	72
D.3 Significant time skip in EOS order book	72
D.4 First significant time skip in XRP order book	73
D.5 Second significant time skip in XRP order book	73
D.6 Third significant time skip in XRP order book	74
D.7 First significant time skip in ETH order book	74
D.8 Second significant time skip in ETH order book	75
I.1 BCH 's DEPTH - main statistics	94
I.2 BTC 's DEPTH - main statistics	94
I.3 EOS ' DEPTH - main statistics	95
I.4 ETH 's DEPTH - main statistics	95
I.5 XRP 's DEPTH - main statistics	95
I.6 BCH 's quote-duration - main statistics	96
I.7 BTC 's quote-duration - main statistics	96
I.8 EOS ' quote-duration - main statistics	97

I.9	ETH 's quote-duration - main statistics	97
I.10	XRP 's quote-duration - main statistics	98
I.11	BCH 's traded amount - main statistics	98
I.12	BTC 's traded amount - main statistics	99
I.13	EOS ' traded amount - main statistics	99
I.14	ETH 's traded amount - main statistics	99
I.15	XRP 's traded amount - main statistics	100

List of Tables

4.1	Observations in the trade datasets for each cryptocurrency	32
4.2	Observations in the order book datasets for each cryptocurrency	32
4.3	Observations in the trade datasets inside the common interval for each cryptocurrency after cleaning	36
4.4	Observations in the order book datasets inside the common interval for each cryptocurrency after cleaning	36
5.1	Explanatory power (in %) of the five first components of the PCA for the whole period and with standardized proxies on 30 minutes intervals	38
5.2	Explanatory power (in %) of the five first components of the PCA for the three first and three last months of the period and with standardized proxies on 30 minutes intervals	38
5.3	Explanatory power (in %) of the five first components of the PCA for the three months before and after the introduction of futures on Bitcoin and with standardized proxies on 30 minutes intervals	39
5.4	Explanatory power (in %) of the five first components of the PCA for the month preceding and following the end of BTC bubble and with standardized proxies on 30 minutes intervals	40
5.5	Aggregate liquidity measures - Summary Statistics	43
5.6	Aggregated standardized liquidity measures in low volatility regimes : Summary Statistics	45
5.7	Aggregated standardized liquidity measures in medium volatility regimes : Summary Statistics	45
5.8	Aggregated standardized liquidity measures in high volatility regimes : Summary Statistics	45
5.9	Liquidity co-movements - Unconditioned upon volatility regime	48
5.10	Liquidity co-movements - Unconditioned upon volatility regime	49
5.11	Liquidity co-movements - Conditioned upon volatility regime	50
5.12	Liquidity co-movements - Conditioned upon volatility regime	51

E.1	Summary statistics of the trades for the five cryptocurrencies	78
E.2	Summary statistics of the first limit ask price and quantity for the five cryptocurrencies	79
E.3	Summary statistics of the first limit bid price and quantity for the five cryptocurrencies	80
F.1	Summary statistics of the proxies for market quality on 30 minutes intervals (1)	82
F.2	Summary statistics of the proxies for market quality on 30 minutes intervals (2)	83
G.1	Explanatory power (in %) of the five first components of the PCA for the whole period and with non-standardized proxies on 30 minutes intervals	85
G.2	Explanatory power (in %) of the five first components of the PCA for the three first and three last months of the period and with non-standardized proxies on 30 minutes intervals	86
G.3	Explanatory power (in %) of the five first components of the PCA for the three months before and after the introduction of futures on Bitcoin and with non-standardized proxies on 30 minutes intervals	86
G.4	Explanatory power (in %) of the five first components of the PCA for the month preceding and following the end of BTC bubble and with non-standardized proxies on 30 minutes intervals	87
H.1	Explanatory power (in %) of the five first components of the PCA for the whole period, with standardized proxies on 30 minutes intervals and with the mean as aggregation rule	89
H.2	Explanatory power (in %) of the five first components of the PCA for the three first and three last months of the period, with non standardized proxies on 30 minutes intervals and with the mean as aggregation rule	90
H.3	Explanatory power (in %) of the five first components of the PCA for the three months before and after the introduction of futures on Bitcoin, with standardized proxies on 30 minutes intervals and with the mean as aggregation rule	91
H.4	Explanatory power (in %) of the five first components of the PCA for the month preceding and following the end of BTC bubble, with standardized proxies on 30 minutes intervals and mean as aggregation rule	92
J.1	Comparison of the average daily adjusted R^2 by periods (1)	102
J.2	Comparison of the average daily adjusted R^2 by periods (2)	103

J.3	Comparison of the average weekly adjusted R^2 by periods (1)	104
J.4	Comparison of the average weekly adjusted R^2 by periods (2)	105
K.1	Results of the tests of means between the average daily adjusted- R^2 by periods	108
K.2	Results of the tests of means between the average weekly adjusted- R^2 by periods	109

Chapter 1

Introduction

During the last decade, cryptocurrencies¹ have become more and more popular, at least as financial assets. Proof of this can be shown by simply looking at the increasing number of cryptocurrencies, the volume of cryptocurrencies traded, the evolution of their market capitalization and price over the last ten years. To give one concrete example, on a period going from October 2016 to October 2017, the market capitalization of Bitcoin (the first decentralized digital currency) went up from \$10.1 to \$79.7 billion whereas its unit price increased from \$616 to \$4 800 US dollars, representing a 680% year-on-year return on investment (Corbet, Lucey, Urquhart, and Yarovaya, 2019). Today there are more than one thousand cryptocurrencies, but Bitcoin still remains the leader on this market. According to the review of literature on cryptocurrencies of Corbet et al. (2019), some elements explaining the growth in the demand for cryptocurrencies are the facts that they have low transaction costs, rely on a peer-to-peer system and do not include governments in their design. This rising popularity also spread to the academic world, where cryptocurrencies became a recurrent research topic. At the time when Corbet et al. (2019) wrote their literature review, they counted 24 papers with significant quotes available on the subject. The specific areas of research of the said papers were common research questions from traditional finance like volatility clustering, asset pricing bubbles, market efficiency, etc. More recently, Facebook initiated a new cryptocurrency project in association with other companies and NGO's called Libra which attracted even more attention to cryptocurrencies and has the potential to change the landscape of the cryptocurrency market again.

In the past few years, cryptocurrencies have been increasingly more used as a tool to raise funds. Sockin and Xiong (2018) point out that in the 2015-2017 period, more than 2 000 Initial Coin Offerings² (ICOs) managed to raise over \$4 billion from the public (according to their research, the top three ICOs in 2017 were Tezos, EOS.IO and BANCOR, each having respectively raised \$208 million, \$200 million, and \$153 million). A study from EY (2017) reported that this amount even exceeded venture capital investments in funding innovative projects related to blockchain technology. Another interesting fact is the geographic repartition of those ICO's. Among the more than 2 000 ICOs launched between 2015 and 2017 discussed by Sockin and Xiong (2018), 1031 were launched in the U.S., 310 in Russia, 260 in Singapore, 256 in China and 196 in Hong Kong. But even though cryptocurrencies can represent a

¹"Cryptocurrencies are peer-to-peer electronic cash systems which allow online payments to be sent directly from one party to another without going through a financial institution" (Corbet et al., 2019, p.3). Academic literature about cryptocurrencies is addressed in section 2.1

²Type of funding using cryptocurrencies

successful tool to raise money, failed ICO's showed us the risks and abuses that can emerge by investing in those assets. Sockin and Xiong (2018) highlight in this way the fact that we still have a need to better understand this technology and have a better assessment of the benefits and risks that their use involves building an adapted regulatory framework. This need for a deeper knowledge has become even more relevant recently with the introduction of Bitcoin futures³. According to the Futures Industry Association, the late introduction of futures on Bitcoin opens financial markets in a way that was never done before and has a high potential to disrupt the economy (Ryznar, 2018). To be precise, Ryznar (2018) refers to the launch of Bitcoin futures trade by the Chicago Board Option Exchange (CBOE) the 10th of December 2017 and by the CME group a week later.

Another subject lying at the core of our master thesis is market quality⁴. Market quality is important for financial markets as it is related to concepts essential for them, such as market efficiency or transaction costs (Marsh and Mazza, 2017). Several facets of market quality have, much like cryptocurrencies, known a growing interest in the recent academic literature. Taking a step back, we can see that a lot of researches in the past 20 years have studied co-movements between several aspects of financial assets (Adjaoute and Danthine, 2001 ; Brockman and Chung, 2002 ; Chordia, Roll, and Subrahmanyam, 2000 ; Domowitz, Hansch, and Wang, 2005 ; Hasbrouck and Seppi, 2001 ; heston1994does ; Heston and Rouwenhorst, 1995 ; Karolyi, Lee, and Van Dijk, 2012). Among them, some focus on co-movements between dimensions that are inherent parts of the concept of market quality, like the literature on liquidity co-movements for example (Brockman and Chung, 2002 ; Brunnermeier and Pedersen, 2009 ; Chordia, Roll, and Subrahmanyam, 2000 ; Domowitz, Hansch, and Wang, 2005 ; Hasbrouck and Seppi, 2001 ; Huberman and Halka, 2001 ; Karolyi, Lee, and Van Dijk, 2012). The presence of a common factor in the movements of the market quality of financial assets is something that policymakers and regulators need to consider to manage their liquidity successfully. Indeed, stronger co-variation between assets would make the propagation of liquidity shocks to other markets easier and could have a significant impact on the stability of financial markets (Klein and Song, 2017). According to Klein and Song (2017), to be aware of this common factor is even more important in times of crisis, as it implies greater transaction costs and an inability to trade assets quickly without significantly impacting prices. In this regard, the financial crisis of 2008 revealed the adverse effects that such common component can have on diversification and was the start of the realization that there was a need for regulatory authorities to reflect those common variations in their risk models (Beaupain, Giot, and Petitjean, 2010). However, a common factor in market quality movements has implications for other actors than policymakers and regulators. The liquidity of a stock and the associated elements that can lead to liquidity risks are important parts of investment decisions (Koch, Ruenzi, and Starks, 2016). In the presence of a common factor in market quality movements, the likelihood of facing systematic liquidity risk and subsequently having to face undiversifiable transaction costs increases (Anagnostidis and Fontaine, 2018). A better knowledge of this factor is thus important as it would change the

³"Futures are contracts to buy or sell a specific quantity of some asset at a specified price and date in the future. In other words, futures are simply contracts for the future delivery of the underlying asset at a certain price" (Ryznar, 2018, p.8). The introduction of Bitcoin futures on financial markets is addressed in section 2.1.6

⁴"Market quality is defined as "a market's ability to meet its dual goals of liquidity and price discovery" (O'Hara and Ye, 2011, p. 463). Academic literature on market quality is addressed in section 2.2"

current paradigm of asset pricing and impact investor's trading strategies.

However, to the best of our knowledge, there are still no article in the academic literature as of today treating the subject of co-movements between the market quality of cryptocurrencies. A few months after we started working on our master thesis, a new article about the Bitcoin's liquidity was released (Marshall, Nguyen, and Visaltanachoti, 2019) and another one about the factor structure in the returns and volatility of cryptocurrencies (Liu et al., 2019) which gets very close to the topic of our work. Nonetheless, liquidity and volatility are not the sole components of market quality and Bitcoin is only one cryptocurrency among many, so academic literature is still lacking a global analysis of market quality in cryptocurrencies. Furthermore, even in the literature treating market quality co-movements of more traditional financial assets, there was no comprehensive analysis of market quality in academic researches before the article of Marsh and Mazza (2017) about common factors in the market quality of European stocks. Most of the body of the literature on this topic is in fact focused on subdimensions of market quality, like liquidity, but those aspects do not depict a complete picture of the concept of market quality by themselves. Moreover, if we take the most documented dimension, commonality in liquidity, despite the number of research papers in this field, there still exists no consensus on the best way to measure liquidity or when there is a consensus on a measure, the necessary data to compute it is sometimes missing, according to Marsh and Mazza (2017).

Through this short introduction on the stakes of the topics of cryptocurrencies and market quality and the state of the academic literature of those respective subjects at the time we wrote this report, highlighted the interest of conducting more research in those areas. In this context, our motivation is to contribute to both the literature on cryptocurrencies and market quality with our master thesis by extending and updating the main findings of the market quality literature to the specific case of cryptocurrency instead of traditional financial assets. The methodology of our approach consists mostly in replicating academic research papers by using cryptocurrency data instead of data from traditional financial assets. The problematic of our master thesis could then be phrased like this:

“What findings of the academic literature about the market quality co-movements of traditional assets can be transferred to cryptocurrencies in 2017/2018?”

Before we move on to the plan of the rest of our work, we would like to clarify the scope of our research. One of the main challenges that we faced in our project was the lack of uniformity between authors in the practical part of the market quality literature. Properly speaking, there are few articles that take a global approach and definition of market quality. Academic studies focus instead more on liquidity or other dimensions of market quality. In the experimental part of our work, we will replicate one article focused on liquidity among others. The limit of this approach is that findings that were true for liquidity might not necessarily be true for market quality. However, liquidity being an important part of market quality, we are convinced that this approach is a valid intuition of the market quality behaviour.

The rest of our master's thesis is divided in six main chapters. Chapter 2 is a detailed review of the academic literature on the subject of cryptocurrencies and

market quality. We present a summarized view of the main findings and define key concepts for the rest of our work. In Chapter 3, we formulate research questions and hypothesis that we need to answer to address our problematic in light of our literature review. We also explain and justify the methodology we chose to tackle them. In Chapter 4, we describe our datasets, elaborate the cleaning process of our datasets and make a small preparatory analysis. Next in Chapter 5, we present the results of the empirical part of our project, answer the hypothesis and research questions and discuss them. Finally, in Chapter 6, we make a concise review of the results we obtained and take a critical viewpoint of our master's thesis by presenting the limits of our approach and areas where we think that further research should be conducted.

Chapter 2

Literature review

The objective of this Chapter is twofold. The first is to define the fundamental concepts necessary to understand our problematic and make our approach as clear as possible. Secondly, as detailed in Chapter 1, our problematic consists in updating, verifying, if the main findings of the market quality co-movements literature to are holds in for cryptocurrencies. To achieve this purpose, we review the academic literature on market quality and on cryptocurrencies to summarize the main findings of those research and present them in a concise way.

Before moving on to the structure of the Chapter, we would like to make an important remark regarding the scope of our literature review. Our review is focused on the academic research relevant in order to address our problematic. The body of academic literature on cryptocurrencies and market quality extends beyond the articles that we will treat in this Chapter. For example, we ignored the research treating the monetary function of cryptocurrencies (or lack thereof), as it does not concern common movements in market quality. Concretely, this chapter is divided into three sections. Section 2.1 is a review of the relevant academic literature on cryptocurrencies and section 2.2 on market quality. Finally, we report the results of two recent papers researching common movements in some dimensions of market quality of cryptocurrencies in section 2.3.

2.1 Cryptocurrency literature

2.1.1 Definitions and key concepts

A cryptocurrency is an electronic cash system that permits payments from one peer to another directly without the need for any financial intermediary. Contrary to other financial assets, a cryptocurrency does not have a physical counterpart as it does not depend on any higher authority like the government and can be divided infinitely. Furthermore, its value is based on the traceability and security of an algorithm, rather than on any other physical asset. Hence its quick growth in popularity and use, which translated in a rise in their prices, their volatility and trading volumes that we mentioned in our introduction (Corbet et al., 2019). In their model, Sockin and Xiong (2018) conceive cryptocurrencies as the subscription to a platform whose goal is to help in the bilateralisation and decentralization of transactions of certain goods and services, executed by households and by means of blockchain

technology. In this model, the cryptocurrency has two functions: one is to serve as an initial financing for the platform and the other is to materialize the membership to the platform.

Similar to classical currencies, cryptocurrencies rely on people's trust in its value and on the belief that it will be accepted by vendors (Ryznar, 2018). On the other hand, a cryptocurrency differs from fiat money¹ through the way it is issued : fiat money is issued by the central bank of a country whereas Bitcoin (for example) is issued by electronic means by a computer program with a predetermined cap (Ryznar, 2018). In this sense, Ryznar (2018) deems the subject of whether bitcoin possesses the basic functions necessary to be a currency (stored value and fungibility) or not is debatable and concludes that it can't be valued using fundamental analysis.

To validate transactions, a cryptocurrency relies on blockchain technology, which can be described in the following way :

"At its root, a blockchain is a public ledger that everyone can inspect, but which no single user controls. The participants in a given blockchain system work together to keep the ledger updated; it may be amended only by strict rules and consensus". (Ryznar, 2018, p.7)

The system uses a hashing algorithm² to check if a transaction will be approved and once a transaction is on the blockchain it is extremely difficult to alter it (Ryznar, 2018). We do not go further into details about the inner workings of the hashing algorithms or about blockchain as it is not relevant to answer our problematic.

2.1.2 Regulation

As of today, there is still no uniformity between countries with respect to the regulation of cryptocurrencies. Hence, some of the most common-spread arguments against the use of cryptocurrencies as currencies are their use as a tool for cybercriminality and their unknown potential to have destabilizing consequences on the world economies (Corbet et al., 2019). Despite this lack of common regulation, we would at least expect uniformity and price consistency across the cryptocurrency exchanges across the world, irrespective of whether they are considered a financial asset or a currency. In the case of Bitcoin however, it appears that some Bitcoin exchanges fail to achieve the law of one price and we can observe different prices for Bitcoin across different exchanges (Pieters and Vivanco, 2017). The reason for such failure can only come from the intrinsic characteristics of the exchanges themselves.

What about the future of cryptocurrencies regulation then? According to Clements (2018) cryptocurrencies are likely to stay around for quite some time and so are fiat currencies. The author foresees a high probability for the distributed ledger technology to become mainstream to enhance payments, settlements and

¹"which is a currency that a government has declared to be legal tender, but not necessarily backed by a physical commodity"(Ryznar, 2018)

²Algorithm mapping an arbitrary size of data to a hash of a fixed size. The mapping is only in one-way and impossible to revert

clearing but will also affect other domains such as data protection, online voting and supply chain management to only name a few. In reaction, new legislation will materialize in each of those domains and they will eventually merge.

2.1.3 Pricing bubbles

The major increases of the Bitcoin price have been said to have bubble-like properties in academic literature Corbet, Lucey, and Yarovaya (2018). In their study, the authors showed the existence of short-term time periods in which block size, liquidity and hash rate³ of Bitcoin and Ethereum affect the prices of these cryptocurrencies. These episodes of extreme volatility do not last for long periods and are often associated with cyber-criminality and regulatory disorientation.

A lot of research in academic literature have tried to determine if the Bitcoin price was in a bubble in the period were it reached sky-high levels(Cheung, Roca, and Su (2015) ; Clements (2018) ; Corbet, Lucey, and Yarovaya (2018) ; Liu et al., 2019 ; Su et al. (2018)). Corbet, Lucey, and Yarovaya (2018) present evidences that Bitcoin was in a Bubble at the end of 2017 and Cheung, Roca, and Su (2015) showed that Bitcoin has known several short bubbles between 2010 and 2014. Cheung, Roca, and Su (2015) also showed the existence of speculative pricing bubbles in Bitcoin and supplied further evidences. By using forward recursive regressions to compute Dickey-Fuller t statistics and comparing them to the critical value of the test defined by Phillips, Wu, and Yu (2011) and Phillips and Yu (2011), Liu et al. (2019) managed to time the late Bitcoin bubble. They show that the bubble started the 24th of May 2017 and ended the 28th of January. Furthermore, using the GSADF test⁴ proposed by Phillips, Shi, and Yu (2015), Su et al. (2018) identify four massive bubbles. The ADF test, or augmented Dickey-Fuller test is a test to see if a time-series has stationary properties. The appearance of bubbles coincides with periods with an increase in price and contrarily, falling prices lead to the burst of these bubbles. This is consistent with Gürkaynak (2008)'s model. In the model, the initiation of a bubble comes from an exogenous shock, whether domestic or foreign. The idea is that a situation of major financial distress can turn into a global bubble when investors lose their trust in government-backed currencies. In this scenario, they will rather buy other types of assets such as cryptocurrencies, causing an appreciation of their price. On the other hand, domestic levers like the expected depreciation of the local currency or changes in the currency regime lead to bubbles that are not likely to contaminate other markets in other countries. In other words, investors hold this alternative asset as a hedging instrument to protect themselves against potential risks or simply use it for speculative purposes. In the end, the bubbles would burst due to intervention of the authorities. Nonetheless, bubbles involve real risks of spill overs, destabilizing other financial markets.

³Rate at which transactions are approved

⁴"The GSADF test repeatedly executes a series sample sequence based on the ADF test"(Phillips, Shi, and Yu, 2015, p.59)

2.1.4 Specificity as an asset class

At the end of the previous section, we mentioned how some investors used cryptocurrencies for hedging their positions or for pure speculation. Exploiting daily and weekly data in a dynamic conditional correlation model Engle (2002), Bouri, Azzi, and Dyhrberg (2016) find that most of the time, Bitcoin can be used as an effective diversification instrument. However, their work indicates that in the case of Bitcoin, there is not sufficient empirical evidence on these diversifications and hedging qualities relatively to other assets, especially against the main bonds, stocks, gold, oil, global commodity index and the US Dollar Index. Bitcoin has shown hedging and safe haven features that deviated between horizons only in a handful of cases.

In spite of the relevance of these results, Bouri, Azzi, and Dyhrberg (2016) warn us against investing in Bitcoin given its specific asset-class properties. First, Bitcoin investments are considerably less liquid than other classical financial assets, although their accessibility for private investors can be tremendously enhanced with the potential rise of corresponding funds and derivatives. Secondly, the turbulent nature of the sample period of the study, where Bitcoin prices displayed high volatility, might suggest that the diversification ability of Bitcoin varies with time.

According to Sockin and Xiong (2018), we can use different valuation approaches and risk attributes to cryptocurrencies given the way we define it (as a medium transaction, as a commodity or as a security). Their research suggests that for the valuation of these cryptocurrencies, we can get helpful insight from the analysis of some quantitative measures like the volume of transactions recorded on their ledgers. The volumes and price of the cryptocurrencies can serve as valid signals of their underlying demands.

2.1.5 Market efficiency

In this subsection, we will first introduce some of the most impactful market efficiency models in academic literature and present cryptocurrencies-related results in a second time.

We begin with the work of Subrahmanyam (1991), who introduced an adverse selection model of market micro-structure where the agents are averse to risk. This contrasts the former models of Kyle (1984) and Admati and Pfleiderer (1988), where market makers and traders are both risk neutral. In the risk-neutral case and when traders are informed, we observe highly correlated signals and a rise in their number diminishes their aggregate expected profits because of the higher competition, resulting in a more liquid market. However, given risk aversion (of the informed traders), a rise in their number can inflate their aggregated expected profits, resulting in a less liquid market. Assuming the acquisition of information is endogenous, higher liquidity trading can hinder the liquidity of the market too by the increase in the number of traders that enter the market. Also, they showed that making private information more precise can lead to improved market liquidity due to a more aggressive competition coming from the risk averse and more "precisely" informed traders. On another note, Subrahmanyam (1991)'s model implies that price efficiency may be inversely proportional to the amount of liquidity trading in the market. Given that information acquisition is not solely endogenous in

security markets due to the presence of insiders. The latter might deter other traders from acquiring costly information and trading (Fishman and Hagerty, 1989).

Next comes the work of Hasbrouck (2002), who simulated different economies : a two-market Roll model (Roll, 1984), two markets with private information and a model of two markets with public and private information. In each setup he tested 2 approaches in their ability to fit the initial structural determination of prices. These approaches are PT/GG (permanent-transitory Granger-Gonzalo) and IS (information shares). Results determined PT/GG and IS approaches did not systematically converge precisely towards the structural determination. The confidence interval resulting from the IS approach contained the true value whereas the PT/GG approach did not. The latter approach sometimes leads to a precise identification when the former approach yields a wide confidence interval. However, the price factor obtained through PT/GG in this case is characterized by a higher serial correlation and volatility. Overall, the IS method seems to bear a more meaningful inference.

Now moving on to cryptocurrencies, Urquhart (2016) tried to find autocorrelations, random walk behaviour, nonlinearities and long-range dependence in Bitcoin returns over the period of 2010 to 2016. This full sample is then split into two sub-samples: before and after 2014. Firstly, the author checks for autocorrelation of returns via a Ljung-Box test (Ljung and Box, 1978) (with H_0 : no autocorrelation). Secondly, he performed runs tests (Wald and Wolfowitz, 1940 and Bartels (1982)) in order to check for independence of returns (with H_0 : independence). Thirdly, he performs a variance ratio test (Lo and MacKinlay, 1988) where under H_0 : the price follows a random walk and the variance of the price difference of order q is equal to p times the variance of the first difference. The variance ratio test parameters p and q are determined automatically via an automatic variance (AVR) test (Choi, 1999) and a wild-bootstrapped AVR test (Kim, 2009) which is even more appropriate for small samples. Fourthly, he uses the BDS (Brock and al., 1996) test, this is a popular non-parametric test for autocorrelation in stock returns. The H_0 is that the data generating processes are i.i.d.. His analysis showed that the Bitcoin market is not weakly efficient when considering the full sample. However, he showed Bitcoin seems to become more efficient, especially when considering the second sub-sample. Nonetheless, these results showed that the inefficiency of Bitcoin remains strong.

Bariviera (2017) improves and completes previous literature on Bitcoin by studying long-memory effect via the Hurst exponent in more detail in three main points: (i) it makes use of the De-trended Fluctuation Analysis method, rather than the R/S method which presents a bias to finding long-range persistence in all the time series. DFA appears to be much less sensitive to deviations from stationarity. (ii) it proposes sliding windows, to dynamically assess the change of informational efficiency with respect to time (iii) it takes into consideration the long-term memory in the daily volatility of returns, which can be used as a proxy for the risk of this market. He found that daily returns experienced a structural change. From 2011 to 2014 the series of returns was mainly persistent ($H > 0.5$), but after, the series appears to be compatible with a white noise. However, daily volatility (defined as the log-difference between daily maximum and minimum price) showed a persistent behaviour for the entire sample period. Also, the long-range dependence magnitude of daily volatility is higher than in daily returns. This particularity seems to indicate an intrinsic characteristic of this synthetic currency market. That is, volatility clustering is a main feature of the Bitcoin market.

Cheah et al. (2018) added that this long-term dependence, for which they had found significant evidence across markets, could provoke a 'system failure' if new regulations were to be introduced in the financial markets. Their results indicate that Bitcoin markets are average to highly inefficient where the estimated memory in prices can aid investors get profits. They also showed through the newly developed fractionally cointegrated VAR⁵ model of Johansen and Nielsen (2012) that contrary to the general thought that the Bitcoin market is virtually detached from the real economy, it is in fact inherently dependent to the dynamics of the real economy. The observed inefficiency of Bitcoin could thus be, to some extent, regulated and interconnected to the real economic system.

Wei (2018) found that, in the case of cryptocurrencies, efficiency is higher and volatility lower in high-liquidity markets as traders are more prone to take advantage of arbitrage opportunities if returns show signs of predictability. Greater transaction costs in markets where turnover is low affects the trader's reaction speed and readiness, resulting in a more inefficient market. Moreover, illiquid cryptocurrencies showed strong return negative-persistence (low Hurst exponent). These results suggest that even though major cryptocurrencies are improving as far as efficiency is concerned, emerging altcoins with fairly low liquidity keep struggling.

Ji et al. (2019) also studied the efficiency of two Bitcoin price indices (USD and CNY), in their papers, they use a more complex metric: the Efficiency Index, which incorporates effects of long-memory, approximate entropy and fractal dimension. The time series spans from foundation of the two indices until July 2017. Starting with the USD market, he observed there were just two longer periods of time in which the market was efficient – from mid-2011 to mid-2012, and between March and November 2014. Taken these two periods out of consideration, market efficiency was rejected on a regular basis. However, the results are not as clear for the CNY market as the examined time frame misses some major bubble-like dynamics pre-2014. Overall, the only efficient periods were those of "cooling off", just after prices skyrocketed. Other than these, the Bitcoin markets were clearly inefficient and therefore, they were predictable. This paper showed the specific contributions of the separate parts of the Efficient Index towards inefficiency. The division between each component of the index can be seen as the contribution of global correlations, local correlations, and complex correlations. For the USD Bitcoin index market, we observe some changes in the structural composition of inefficiency. Up to 2014, entropy (thus complex correlations) played a major role which became less and less important in recent years. Regarding the global correlations, their contribution always stayed relatively the same over all periods, and the original proportion of entropy has faded in favour of fractal dimension. In the last years, inefficiency has been essentially composed of the local correlations, in other words short-term peaks and troughs. Ji et al. (2019)'s results favour inefficiency. One explanation they propose is that the market is still very shallow and thus it does not attract big players, so there is no real deepening of the market. Although, this is likely to change with the introduction of Bitcoin-based derivatives and assets tracking Bitcoin price. Still, the liquidity of Bitcoin markets remains low relative to the equity or currency markets. Besides, the Bitcoin price is still strongly dependent on exogenous factors related to the legal and security domains. The nice features that are low/no regulations and decentralization come at considerable cost. In addition, conditions vary quite a bit depending on the exchange market, that is, there can be costly transaction fees (i.e.

⁵Value at Risk

withdrawal fees) or (in)ability to pull the funds and turn them into the fiat currency. Another difficulty in computing the effective profits comes when it's time account for taxation. Indeed, countries treat this matter differently and the debate about the value-added tax and income tax on cryptocurrencies is still continuing.

Sensoy (2019) studied the weak-form efficiency of Bitcoin prices at high-frequency (intraday prices) with respect to euro and US dollar. In his paper, he performed permutation entropy using a rolling window for his test. The finding can be regrouped in four main points. First, both markets have seen improvements in efficiency at the intraday level since 2016, this supports and completes the works of Urquhart (2016) and Bariviera (2017) on the recent improvement in efficiency at the daily level. Nevertheless, the results of Sensoy (2019) show that this improvement follows a cyclical structure in the case of BitcoinUSD (Bitcoin paired to USD). On the contrary, the increase is more progressive in the case of BitcoinEUR (Bitcoin paired to euro). Second, BitcoinUSD market barely shows to be more efficient than BitcoinEUR market since 2016. This could indicate that on the BitcoinEUR market one could make more profit out of trading at the intraday level. Third, irrespective of the choice of the currency, the higher the frequency the lower the efficiency. Fourth, liquidity (volatility) has a substantial positive (negative) influence on the informational efficiency of Bitcoin prices.

Tiwari et al. (2018) use estimators with both algorithmic efficacy and constructed efficiency index and show that Bitcoin market is efficient with some exception to the timeframe from April to August 2013 and from August to November 2016.

Finally, Vidal-Tomás and Ibañez (2018) discuss the semi-strong efficiency of Bitcoin in the Bitstamp and Mt.Gox markets, showing how the cryptocurrency reacts to monetary policy and to Bitcoin events. They noted that, with time, Bitcoin became more efficient with regard to its own events. Alternatively, Bitcoin is independent of monetary policy news. Hence, there should not be any kind of control on the cryptocurrency.

2.1.6 Introduction of futures

For traditional financial assets, stock futures let traders have positions with leverage so we would expect most informed traders be attracted by this feature. Despite the intuition, there is considerable evidence that the information share of the stock futures is remarkably limited and the stock spot market seems to play an important function in price discovery. Aggarwal et al. (2011) explored this phenomenon and made two major observations. First, the liquidity of the stock futures plays a major role in determining price discovery. The securities where the spot market plays a great role usually are the securities with the least liquid stock futures. The increased transactions costs seem to offset the benefits of leveraged trading. The second observation is that in situations where prices make large movements, stock futures seem to play a significant role in those movements.

In December 2017, both the Chicago Board Options Exchange and the Chicago Mercantile Exchange introduced futures on Bitcoin. Baur, Dimpfl, and Kuck (2018) researched to what degree this introduction offers useful information for the price

discovery of Bitcoin. Based on Hasbrouck (1995) and Gonzalo and Granger (1995) information share methodology they found that the spot price guides the futures price. This is the opposite for the majority of financial instruments as most of the time, it is the futures price that guides the spot price.

Bitcoin futures present new levels of risk for the economy, hence the necessity to regulate them. Ryznar (2018) investigated the potential legislation for Bitcoin futures and concluded that if Bitcoin futures trading were to proceed, it should be supervised. Meanwhile, some observers believe Bitcoin will collapse. Undoubtedly, the cryptocurrency has its share of flaws that could hinder its development, especially transaction processing takes a long time, the fees are quite significant, and volatility is rather high. Anyway, lessons from Bitcoin are transferable to other cryptocurrencies and Ryznar (2018) recommends prudence. Though the products are new, their risks are well-known. These familiar hazards should be tackled to prevent any financial distress.

2.1.7 Other aspects of cryptocurrencies

The promise of financial privacy has greatly contributed to the success of cryptocurrencies. This being said, some works that have investigated this topic suggest this might be compromised. Reid and Harrigan (2013) explained that through the use of an adequate network representation, one could associate many public-keys with each other's, and further associate these with external sources to identity information. With right tools, the full activity of users could be described in detail. This can be achieved via passive analysis alone. Even more, with active analyses, where an interested party can deploy marked Bitcoin and interact with other users can be leveraged to increase the level of information. The authors think that large centralized services are able to recognize and to track substantial portions of user activity. Technical members of the Bitcoin community have answered that strong anonymity is not a primary design objective of the system. However, the average user must be aware of this, especially when sending Bitcoins to users and to institutions with whom they would rather not to be openly associated with. Androulaki et al. (2013) simulated Bitcoin transactions in day-to-day use in a university setting and found that 40% of profiles could be recovered, and the users clearly identified even when following privacy good practices. Biryukov and Pustogarov (2015) even show a profound anonymity failure, it appears to be technically possible to match the addresses of Bitcoin users with their actual IP addresses.

2.2 Market quality

Market quality is defined as "a market's ability to meet its dual goals of liquidity and price discovery" (O'Hara and Ye, 2011 p. 463). In other words, a more liquid market in which transaction costs are lower and where price discovery is more efficient in comparison to other markets is considered of higher quality (O'Hara and Ye, 2011). According to Marsh and Mazza (2017), whereas the definition of the concept is clear and accepted, as of today there is still no consensus in the academic world about the best way to measure market quality (or when there is a consensus,

the data to compute the measure is not always available⁶). We have observed this lack of uniformity ourselves while doing our literature review. First, most of the academic research generally focus on one of the dimensions of market quality like liquidity rather than adopting a general definition of the concept. It is therefore difficult to define a set of proxies that can capture all the facets of market quality and the second element of divergence resides in the choice of proxies to measure the dimensions of market quality. They vary from one study to another either because a study uses intraday data and another daily data, or either by preference of an author for one proxy over another. In our opinion, the most general approach used to capture market quality available in the current literature is the one proposed by Marsh and Mazza (2017). In their study, they measure market quality through a large set of proxies designed around five dimensions: adverse selection, cost of immediacy, liquidity, price efficiency and volatility.

Adverse selection

When asymmetric information affects market participation. Normally, traders meet randomly in the market to trade assets but with adverse selection, only the current asset owner can observe its quality, hence making trading more difficult (Chiu and Koepl, 2016).

Cost of immediacy

Can be defined as the return of transaction from the dealer's perspective, in other words the cost in order that the dealer accepts a quick shift in his inventory (Dick-Nielsen and Rossi, 2018).

Liquidity

The degree to which transactions happen at short notice and fair price (Dick-Nielsen and Rossi, 2018).

Price efficiency

Theory that price reflects complete information from all the market participants. Statistically, efficient prices should follow a Random Walk process (Dodd and Gilbert, 2016).

Volatility

The Volatility of the price of an asset is a statistical measure of the dispersion of the asset's price.

⁶Some research use daily data and measures while other use intraday data

Now that we have introduced the concept of market quality, we will review the main findings of related academic research in the following subsections. First in sub2.2.1, we will go over the literature searching the presence of a common factor in market quality movements. We overview the sources of this commonality factor identified in academic literature and the implications of commonality on asset pricing. Then in subsection 2.2.2, our topic is the impact of high-frequency trading on the commonality factor, a topic emerging in recent research. Finally, we discuss the influence of volatility regime market capitalization on commonality in subsection 2.2.3.

2.2.1 Commonality

We refer to commonality as the tendency of individual assets to move together in aggregate Marsh and Mazza (2017). Several aspects of individual shares have been proven to exhibit commonality. One of the first research conducted on the topic was the study of Heston and Rouwenhorst (1994), where the authors proved that returns of stocks of the euro-area contained common components. However, we will only interest ourselves to the part of the commonality literature focusing on market quality. By extension to the definitions of commonality and liquidity, “Liquidity commonality refers to market-wide liquidity movements” (Anagnostidis and Fontaine, 2018, p.1). Chordia, Roll, and Subrahmanyam (2000) were among the first to interest themselves in liquidity co-movements. In their study, they have shown a high commonality degree in individual U.S securities’ liquidity. After them, multiple other researches have also concluded in the presence of a common factor in liquidity movements both at the exchange global level for different types of market structures. Other significant findings were the fact that commonality in liquidity can be observed in an order-driven market structure (Brockman and Chung, 2002) and is stronger in times of financial crisis (Hameed, Kang, and Viswanathan, 2010). Overviewing this stretch of the literature as a whole, Anagnostidis and Fontaine (2018) make three important observations. First is the fact that so far, a large part of the research has focused on liquidity commonality only inside financial markets. Another point is the fact that most of the studies have been focused on U.S. equity market, while there are less focused on Euro-area stocks. Last is that collectively, the empirical findings conclude strongly in the presence of significant commonality in liquidity inside financial markets and highlight the time-varying nature of liquidity.

More recently, Marsh and Mazza (2017)⁷ researched market quality commonality using a comprehensive and more general definition of the concept, using a variety of proxies to capture the dimensions detailed in section 2.2). The results of their study also suggest that there exists a large commonality in market quality movements both for sector aggregates and country indexes in their datasets, which is a consistent extension of the findings of previous research.

In order to assess commonality, two main methodologies were used in academic research. The first is the ‘market model’ regressions used by Chordia, Roll, and Subrahmanyam (2000) where liquidity measures of an asset are used in regressions against market measures of liquidity and control variables. The second is the

⁷The research is in fact a preliminary version not yet published. We contacted one of the authors of the article (Paolo Mazza) who gave us his accord to quote the preliminary version of his work

one of Hasbrouck and Seppi (2001), which consists in using principal component and canonical correlation analyses to determine the existence of common factors.

Sources and pricing of commonality in market quality

Following the research treating the existence of a common factor in market quality movements across stocks, several authors have tried to investigate the sources of this factor (Brockman, Chung, and Pérignon, 2009 ; Fernando, 2003 ; Hameed, Kang, and Viswanathan, 2010 ; Huberman and Halka, 2001 ; Kamara, Lou, and Sadka, 2008 ; Karolyi, Lee, and Van Dijk, 2012). In the related academic literature, researchers have proposed various hypotheses to explain liquidity co-movements, like heterogeneity in the liquidity demand (Fernando, 2003), volatility of the market (Hameed, Kang, and Viswanathan, 2010), macroeconomic announcements (Brockman, Chung, and Pérignon, 2009), country-level factors related to either supply or demand of stocks (Karolyi, Lee, and Van Dijk, 2012), ownership of institutional stocks (Kamara, Lou, and Sadka, 2008) or trading effects (Huberman and Halka, 2001) to name a few. Taking a step back, the literature as a whole has used mainly two established price variation frameworks of financial markets to examine and analyse liquidity commonality : the asymmetric information framework and the inventory control framework (Anagnostidis and Fontaine, 2018). In their paper, Klein and Song (2017) classify all the hypotheses regarding the sources of commonality sources in two groups, by saying that the results from the body of literature suggests that liquidity commonality can emerge through either demand channels (Koch, Ruenzi, and Starks, 2016 ; Kamara, Lou, and Sadka, 2008) or supply channels (Coughenour and Saad, 2004). To illustrate this classification, we take the examples proposed by Anagnostidis and Fontaine (2018) for the supply channels (*"For example, in quote driven systems, liquidity commonality may emerge from common inventory adjustments by market makers who handle common pools of shares"* (Anagnostidis and Fontaine, 2018, p.2)) and for demand channels (*"Liquidity commonality may also emerge in order-driven systems due to common security price variations, within the market level, as well as due to common market trends that, in turn, lead to common portfolio adjustments by investors."* (Anagnostidis and Fontaine, 2018, p.2))

Moshirian et al. (2017) make a more detailed analysis of all the market-level determinants of liquidity commonality. They divide market-level determinants into three categories: economic and financial factors, environment information and investor protection factors, cultural and behavioural factors. They hypothesize that liquidity commonality is not only determined by some specific market-level determinants but also by some specific firm-specific characteristics. The results of their research show that both market and firm factors have significant effect on liquidity commonality. More specifically, they find that liquidity co-movements are higher in environments with "weaker and more volatile economic and financial conditions, minimal investor protection, and information opacity" (Moshirian et al., 2017, p. 36) and that at the firm-level, there is a significant association between liquidity commonality and firm transparency, correlated trading and stock performance (Moshirian et al., 2017). As such, they found that a firm with greater transparency has lower liquidity commonality (suggesting an increase of commonality with information asymmetry), a positive relationship between liquidity commonality and analyst forecast error and a negative one between liquidity commonality and analyst coverage. Some studies have also tried to analyse investor sentiment as a factor

driving liquidity commonality, but results are mixed (Lee, Tseng, and Yang, 2014).

Another area of focus of the literature is the relative contribution of industry and country returns commonalities in Europe due to the behavioural change of fund manager after the start of the euro currency (Marsh and Mazza, 2017). The same authors point out that the liquidity commonality literature generally focus on a one-country level in their research and lacks international analysis present in the cousin literature. This particular aspect was investigated by Marsh and Mazza (2017) for market quality and they concluded in the presence of large commonality in market quality movements for both sector and industry aggregates.

As explained in Chapter 1, the topic of our thesis has a significant importance for the asset pricing theory (related to cryptocurrencies in our case). In the common asset pricing theory, the expected return of a stock depends on the variation of other states variable (Beaupain, Giot, and Petitjean, 2010). The liquidity of a market, portraying the co-movements between the liquidity of the market's stocks, has been proven to be one of those state variables included in the price of a stock (Acharya and Pedersen, 2005 ; Amihud and Mendelson, 1986 ; Amihud, 2002 ; Gibson and Mougeot (2004) ; Jacoby, Fowler, and Gottesman (2000) ; Lee, Tseng, and Yang, 2014 ; Moshirian et al., 2017 ; (Pástor and Stambaugh, 2003)). However, academic research in this area is still in its infancy according to Moshirian et al. (2017), as we do not have a good understanding of the role of liquidity commonality in the pricing of assets.

2.2.2 Impact of high frequency trading

More recently with the rise of high-frequency trading, some articles have examined the impact that high frequency trading can have on liquidity and liquidity commonality. Most of the research are focused on the impact of high-frequency trading on the liquidity of individual liquidity (we won't review them), but some have analysed the role of high-frequency traders on co-movements in liquidity (Anagnostidis and Fontaine, 2018). This topic is relatively less relevant for cryptocurrencies as high frequency trading is not widespread and made difficult by the long time necessary to validate cryptocurrency transactions. However, an article from the Financial Times(Nicole, 2017) mentioned that the frequent price swings of cryptocurrencies are attracting more and more high-volume traders and that several large high-frequency trading companies have already started trading cryptocurrencies. Furthermore, some exchanges are beginning to offer client servers so we expect some forms of high-frequency trading to become common place in the cryptocurrency market.

High-frequency trading (or HFT), is a method of trading where automated computers act as traders through lightning-fast speed interactions with electronic trading platform (Benos et al., 2017). They have known a rapid growth recently and are now a common feature of modern markets. High frequency traders rely on a less varied set of trading strategies (Brogaard, 2010), like using arbitrage strategies to exploit a momentary mispricing between products or directional trading strategies based on new information (Breckenfelder, 2019). In their research paper, Klein and Song (2017) make a review of the advantages of high frequency traders exposed so far in academic literature. The main idea is that they improve the overall market

quality by treating new information more efficiently and by reinforcing the competition:

“They increase market liquidity (Brogaard 2010, Jovanovic and Menkveld 2016), reduce short-term volatility, at least during normal market conditions (Hassbrouck and Saar 2013, Hagströmer and Nordén 2013), and contribute positively to price discovery (Brogaard, Hendershott, and Riordan 2014). They reduce the trading costs of retail traders (Malinova, Park, and Riordan 2016), keep fragmented markets virtually consolidated (Menkveld 2013) and might even increase social welfare (Jovanovic and Menkveld 2016).” (Klein and Song, 2017, p.0)

In their research, O’Hara and Ye (2011) show that through market fragmentation, new trading platforms reduce effective spread of transactions costs, transaction speed, improve competition and overall efficiency. They also find that the magnitude of the impact varies according to listing and size regimes, fragmentation being particularly beneficial for small stocks. Despite those advantages, they also involve some market-level risks of increased liquidity spill overs:

“In addition to an increase in adverse selection costs for other traders (Biais, Foucault, and Moinas 2015; Foucault, Kozhan, and Tham 2017), and the likely contribution of HFTs to high volatility during the Flash Crash (Kirilenko et al. 2017; Easley, López de Prado, and O’Hara 2011), the July 2011 International Organization of Securities Commissions (IOSCO) Technical Committee report emphasizes the effect HFT activity might potentially have on the transmission of extreme shocks across different markets and asset classes.¹ In their theoretical paper, Cespa and Foucault (2014) show that cross-asset learning leads to liquidity spill overs across asset classes, and a small drop in liquidity of one asset can even cause a market wide liquidity crash.” (Klein and Song, 2017, 2018, p.0)

What about the impact on common movements of market quality then? Intuitively, as high frequency traders use the same type of trading algorithms and their trading strategies present common characteristics (Brogaard, 2010), we expect them to favour commonality in market quality through common co-movements in the demand and supply of stocks. In this respect, Klein and Song (2017) have shown through the staggered entrance of an alternative platform to trade European equities using high frequency trading (Chi-X) that liquidity co-movements in the aggregate national home market significantly increased after the introduction of the platform in the country. Their findings are aligned with the idea multimarket high-frequency traders provoke stronger network-wide co-movements in liquidity, subsequently making liquidity shocks easier to propagate through markets. In their research, Anagnostidis and Fontaine (2018) obtained results that lead to similar conclusions. They find that investors would have to pay higher transaction costs in the absence of high-frequency trading and that high-frequency trading implies strong liquidity commonality that increases the risk of having more non-diversifiable costs related to liquidity for traders (similarly to the results of Jain, Jain, and McInish (2016) conducted on the Japanese market).

2.2.3 Volatility regimes and liquidity distributions

In their study, Beaupain, Giot, and Petitjean (2010) study the impact of volatility and market capitalization on liquidity co-movements. Using a restricted selection of stocks listed on the New York Stock Exchange (NYSE) to class A shares, they show that the magnitude of common liquidity movements has a positive relationship with market capitalization on average (in other words, common liquidity movements are more intense among stocks with high capitalization and less intense for among stocks with small capitalization). However, the degree of pervasiveness is the same whatever the market capitalization.

Furthermore, Beaupain, Giot, and Petitjean (2010) also condition the previous analysis of the common movements in liquidity between portfolios of different capitalization upon volatility by defining three volatility regimes (low volatility, neutral volatility and high volatility). They show that liquidity co-movements for large caps are greater in quiet markets (common variations of adverse selection costs and inventory dominates the specific variations of a stock) and less important in stressful markets (stock-specific variations costs dominate). The same trend is observed for medium capitalization portfolios, but with less sensitivity to volatility. Finally, for portfolios of stocks with small capitalization, common liquidity movements are weak for both volatility regimes.

2.3 Market quality of cryptocurrencies

Even though the current academic literature on market quality of cryptocurrencies has barely been researched, two working papers came out during the time we spent working on our master's thesis on the topic. The first is a study of Bitcoin Liquidity (Marshall, Nguyen, and Visaltanachoti, 2019) and the second a study of the common factor structure in the returns and realized volatility of cryptocurrencies (Liu et al., 2019).

In their research, Liu et al. (2019) highlights the fact that the knowledge about the characteristics of cryptocurrencies as financial assets are still lacking in academic literature. They use high frequency trades data and quotes of the nine most liquid cryptocurrencies against Bitcoin to document the factor structure of volatility and returns, test whether major macroeconomic factors drive them and the impact of the Bitcoin Bubble. Among the stylized facts that the study shows, we note that they document stronger factor structure for the volatility compared to returns and that the factor structure model had a stronger explanatory power during the bubble.

Marshall, Nguyen, and Visaltanachoti (2019) also point out the lack of literature treating the subject of cryptocurrency's liquidity, in sharp contrast with the profusion of research on the liquidity of traditional financial assets. In spite of the fragmented nature of the Bitcoin market, they find strong evidences of a systematic component in the liquidity of Bitcoin. They also show that Bitcoin is not as strongly influenced as traditional assets by factors influencing stocks liquidity and that liquidity plays a key role in the price formation for Bitcoin.

Chapter 3

Methodology

The objective of the present Chapter is to explain the underlying approach to the empirical chapter. It also aims at providing a justification of the methods we chose in our attempt to address the problematic of our research. Concretely, this Chapter is divided into two sections. In light of our literature review, we refine our problematic into several research questions and empirical hypotheses in section 3.1. In section 3.2, we describe the methods we followed to test our hypotheses and answer our research questions.

3.1 Research questions and hypotheses

As developed in Chapter 1, the subject of our work is the co-movements in the market quality of cryptocurrencies. More specifically, our problematic consists in verifying whether the main findings of literature about market quality co-movements are transferable to recent cryptocurrency data. As stated in Chapter 1, the specific formulation of our problematic is :

“What findings of the academic literature about the market quality co-movements of traditional assets can be transferred to cryptocurrencies in 2017/2018?”

In order to answer our problematic, we already reviewed the main findings and discoveries of the academic literature about cryptocurrencies and market quality in Chapter 2. Considering our review of literature and given the data available for us to conduct our research, the themes treated in academic literature that are relevant and possible to investigate with our data are the presence of a common factor in market quality of cryptocurrencies, its evolution across several periods and the influence of volatility regimes on liquidity of cryptocurrencies.

3.1.1 Commonality in market quality

As explained in subsection 2.2.1, the presence of a common factor in market quality or dimensions of market quality such as liquidity is a well explored subject in academic literature on traditional financial assets but is barely researched for

cryptocurrencies (only the liquidity of Bitcoin and the common factor structure of cryptocurrencies's volatility). Therefore, we investigate the presence of a common factor in market quality movements of cryptocurrencies. Given the specific properties of cryptocurrencies as financial assets detailed in subsection 2.1.4 and the fact that cryptocurrencies markets are still immature, we expect to find the presence a common factor in market quality movement of our set of cryptocurrencies.

Research question 1 : There is a common factor in the market quality movements of our set of cryptocurrencies within our period of interest.

Furthermore, given the distinctive features of cryptocurrencies, we expect differences in the intensity of the common movements of market quality of cryptocurrencies compared to other traditional assets. We would a priori assume the commonality factor to be more important for cryptocurrencies given their properties compared to traditional assets (more illiquid, more volatile, less price efficient, etc.). However there are few academic researches with which we could compare our results (Marsh and Mazza (2017) work on a very different period from ours so the comparison would not be relevant) and reproducing our analysis on traditional assets to produce a comparison could be the topic of another master's thesis in itself. Consequentially, it is difficult to translate this intuition into a hypothesis that can be verified experimentally. However as Liu et al. (2019) and Christoffersen, Lunde, and Olesen (2019) also apply Principal Component Analysis to assess commonality in returns and realized volatility of cryptocurrencies and commodities respectively, we can still compare a part of our results.

Research question 2 : If a common factor exists in the Realized Volatility and Returns movements of our set of cryptocurrencies, the first component of the PCA for those proxies on our set of cryptocurrencies are more important than those documented by Christoffersen, Lunde, and Olesen (2019) for commodities.

3.1.2 Improvement of market efficiency, introduction of futures on Bitcoin and Bitcoin bubble

Among the most discussed topics of the cryptocurrency literature, three subjects attracted our attention in particular. First is that the cryptocurrency market has been shown to improve in efficiency over time as discussed in subsection 2.1.5. It could be interesting to see if this trend can also be observed in our datasets in terms of market quality. A comparatively more efficient market would translate into lower co-movements of market quality, so we would naturally expect the commonality factor to decrease from the beginning to the end of the studied period. We thus make two snapshots of liquidity at the beginning and the end of our working period to compare the results using PCA.

Research question 3 : The principal components of the PCA for every proxy of market quality are higher for our set of cryptocurrencies in the three first months of our period of interest compared to the last three months of the period.

The second and third topics of academic literature we chose to explore is the introduction of futures on Bitcoin and the bubble of Bitcoin. For the first event, some articles highlighted the potential of futures product to disrupt markets. As some evidences in the literature like Kumar, Sarin, and Shastri (1995) have shown that option listing improves market quality, we expect the commonality factor to decrease after the introduction of Bitcoin. For the second event and as discussed in subsection 2.1.3, the Bitcoin was in a bubble for a part of our working period. Liu et al. (2019) dates the end of the Bubble to the 28th of January 2018. They also show that their model was more powerful in explaining variations in returns and volatilities during the period of the bubble, so we expect commonality in market quality to be higher in the period preceding the end of the bubble compared to the period following it. Like for the improvement of market efficiency, we make two snapshots of market quality commonality before and after the introduction of futures products and the end of Bitcoin bubble to compare the results using PCA.

Research question 4 : The principal components of the PCA for every proxy of market quality are higher for our set of cryptocurrencies during the three months preceding the introduction of Bitcoin futures compared to the three months following the introduction of Bitcoin futures.

Research question 5 : The principal components of the PCA for every proxy of market quality are higher for our set of cryptocurrencies during the month preceding the end of Bitcoin bubble compared to the month following the end of Bitcoin bubble.

3.1.3 Influence of volatility regimes

In subsection 2.2.3, we discussed the impact of volatility regimes on liquidity commonality of traditional financial assets for small, middle and large caps. The impact of the volatility regimes on liquidity commonality is exploitable experimentally with our datasets but the analysis conditioned on market capitalization would not be relevant as we only have data for five cryptocurrencies from one exchange. Intuitively, we would expect the behaviour of our group of cryptocurrencies to be similar to the high market capitalization portfolio described in the article of Beaulpain, Giot, and Petitjean (2010), meaning a higher liquidity commonality in quiet markets given the high capitalization of cryptocurrencies described in Appendix C.

Hypothesis 1 : The commonality in liquidity in our set of cryptocurrencies is higher in low volatility regimes compared to high volatility regimes during the period of interest.

3.1.4 Evolution of liquidity commonality

In order to complement the comparison of PCA results across periods discussed in subsection 3.1.2, we test for the significativity of the evolution of commonality across the same periods, formulating the three following hypotheses (the economic intuitions behind the hypotheses is explained in research questions 3 to 5):

Hypothesis 2 : Commonality in liquidity is significantly higher for our set of cryptocurrencies in the three first months of our period of interest compared to the last three months of the period.

Hypothesis 3 : Commonality in liquidity is significantly higher for our set of cryptocurrencies during the three months preceding the introduction of Bitcoin futures compared to the three months following the introduction of Bitcoin futures.

Hypothesis 4 : Commonality in liquidity is significantly higher for our set of cryptocurrencies during the month preceding the end of Bitcoin bubble compared to the month following the end of Bitcoin bubble.

3.2 Description of our methodology

3.2.1 Assessing commonality in market quality

In order to assess commonality in market quality movements, we reproduce the methodology of Marsh and Mazza (2017), which is in fact a combination of approaches used in academic literature. The originality of their approach lies in their choice of proxies to capture market quality. They proceed in four steps :

1. *Merge the trade and order book datasets*

Marsh and Mazza (2017) work with intraday data separated into trade datasets and order book datasets. The very first step of their methodology consists in merging the trades and order books for each cryptocurrency by matching a trade with the quote prevailing at the time of the said trade. Starting by merging the two datasets allow them to compute all their proxies more easily as some are built from variables of both trades and quotes. However, it would not be consistent for us to use their merging rule as there is a discrepancy in our datasets between the frequency of observations for quotes and trades. Indeed, in the trade datasets we have generally an observation when they occur whereas we have only one observation per minute at the beginning of the order book datasets of each cryptocurrency. In our case, matching a trade with the prevailing quote at the time of the trade would result with the same quote being assigned to a lot of trades, which would introduce a bias in some of our proxies. A workaround consists in merging the two datasets based on time aggregates. We separately aggregate the data of the two datasets on 30 minutes intervals and then merge them using the time as key. As the authors later aggregate the data in 30 minutes intervals, our workaround has almost no consequence except for the Percentage Price Impact proxy that can no longer be computed with the same formula. When constructing their time aggregates, Marsh and Mazza (2017) use the mean as aggregation rule. We will do the same as robustness check but our main methodology will use the median instead, as explained in section 4.3.

2. Construct the proxies for the dimensions of market quality and aggregate them on 30 minutes intervals

As discussed in section 2.2, Marsh and Mazza (2017) summarize the concept of market quality through five dimensions. In their empirical part, they compute several proxies for each dimension. Their analysis in their research is focused around six proxies, the rest of the proxies being meant as robustness checks. In our framework, we decided to use only their six main proxies, which are Realized Spread for Cost of Immediacy, Percentage Price Impact for Adverse Selection, a modified version of Amihud (2002) for Liquidity, the absolute value of the returns autocorrelation of the midpoint prices for Price Efficiency, Realized Volatility and Quote Return Volatility of the midpoint prices for Volatility. We compute all proxies by following the definitions and formulas used by Marsh and Mazza (2017) except for Percentage Price Impact, where we replace the midpoint price by the trade price because we were unable to merge the datasets with the same rules as the authors due to the discrepancy between the frequency of observations of the trade and order book datasets¹. Again, we use the median instead of the mean as aggregation rule (see section 4.3). All the details of the definitions and formulas for each proxy can be found in Appendix A.

3. Standardization of the proxies

We reproduce the standardization procedure applied by Marsh and Mazza (2017), in accordance with Hasbrouck and Seppi (2001) and Corwin and Lipson (2011). This standardization procedure aims at removing intraday patterns by using the interval-specific mean and standard deviation of each cryptocurrency. The standardized variables are computed as follow :

$$z_{ipd} = \frac{x_{ipd} - \mu_{ip}}{\sigma_{ip}} \quad (3.1)$$

Where z_{ipd} is the standardized variable observed for asset i on interval p of day d , x_{ipd} is the variable before standardization for asset i on interval p of day d , μ_{ip} and σ_{ip} are the mean and standard deviation of asset i for the time interval p . As cryptocurrencies are traded all the time, we have 48 intervals per day (and thus 48 interval-specific means and standard deviations for each cryptocurrency).

4. Apply Principal Component Analysis (PCA)

Like suggested by Hasbrouck and Seppi (2001) and used by Marsh and Mazza (2017), we use Principal Component Analysis (PCA) to measure the co-movements of each proxy of market quality across our five cryptocurrencies. PCA is a descriptive method that uses the variance-covariance matrix of the input variables to create new variables called principal factors that should represent the input dataset in a better way (Marsh and Mazza, 2017). If we note $x_{i,t}$ the market quality proxy for cryptocurrency i at time t and our dataset contains

¹The formula of Percentage Price Impact used by Marsh and Mazza (2017) use the midpoint price and the direction of a trade at time t , which comes respectively from the order book datasets and trade datasets

N cryptocurrencies and T periods, then our dataset can be represented by a matrix $X = [X_1, X_2, \dots, X_N]$ where each X_i is a vector of length T . In our case, there are as much principal factors (noted F_j) as original variables N and they are constructed by linear combination of the original variables :

$$F_j = u_{1j}X_1 + u_{2j}X_2 + \dots + u_{Nj}X_N \quad (3.2)$$

The first component is constructed in a way that maximizes the explanatory power by maximizing the variance along its axis (Marsh and Mazza, 2017). The second factor is also built in such way that it maximizes the variance along its axis and is orthogonal to the first factor. The total variance of the input and output datasets are the same however the first k principal components are able to explain more variance than k input variables because of the way they were built. In other words, we are able to capture a large proportion of the variance with a smaller number of components. In our application, the greater the variance captured by our principal components for a given proxy the greater the commonality across our five cryptocurrencies for that proxy. To obtain our principal factors, we extract the eigenvalues and eigenvectors of the variance-covariance matrix of the original dataset. The first and highest eigenvalue (λ_1) corresponds to the first principal factor's variance, the second highest (λ_2) corresponds to the variance of the second principal component, and so on. Finally, we assess the importance of a factor by the ratio of the proportion of total variance that the factor explains :

$$\frac{\lambda_j}{\sum_{j=1}^N \lambda_j} \quad (3.3)$$

In our case, the presence of commonality should translate into a first factor to which most stocks correlate positively.

We extend the methodology by adding two steps of our own in order to compare snapshots of market quality commonality at different periods and make some simple robustness checks.

5. Reiterate PCA over other periods

We reiterate step 4 on different time intervals in order to have snapshots of commonality at different times. More specifically, we conduct PCA over the first three months of the sample (from the 11th of August 2017 until the 11th of November 2017) , the last three months (from the 21th of April 2018 until the 21th of July 2018), the three months preceding the introduction of futures on Bitcoin (from the 10th of September 2017 until the 10th of December 2017), the three months following it (from the 10th of December 2017 until the 10th of March 2018), the month preceding the end of Bitcoin bubble (from the 28th of December 2017 until the 28th of January 2018) and the month following it (from the 28th of January 2018 until the 28th of February 2018). Our reference for the date of the introduction of futures on Bitcoin is the day where the CBOE launched trading in Bitcoin futures products, the 10th of December 2017 Ryznar (2018). The reference for the end of Bitcoin bubble is the 28th of January 2018, as dated by Liu et al. (2019).

6. Robustness checks

In order to test if our results are robust to small changes in methodology, we also reproduce all the previous steps by applying the mean instead of the median for the time aggregates and by checking the results without standardization procedure.

3.2.2 Assessing the impact of volatility regimes on liquidity commonality

In order to assess liquidity co-movements conditioned on volatility regimes, we replicate a part of the methodology of Beaupain, Giot, and Petitjean (2010). We compute six proxies of liquidity for each cryptocurrency and regress them against market-wide liquidity proxies and other variables. The regressions are conditioned upon three volatility regimes: high, medium and low. The whole procedure is divided into 9 steps:

1. Merge trade and order book datasets.

As we will use regressions in this part of the work to assess liquidity commonality conditionally to volatility regimes, we merge the trade and order book datasets at the beginning of the methodology despite the discrepancy in frequency of observations in the beginning of the period. We merge the trade and order book datasets of every cryptocurrency by matching a trade with the prevailing quote at the time of the trade. In order to take into account the time jumps caused by missing quotes, we drop all the trades for which we do not have a quote 60 seconds before the time of the trade.

2. Compute six aggregated liquidity proxies over 15-minute intervals.

The liquidity proxies are constructed by :

- (a) Computing four base liquidity proxies for every trade and computing quote duration.
- (b) Aggregating the proxies over 5 minutes intervals using the median as aggregation rule.
- (c) Compute the six final liquidity proxies over 15 minutes intervals using the 5 minutes proxies and different weighting methods.

In (a), we compute two ex-ante and two ex-post liquidity measures for every trade. The ex-ante measures are the Proportional Quoted Spread (PQS) and the Quoted Depth (DEPTH). They are defined as follows:

$$PQS_t = \frac{p_t^A - p_t^B}{p_t^M} \quad (3.4)$$

$$DEPTH_t = \frac{q_t^A + q_t^B}{2} \quad (3.5)$$

Where at time t , p_t^A and p_t^B are the best quoted ask and bid prices respectively and p_t^M is the best mid-quote, defined as the simple average of the p_t^A and p_t^B . Similarly, q_t^A and q_t^B are the quantities of a given cryptocurrency available for the best ask and bid respectively. The ex-post liquidity measures are the Proportional Effective Spread (PES) and the Proportional Traded Spread (PTS), defined as :

$$PES_t = 2D_t \frac{p_t^{trade} - p_t^M}{p_t^M} \quad (3.6)$$

$$PTS_t = \frac{PES_t}{2} \quad (3.7)$$

Where D_t is the direction of the trade at time t (+1 for buy orders and -1 for sell orders) and p_t^{trade} is the trade price. Next is the inclusion of a quote-duration feature in our datasets by counting the time quotes remained unchanged in the order books. We make two assumptions to compute this new feature. First, given that most of the snapshots of our order-books are given with a one-minute frequency in the beginning of the period (later on the frequency of observations increases), we consider that the time-step between two consecutive snapshots of the order-books should be at most 60 seconds. We do this because we have to deal with missing snapshots of our order-books (sometimes hours or even days are missing). Whenever there are missing observations like these, the time difference around these can be quite large. Thus, after spotting missing values, we replace this disproportionate time-steps with the mode of time steps between observations. Our second assumption is that whenever we have following valid observations were the quote stays the same, we update the duration of the unchanged quotes. This is done after computing all quote duration in the following way : for every block of x consecutive valid but unchanged observations indexed $\{t - x + 1, t - x + 2, \dots, t\}$

$$\forall i \in \{t - x + 1, t - x + 2, \dots, t - 1\} : \text{duration}_i = \text{duration}_t$$

In other words, we assign the same quote duration to all the observations in the same block (and the duration of the block is the sum of the duration of the unchanged quotes). An important remark with respect to the quote duration is that the precision of this feature is limited by our datasets as quotes are provided every minute at the beginning of the period. It is possible that some quotes changed between two observations and we missed it due to the low frequency of observations.

After having computed our base liquidity proxies and added a quote duration feature (a), we now aggregate the proxies on 5 minutes intervals (b). This intermediary step fulfils two objectives. First is the reduction of the amount of data to treat before moving on to the weighted proxies. The initial amount of data being very demanding in terms of computing power. Second is to reduce the weights of outliers by using the median as aggregation rule. We aggregate our data into 5 minutes intervals and only keep the intervals containing at least three observations ².

²We have 345 trading days $\times$ 288 five-minute intervals per day = 99360 intervals. By rejecting the intervals with less than 3 trades we manage to keep 97% of the intervals. The rejected intervals ranged from 1% to 5% of the total number of 5 minutes intervals for each cryptocurrency. Moreover, there is no

Lastly comes (c), where we compute our final proxies by aggregating the outputs of (b) over 15 minutes intervals using different weighting methods : equal weights, size weights and time weights. We use equal weights to compute the Equally-Weighted Proportional Quoted spread and the Equally-Weighted Depth hereafter referred to as EWPQS and EWDEPTH. We then use size-weighted averages for one ex ante liquidity proxy. Specifically, PQS_t is weighted by $DEPTH_t$ to compute the Size-Weighted Proportional Quoted Spread SWPQS. PES_t and PTS_t are weighted by the $amount_t$ to compute the Size-Weighted Proportional Effective Spread SWPES and Size-Weighted Proportional Traded Spread SWPTS. Note that, in the case of SWPTS, the process is slightly different. We compute the weighted averages of our proportional traded spreads PTS and weight them by their respective trades amount. We then add these two weighted averages to construct our SWPTS. The last aggregation method is time-weighting, where we weight each PQS_t by the number of seconds its corresponding quote remained unchanged to compute the Time-Weighted Proportional Quoted Spread TWPQS.

3. Assessing daily volatility regimes.

This step consists in adding a flag for each day to indicate if the volatility regime of the day was low, medium or high. To do that, we use the daily values of the VCRIX³ which is a volatility index on cryptocurrencies. Just as VIX is induced from option prices in the S&P500 index, VCRIX is calculated from option-prices on the CRIX. CRIX stands for crypto-currency index and it is made freely available and hosted by Humboldt University of Berlin. We classify the daily volatility into one of the three regimes using an arbitrary threshold rule. We take the 33rd percentile and the 67th percentile of the VCRIX as lower and upper bounds respectively (if the value of the volatility for a given day is in the 67th percentile over the period, we assign a high volatility regime flag, low volatility regime for the 33rd percentile and medium in-between).

4. Compute market-wide variables and individual volatility.

For each of the six liquidity proxies and for each cryptocurrency, we calculate the simple average of that specific liquidity proxy over all other cryptocurrencies and call these "market liquidity proxies". To simplify the notation, if $L_{j,t}$ denotes a liquidity proxy for cryptocurrency j at time-interval t , then the market liquidity proxy associated will be noted $L_{M,t}$. Similarly, we calculate the log-returns of our time series with the same notation convention : we note $R_{j,t}$ the log-return on cryptocurrency j at time t and $R_{M,t}$ the corresponding market log-return. In addition, we compute $V_{j,t}$ the concurrent change in the volatility of cryptocurrency j at time t , which is in fact proxied by its squared log-return $R_{j,t}^2$.

5. Standardization

clear pattern in the time distribution of the rejected intervals. Mostly, they do not account for extreme values in any of the variables.

³www.thecrix.de

Before each regression detailed in the next step, we pre-process all the independent variables by removing intraday seasonality. We follow the same standardization procedure explained in step 3 of subsection 3.2.1 but using 15 minutes intervals.

6. *Regressions on liquidity proxies, unconditioned upon volatility regimes.*

We regress each proxy for each cryptocurrency using the equation below. The goal is to assess liquidity co-movements unconditioned upon volatility regimes.

$$L_{j,t} = \alpha_j + \beta_{1,j}L_{M,t} + \beta_{2,j}L_{M,t-1} + \beta_{3,j}L_{M,t+1} + \gamma_j V_{j,t} + \sum_{i=-1}^{+1} \delta_{i,j} R_{M,t+i} + \epsilon_{j,t} \quad (3.8)$$

Where $L_{j,t}$ denotes a liquidity proxy for cryptocurrency j at time-interval t , $L_{M,t}$ is the associated market liquidity proxy at time t , $V_{j,t}$ is the concurrent change in the volatility of cryptocurrency j at time t and $R_{M,t}$ is the market log-return at time t .

7. *Regressions on liquidity proxies, conditioned upon volatility regimes.*

We regress each proxy for each cryptocurrency using the following equation :

$$L_{j,t} = \sum_{k=1}^3 D_{k,t} (\alpha_{j,k} + \beta_{j,k} L_{M,T} + \gamma_j L_{t-1} + \phi_{j,k} V_{j,t} + \sum_{i=-1}^{+1} \delta_{i,j,k} R_{M,t+i}) + \epsilon_{j,t} \quad (3.9)$$

where $D_{k,t}$ is a dummy variable that takes the value of 1 when the volatility of the day is of type k , we use $k = 1, k = 2, k = 3$ for low volatility, medium volatility and high volatility regimes respectively. This equation differs from the regression in equation 3.8 in the fact that it includes a dummy variable $D_{k,j}$ as an indicator of volatility regimes (1 for low volatility, 2 for medium volatility, 3 for high volatility). It excludes Lag and Lead market-liquidity proxies and includes a 1-lag version of the dependent variable as a regressor. Furthermore, for any market-liquidity proxy variable, we define $\beta_{j,3}$ as the slope for that given variable and use it as a base category. We then define $\beta_{j,2}$ and $\beta_{j,1}$ as increments with respect to the base category $\beta_{j,3}$.

8. *Hypotheses test*

For each regression irrespective of the volatility regime presented in equation 3.8, we make three statistical tests. The first test is a Wald test, testing $H_0 : \beta_{1,j} + \beta_{2,j} + \beta_{3,j} = 0$. This test brings information on the pervasiveness of the common movements by getting rid of rapid variations and testing whether smoothed-out liquidity co-movements exist. The second test is another Wald test, this time testing $H_0 : \gamma_j = \delta_{-1,j} = \delta_{0,j} = \delta_{1,j} = 0$. This test evaluates the influence of all the other variables not directly related to market liquidity from regression equation 3.8. The third test is simply a F -test, which is here to test whether the fit of the intercept-only model is equivalent to the fit of the whole model.

Now for all the regressions conditioned on volatility discussed in equation 3.9, we only make two unilateral F-tests. The t -statistic of coefficient $\beta_{j,3}$ has the same meaning as before, but now for $\beta_{j,2}$ and $\beta_{j,1}$, the t -statistics comes from tests that check if the difference between the slope for their category and the slope of the reference category is null or not. Since these statistics are used to test $H_0 : \beta_{j,1} = \beta_{j,3}$, we can use them to test $H_0 : \beta_{j,1} < \text{or } > \beta_{j,3}$ against its corresponding opposite one-sided alternative hypothesis.

9. Robustness checks

In order to test if our results are robust to small changes in the methodology, we repeat all the previous steps without the standardization procedure, by applying the mean instead of the median as aggregation rule and by working with 30 minutes intervals instead of 15 minutes.

3.2.3 Evolution of the commonality in liquidity across periods

In order to test hypotheses 2, 3 and 4 and see if there are significant evolution of commonality in liquidity across the three pairs of periods, we proceed in two steps.

1. Daily regressions on liquidity proxies

We use the same liquidity proxies detailed in subsection 3.2.2 and use equation 3.8 to make daily regressions for each day belonging to our working period.

2. Tests of means

Using the R^2 of the daily regressions computed in the previous test, unilateral tests of means with a t -stat to reject or accept the null hypothesis.

3. Robustness check

We repeat the same procedure but with weekly regressions instead of daily regressions in order to verify that our results are consistent.

In total, we compare the power of the market model across three pairs of periods by making three tests of means. If we note μ_1 the mean of the R^2 of the daily regressions during the first three months of our period of interest, μ_2 during the last three months, μ_3 during the three months preceding the introduction of futures on Bitcoin, μ_4 during the three months following it, μ_5 during the month preceding the end of Bitcoin bubble and μ_6 during the month following it, we conduct the following tests :

$$\begin{aligned} & \underline{\text{Test1 :}} \\ & H_0 : \mu_1 < \mu_2 \\ & H_1 : \mu_1 > \mu_2 \end{aligned} \tag{3.10}$$

Test2 :

$$H_0 : \mu_3 < \mu_4 \quad (3.11)$$

$$H_1 : \mu_3 > \mu_4$$

Test3 :

$$H_0 : \mu_5 < \mu_6 \quad (3.12)$$

$$H_1 : \mu_5 > \mu_6$$

Chapter 4

Data

In this Chapter, we provide an overview of the datasets that we used all along the practical part of our research and give some insights that we learned by cleaning and exploring them. To explore and clean our datasets, we used a data management tool called Anatella¹. Anatella is an ETL software (Extract Transform Load) used to extract and process data. Anatella is very fast and powerful to process data in comparison to alternatives like R or Python. All the Anatella workflows used in the framework of our thesis can be found in the "Main Scripts" folder of our Dropbox². Our graphs were realized with gnuplot, a command-line graphing utility.

This Chapter is divided into three sections. First, we begin with a description of our datasets, variables and number of observations in section 4.1. Next in section 4.2, we review the preparatory analysis of our datasets. Finally, we review the steps we went through in order to clean our datasets in section 4.3.

4.1 Description of the datasets

For the empirical part of our work, we use intraday data on the trades and order book³ of the five main cryptocurrencies in terms of market capitalization in January 2018 (namely Bitcoin, Ethereum, Bitcoin Cash, EOS and Ripple) paired with US dollar on the Bitfinex exchange⁴. The list of all the variables and their descriptions for the trade and order book datasets are listed in Appendices B.1 and B.2 respectively and Appendix C contains a short description, factual information and recent market capitalization evolution for our set of cryptocurrencies. The source of the datasets are the records of the Bitfinex exchange itself. The time interval of the observations varies for the different cryptocurrencies (the trading of all cryptocurrencies was not launched at the same time on the Bitfinex exchange). Subsequently, the number of observations that we have for each cryptocurrency also differs.

Tables 4.1 and 4.2 summarize the dates of the first, last and total amount of

¹<http://www.anatella.com/>

²https://www.dropbox.com/sh/boxaiiaql9pbtcw/AAD_hK0Qx-XJ01u7jwQt0yMsa?dl=0

³Order book are list of buy and sell orders of a security (here on the Bitfinex exchange)

⁴Bitfinex is a trading platform headquartered in Hong-Kong and dedicated to digital assets and cryptocurrencies like Bitcoin, EOS, Ripple, etc. (<https://www.bitfinex.com/>). Note that the timezone of the observations is thus in UTC +8.

observations for each cryptocurrency in the trade and order book datasets respectively as well as their respective tickers .

Cryptocurrency	Ticker	Datetime first obser- vation	Datetime last obser- vation	Number of observa- tions
Bitcoin	BTC	01/04/2013 00:07:49	04/09/2018 23:59:46	53 120 202
Bitcoin Cash	BCH	10/08/2017 16:15:27	04/09/2018 23:57:34	10 735 889
Eos	EOS	10/08/2017 16:49:18	04/09/2018 23:58:54	14 761 953
Ripple	XRP	10/08/2017 15:02:28	04/09/2018 23:58:35	14 339 079
Ethereum	ETH	09/03/2016 16:04:35	04/09/2018 23:59:10	25 839 515

TABLE 4.1: Observations in the trade datasets for each cryptocur-
rency

Cryptocurrency	Ticker	Datetime first obser- vation	Datetime last obser- vation	Number of observa- tions
Bitcoin	BTC	15/05/2015 00:00:25	21/07/2018 23:59:12	1 538 292
Bitcoin cash	BCH	10/08/2017 17:58:00	21/07/2018 23:59:09	615 415
Eos	EOS	10/08/2017 17:58:00	21/07/2018 23:59:11	636 850
Ripple	XRP	10/08/2017 17:58:00	21/07/2018 23:59:18	616 424
Ethereum	ETH	28/04/2016 20:59:52	21/07/2018 23:59:14	1 048 678

TABLE 4.2: Observations in the order book datasets for each cryp-
tocurrency

4.2 Preparatory analysis

As preparatory analysis of our datasets, we computed summary statistics for the key variables of the trades and order books and made graphical analyses.

The tables with the summary statistics can be found in Appendix E and the graphs in our Dropbox under the "Plots trades"⁵ and "Plots order book"⁶ folders.

First, let's begin with the analysis of the trade datasets. By comparing the graphs, we made for the evolution of the trade prices of the five cryptocurrencies, two groups distinguish themselves based on the scale of price variations : XRP and EOS on one side, BTC, ETH and BCH on the other side. There is indeed a big difference in the average trade price of those two groups during our period of interest (the mean trade price on the period were \$1.0422 and \$9.5924 for XRP and EOS whereas we have an average trade price of \$617.6807, \$1 354.4715 and \$9 482.3155 for ETH, BCH and BTC respectively). From the graphical analysis of the evolution of the trade prices, we can make three main observations. First, at the end of the period, the prices of every cryptocurrency we were analysing were higher compared to their prices at the beginning of the period. Second, we cannot notice a lot of fluctuations in the trade price of every cryptocurrency (this can also be observed in the summary statistics where we have high standard deviations for the trade prices compared to their mean). Lastly, we can highlight two periods where all the trade prices increased (in different intensity and with some small variations around the beginning and end of the period) : from December 2017 until mid-January 2018 and from mid-March 2018 until April 2018.

With respect to traded amounts and volumes, it is difficult to interpret the data at least visually. Naturally, as a consequence to the huge disparity in trade prices between the two groups discussed above, we also find large differences in the average amount of cryptocurrencies traded between the two groups (when the unit price is higher, the average traded amount is lower ; for example, the average quantity of a trade of BTC over our period was of 0.4360 while it was 1 602.5399 for XRP). A fairer metric to have an idea of the respective popularity of the five cryptocurrencies as financial assets would be the volumes of the trades. In this regard, BTC comes first with a mean trade volume of \$712.2091, followed by BCH (\$2 265.9407) and ETH (\$2 191.0476). For BTC, ETH and BCH, we cannot observe a clear trend in the trade volumes plots as they are highly volatile (in the case of BTC for example, the volume of a trade has a standard deviation of \$1 6856.1428 whereas the mean traded volume is \$3 712.2091). On the other end for XRP and EOS, we can clearly see two trends visually with two increases in traded volumes from end of December 2017 until mid-January 2018 and from mid-March 2018 until April 2018. We can also see that the traded amounts are highly volatile, but we cannot pick a clear visual trend.

Now let's move on to the analysis of the order book datasets. For the bid and ask prices graphs, we find the same trends discussed above for the trade prices as expected (two peaks at the mentioned periods with varying length and intensity between cryptocurrencies ; for example, the first peak in prices for EOS is less intense compared to other cryptocurrencies but the second is more intense than for the others). On the whole period, bid and ask prices also increased and are higher than at the beginning of the sample. We note no visual differences between bid and ask prices graphically and between the first, fifth and tenth limit (the differences are too small to observe at this scale ; see the difference between ask and bid statistics in Appendix E).

⁵<https://www.dropbox.com/sh/cwnbgcq5cc7udyx/AAACGud1A4cVKgXXoEJQtVB8a?dl=0>

⁶https://www.dropbox.com/sh/parzcy8imdugqhp/AABzNqfr_OdmaEqOn-kzJGDca?dl=0

Again, it is difficult to interpret graphically the data for the quantities, but we still notice the difference in scale between the two groups of cryptocurrencies discussed above. Very interesting on the other hand is that we find large drops in quantities and even probably missing data at the end of January 2018 for every cryptocurrency. This potential lack of data needs to be investigated and taken care of otherwise some issues could arise in the empirical chapter with regard to the continuity of data.

4.3 Data cleaning

From the investigation of summary tables 4.1 and 4.2, we notice that the number of observations for Bitcoin and Ethereum are more important in comparison to the other cryptocurrencies in both the order book and trade datasets. This was expected as those two assets are older and were traded before the others on Bitfinex, so it is normal that the timeframe of the observations for those two cryptocurrencies starts earlier. For the practical part of our research, we will need to have a common time interval for which the data for every cryptocurrency is available. We thus restrict our datasets so as to only consider a time interval for which we have observations for both order book and trade datasets for all five cryptocurrencies and exclude the rest of the data. This gives us a common working interval going from 11/08/2017 at 00:00:00 until 21/07/2018 at 23:59:59. The first datetime for which we have observations for every cryptocurrency in all datasets comes a little bit before our chosen start datetime (we could have taken the start of the common interval the 10/08/2017 at 17:58:00). However, for clarity and because the data at the very beginning of the common period is not complete for all assets, we prefer to start the working interval by a complete day, in other words to 11/08/2017 at 00:00:00.

Next, in order to make a proper cleaning of our datasets, we need to tackle the two following elements : missing values and outliers. First of all, we methodically checked for the presence of missing and impossible values inside our datasets (we already spotted potential gaps in the data at the end of the end of January 2018, so we paid particular attention to this period). By impossible values, we refer to values in the datasets that cannot be correct with absolute certainty (like a negative price for example) and by missing values we refer to significant time skips in the observations that could engender issues in the empirical part of our work (like having insufficient data to compute our proxies for every 15min and 30min intervals). We did not find impossible values in the trade datasets but well in the book order datasets. In the book order data of every cryptocurrency, we found observations with negative values in either bid prices, ask prices, bid quantities or ask quantities. We cannot identify the source of those negative values with certainty but the only possibility we see would be a problem of reporting on the exchange side. We removed those observations from our datasets. Tables 4.3 and 4.4 present the ticker and adapted summary of the first, last, total number of observations and average number of observations per day in the trade and order book datasets respectively for each cryptocurrency after restricting the datasets to the common working interval and removing the spotted negative values inside the order book datasets. Starting from this Chapter, we will often refer to our cryptocurrencies by their tickers.

With respect to missing values, we first decided to investigate graphically the daily frequency of the observations in our datasets inside our working interval. The graphs and scripts of the graphs can be found in our Dropbox under the "Frequency plots" folder⁷. For the trade datasets, the number of observations for each day (the number of trades per day referenced on the Bitfinex exchange) fluctuates greatly for each cryptocurrency over the period but the only real concern are the days with a low number of observations (for example, there are two days in EOS and XRP with less than 3 000 trades). After a double-check, we found out sufficient data to compute our proxies for each 30 minutes intervals in the trade datasets. The situation for the order book datasets is quite different however. First of all, we can observe a change in the frequency of quotes in the middle of April 2018. At the beginning of our period, Bitfinex provided a quote every minute and the frequency of observations is nearly doubled during the month of April 2018. Moreover, we discovered a concerning common trend in all the order book plots : for every cryptocurrency, there is a gap in the frequency of observations in a period ranging from mid-January 2018 until the end of the month (the precise days of beginning and ending of the period vary from one asset to another) where the number of observations fall to zero (for several days in a row in some cases). If it remains untreated, this lack of data will inevitably lead to problems in the empirical part to compute our proxies and apply the PCA or the market model. Screenshots of the most significant time skips within our datasets for each cryptocurrency can be found in Appendix D. In our opinion, the easiest option was to deal with those discontinuities in the datasets after the merging of the trade and order book datasets, after construction of the proxies (as explained in section 3.2) but naturally before applying the PCA or the market model. Our approach simply consists in restricting the merged datasets to intervals for which we have a value for every cryptocurrency and exclude the time intervals for which we could not compute every proxy for at least one asset. On the 16 560 intervals of 30 minutes inside our period of interest, 850 were missing (the list of the missing 30 minutes intervals can be found in a csv file in our Dropbox under the "Missing intervals" folder⁸). The intervals with missing data are mostly at the end of January 2018 and beginning of February 2018 as expected. In their paper, Liu et al. (2019) also face the same problem of missing order book data in their intervals. They suspect that the source of the problem might be the rise in the price of the cryptocurrencies in the concerned periods, leading to extremely active trading and a failure from the data providers to keep up with the order book developments. In the same research, Liu et al. (2019) use linear interpolation to estimate the values for small missing intervals and a modified approach of Hansen and Lunde (2005) to account for larger breaks. In our case, we think that simply removing the missing intervals without interpolating them is sufficient as we are investigating common movements between cryptocurrencies (the time jumps will be the same for every cryptocurrency with our approach).

Now with respect to outliers, we considered applying standards outliers treatment methods, such as winsorizing⁹ or trimming¹⁰ our datasets but after discussions, we decided to opt for the alternative of using the median instead of the mean in the construction of the 15 minutes and 30 minutes aggregates as discussed in section 3.2. By taking the median instead of the mean of a variable or proxy in the

⁷https://www.dropbox.com/sh/6kxt9nzkdq1sbtn/AAD52VtR60TQWy_3omdxZ2S6a?dl=0

⁸<https://www.dropbox.com/sh/ie6zdklhvo9652i/AABBs6nkD8chf19vF0A43Bnra?dl=0>

⁹Replacing extreme values in the dataset

¹⁰Removing extreme values in the dataset

construction of the time aggregates, we reduce the weights of the outliers inside our datasets. We favoured this method over winsorizing or trimming because it allows us to avoid deleting any data and hence possible relevant information or introducing bias.

Cryptocurrency	Ticker	Datetime first obser- vation	Datetime last obser- vation	Number of observa- tions	Average observa- tions per day
Bitcoin	BTC	11/08/2017 00:00:02	21/07/2018 23:59:56	35 913 013	104 096
Bitcoin Cash	BCH	11/08/2017 00:00:14	21/07/2018 23:59:51	9 976 566	28 918
Eos	EOS	11/08/2017 00:00:18	21/07/2018 23:59:49	12 749 502	36 955
Ripple	XRP	11/08/2017 00:00:20	21/07/2018 23:59:07	13 482 419	39 080
Ethereum	ETH	11/08/2017 00:00:03	21/07/2018 23:59:44	19 866 751	57 585

TABLE 4.3: Observations in the trade datasets inside the common interval for each cryptocurrency after cleaning

Cryptocurrency	Ticker	Datetime first obser- vation	Datetime last obser- vation	Number of observa- tions	Average observa- tions per day
Bitcoin	BTC	11/08/2017 00:00:00	21/07/2018 23:59:12	630 580	1 833
Bitcoin Cash	BCH	11/08/2017 00:00:00	21/07/2018 23:59:09	614 857	1 787
Eos	EOS	11/08/2017 00:00:00	21/07/2018 23:59:11	636 329	1 855
Ripple	XRP	11/08/2017 00:00:00	21/07/2018 23:59:18	615 906	1 796
Ethereum	ETH	11/08/2017 00:00:00	21/07/2018 23:59:14	582 566	1 703

TABLE 4.4: Observations in the order book datasets inside the common interval for each cryptocurrency after cleaning

Chapter 5

Empirical analysis

In the following Chapter, we present the results obtained by applying the methodologies detailed in Chapter 3 on the datasets overviewed in Chapter 4 and discuss them. All the results of our research questions related to commonality in market quality are analysed in section 5.1, the study of impacts of volatility regimes on liquidity commonality in section 5.2 and the results of the hypotheses tests regarding the evolution of the explanatory power of the model in section 5.3.

5.1 Results : commonality in market quality

Before reviewing the results of the PCA, we follow Marsh and Mazza (2017) and compute the mean, standard deviation, median, skew and kurtosis for each proxy of market quality. The descriptive statistics for the proxies can be found in Appendix F.

The Realized Spread is lower than what we expected and is quite stable for all the cryptocurrencies judging by the standard deviations. We interpret this as a relative stability of the cost of immediacy over the whole period. The skew is low for all cryptocurrencies (many small values) except for BCH. The Realized Spread of BCH has very high Kurtosis (very high peaks in the distribution) while EOS and XRP approach Gaussian-like peakness. We find slightly negative values for the Percentage Price Impact, which is probably abnormal. This might be resulting from our change in the proxy's formula that does not take bid-ask bounces into account. In any case, the summary statistics indicates low adverse selection for every cryptocurrency. The liquidity appears to be quite good over the whole period as a unitary volume change has almost no impact on price in average. We find no surprise in the results for Return Autocorrelation, which are just a bit more volatile compared to a traditional asset. Price efficiency thus appears relatively good. As expected, the Realized Volatility is very high for 30 minutes intervals with high skew and kurtosis.

Now that we have reviewed the summary statistics for our proxies, we will expand on the results of the PCA in the following subsections and discuss their implications on commonality in market quality. In table 5.1, we present the results of the PCA applied over the whole working period. We report the explanatory power of the first five components for each proxy. Tables 5.2, 5.3 and 5.4 present the same results but for different periods with a comparison between the beginning and end

of the sample, a comparison of the results before and after introduction of futures on Bitcoin, before and after the end of the Bitcoin bubble respectively.

Measure	Components				
	1	2	3	4	5
Realized Spread	59.368	16.863	9.307	8.303	6.158
Percentage Price Impact	28.467	19.846	18.773	17.522	15.392
Realized Volatility	58.999	13.846	11.202	9.080	6.873
New Amihud	38.106	18.243	16.526	14.894	12.232
Return Autocorrelation	30.917	18.434	18.062	17.425	15.163
Quote Return	64.476	12.136	10.436	8.191	4.761

TABLE 5.1: Explanatory power (in %) of the five first components of the PCA for the whole period and with standardized proxies on 30 minutes intervals

Period	Measure	Components				
		1	2	3	4	5
Beginning	Realized Spread	43.784	17.669	16.108	12.758	9.681
	Percentage Price Impact	31.704	30.706	17.687	12.442	7.461
	Realized Volatility	46.038	22.915	15.213	9.342	6.492
	New Amihud	31.481	25.523	18.682	14.353	9.961
	Return Autocorrelation	25.109	20.238	19.570	19.008	16.075
	Quote Return	48.839	18.938	15.534	11.937	4.752
End	Realized Spread	67.611	15.961	13.819	1.517	1.092
	Percentage Price Impact	50.642	39.774	4.635	2.675	2.274
	Realized Volatility	55.652	22.203	11.276	8.274	2.596
	New Amihud	60.805	26.109	11.935	0.790	0.361
	Return Autocorrelation	33.878	17.315	17.232	16.550	15.024
	Quote Return	72.518	11.027	7.404	5.015	4.036

TABLE 5.2: Explanatory power (in %) of the five first components of the PCA for the three first and three last months of the period and with standardized proxies on 30 minutes intervals

5.1.1 Analysis of the results for the whole period

As explained in Chapter 3, one of the drawbacks of PCA is that the principal factors have no clear economical interpretations. Academic literature proposes

Period	Measure	Components				
		1	2	3	4	5
Before futures	Realized Spread	38.862	25.646	13.353	12.353	9.786
	Percentage Price Impact	35.393	24.779	15.182	14.016	10.630
	Realized Volatility	55.522	19.605	13.080	7.348	4.444
	New Amihud	37.906	25.079	16.727	12.822	7.466
	Return Autocorrelation	25.954	19.932	19.418	18.777	15.919
	Quote Return	56.788	18.628	13.186	6.635	4.763
After futures	Realized Spread	62.214	17.922	10.810	5.683	3.370
	Percentage Price Impact	50.026	36.573	10.019	1.771	1.611
	Realized Volatility	70.902	10.959	7.654	6.476	4.010
	New Amihud	57.429	24.846	15.855	1.145	0.725
	Return Autocorrelation	32.348	18.894	17.330	16.831	14.597
	Quote Return	72.339	10.934	7.235	5.454	4.039

TABLE 5.3: Explanatory power (in %) of the five first components of the PCA for the three months before and after the introduction of futures on Bitcoin and with standardized proxies on 30 minutes intervals

to look at the factor loadings (correlation between principal factors and original variables) to interpret the results. We notice that all the five cryptocurrencies are highly correlated positively with the first principal factors for each of the proxies. Those results therefore suggest high level of commonality across our set of cryptocurrencies for the six dimensions of market quality measured and within our period of interest and confirm *Research question 1*. However, compared to the results obtained by Marsh and Mazza (2017), it is not as evident to tell which proxies co-moves the most significantly across our assets by referring to the first principal factor to assess the intensity of co-movements. In their case, co-movements of Quote Returns and Realized Volatility were clearly more significant but for us there are slight variations in the results between the different periods (see tables 5.1, 5.2, 5.3 and 5.4). Nonetheless, on the whole working period the proxies that appears to comove (in other words have a bigger correlation between the first principal component and the original variables) are Quote Return, Realized Spread and Realized volatility. Whatever the proxy or period considered, the proportion of variance explained decreases significantly starting with the second principal component as shown by Marsh and Mazza (2017).

A part of our results can still be compared to Liu et al. (2019) as they investigate common factor structure for returns and volatility. Even though their proxies and ours are slightly different and the period are not identical, the proxies measure the same concepts over very close periods so the comparison should hold. The proportion of variance explained by the first component of the PCA for our Quote Return proxy is higher than theirs (64,48% in our results against 39,81% in theirs)

Period	Measure	Components				
		1	2	3	4	5
Before BTC bubble	Realized Spread	52.863	21.515	13.702	8.481	3.439
	Percentage Price Impact	53.439	38.469	5.874	1.137	1.081
	Realized Volatility	70.422	12.606	9.837	4.623	2.512
	New Amihud	66.161	22.689	7.375	2.742	1.032
	Return Autocorrelation	29.980	20.152	18.719	16.736	14.413
	Quote Return	70.422	12.606	9.837	4.623	2.512
After BTC bubble	Realized Spread	63.666	17.838	13.860	2.720	1.916
	Percentage Price Impact	75.969	10.978	5.683	4.095	3.274
	Realized Volatility	74.620	9.198	7.339	5.979	2.864
	New Amihud	62.116	21.821	14.581	0.816	0.667
	Return Autocorrelation	38.222	17.191	16.065	14.800	13.722
	Quote Return	77.859	7.456	5.734	4.840	4.110

TABLE 5.4: Explanatory power (in %) of the five first components of the PCA for the month preceding and following the end of BTC bubble and with standardized proxies on 30 minutes intervals

and the cumulative explained ratio is also higher (95,22% against 69,63%). In a recent study, Christoffersen, Lunde, and Olesen (2019) found small evidences of the presence of a common factor structure in commodity futures returns and Realized Volatility. In their results for the returns, the first four principal components explain 65.3% of the cross-sectional variation, which is close to the results of Liu et al. (2019) (cumulative explained variation is higher in our results). Important to note however is the fact that their first components of the PCA for the returns capture only 31.4% of the explained variance for commodities. In the case of cryptocurrencies, our first principal component captures 64,48% and the one of Liu et al. (2019) 39,81%, which is respectively 33,08% and 8,41% higher than in the commodity universe. Now regarding Realized Volatility, we find similar results to the three researches for the proportion of variance explained by the first four principal components, but they differ in the intensity of the first principal component. The first principal component and proportion of variance explained by the first four principal components are 70.15% and 84.8% for Liu et al. (2019), 54% and 79% for Christoffersen, Lunde, and Olesen (2019) and 59% and 93% for us. Again, in the case of commonality in Realized Volatility, the commonality seems to be higher for the cryptocurrencies compared to commodities. All of this put together, we are now able to answer to *Research question 2* by saying that for Returns and Realized Volatility, the common factor structure seems to be more important for our set of cryptocurrencies and the one of Liu et al. (2019) compared to the commodity universe of Christoffersen, Lunde, and Olesen (2019) on their respective periods. Again, those results are to be balanced with the facts that the periods of the three researches differ slightly and the construction of the proxies are not identical, so a more thorough investigation is required.

5.1.2 Comparison between the beginning and the end of the period

With respect to the comparison of the results of the PCA between the three first and three last months of the period of interest, we expected to find a general diminution of the commonality in market quality, translated by first principal components smaller given the fact that the academic literature highlights the improvement of the market efficiency of cryptocurrencies over time. Results of table 5.2 show instead that the first principal components for every proxy have increased. We can thus answer to *Research question 3* negatively by saying that the first principal components of every proxy are higher in the last three months of our working period compared to the first principal components of the three first months.

5.1.3 Comparisons before and after the introduction of Bitcoin futures

One of our research questions treated the comparison between the commonality in market quality three months before and after the introduction of futures on Bitcoin. If we look at table 5.3, we can see that the first principal factors have increased over the period for every proxy, which suggests an increase in market quality commonality in the period following the introduction of Bitcoin futures, contradicting *Research question 4*. The explanatory power of the first component for the three months preceding the introduction of futures are almost all lower than the explanatory power of the principal component of the whole period (except for PPI) and are close to the level during the first three months. It is the opposite for the principal components for the three months following the introduction of futures (higher explanatory power compared to the whole period and closer to the levels during the last three months of the period).

5.1.4 Comparison before and after the end of the Bitcoin bubble

For this comparison, we expected to observe stronger commonality before the end of the BTC bubble compared to after the bubble, translated by decreasing first principal components. However, in table 5.4, we see that for all the proxies except Liquidity, the first principal components have increased, which contradicts *Research question 5*. Similarly to the comparison between the commonality between the periods preceding and following the futures introduction, the results suggest that the level of commonality during the month preceding the end of the bubble was lower to the commonality on the whole period and closer to the commonality during the first three months. The opposite holds for the commonality after the end of the bubble. Furthermore, the level of commonality during the three months preceding futures introduction was lower to either the month preceding or following the end of the BTC bubble. Comparing results after the introduction of futures with table 5.4 does not give us a clear pattern (PPI, Realized Volatility, Return Autocorrelation and Quote Return are closer to the explanatory power of the first principal component before the end of the BTC bubble whereas Realized Spread and Liquidity are closer to the levels after the end of the BTC bubble).

5.1.5 Results of robustness checks

The results of the PCA applied over the different periods with non-standardized proxies and by using the mean as aggregation rule can be found in Appendices [G](#) and [H](#) respectively.

Without standardizing the proxies, the commonality seems to be more important for Realized Spread, Percentage Price Impact and Liquidity while being slightly lower for the other proxies. However, the comparisons of the results by period seems to be less clear (for certain proxies, the trend is even reverting). The standardization of the proxies thus appears to have an important impact on market quality commonality, and this impact seems to vary depending on the proxy. With respect to the results obtained by using the mean instead of the median, we observe slightly higher commonality for some of the proxies (Realized Spread, Percentage Price Impact, Realized Volatility, Quote Return) whereas it stays almost identical for the others (which suggests that some proxies might be more subjects to outliers than others). The comparisons of the results by periods give the same trends in general. Our results obtained by using the median as the aggregation rules seems to be robust in this aspect.

5.2 Results : impacts of volatility regimes on liquidity commonality

In this section, our goal is to assess whether or not we observe liquidity co-movements among our five cryptocurrencies. We start with a general analysis, disregarding the effect of volatility regimes. Then, we proceed to include the level of volatility of the day into the analysis. As mentioned before in the methodology, we follow Beaupain et al. (2010) methodology. That is, we use the same *market model* time-series regressions à la Chordia et al. (2000), where liquidity proxies are regressed against their corresponding market-wide liquidity proxy. However, there is one major difference. We determine the daily volatility regime (low, medium or high) not through the Markov switching methodology but using a threshold rule on the VCRIX, an index of cryptocurrencies' volatility. For both regression equations used in this section, all variables and parameters are defined in section [3.2.2](#).

5.2.1 Summary statistics for the proxies

Table [5.5](#) gives some cross-sectional statistics of the global means for each of the six liquidity measures defined above. Like in academic literature, we also observe right-skewness in the distribution of our proxies. Indeed, all the medians for all proxies are lower than their respective means. Again, we observe that depth shows the greatest variance among proxies.

We make four main observations regarding the summary statistics of our proxies. First is that contrary to what could be expected from the literature regarding stocks, the proportional effective spread and the proportional traded spread appear to be larger than the proportional quoted spread in our datasets. This would suggest

Liq. Proxy	Mean	Median	Stdev	Min	Max	AC(1)
EWPQS	0.0008	0.0004	0.0011	-0.0748	0.0249	0.76
EWDEPTH	1 767.9540	28.8419	5 489.4700	0.0200	606 577.8500	0.50
SWPQS	0.0008	0.0004	0.0012	-0.0683	0.0416	0.70
SWPES	0.0012	0.0007	0.0017	-0.0874	0.1215	0.56
SWPTS	0.0017	0.0007	0.0015	-0.0731	0.0646	0.52
TWPQS	0.0008	0.0004	0.0011	-0.0748	0.0249	0.76

TABLE 5.5: Aggregate liquidity measures - Summary Statistics

some form of price deterioration. Second is that we observe for all our proxies except EWDEPTH that the series contain negative spreads. Third is that overall, SWPTS are tighter than the corresponding SWPES and proportional quoted spreads are the highest when weighted equally. This means that higher spreads are linked to lower quoted sizes and that higher spreads are more ephemeral. However, this does not hold when conditioning the regressions on volatility regime. Finally, we can see that spread measures display a one-lag autocorrelation ranging between 50% and 76%.

Figure 5.1 is a heatmap representing correlations between our proxies. We find that most of our liquidity proxies are positively correlated to each other's. The only special case is EWDEPTH. This proxy is almost uncorrelated to all the others. More interestingly, we observe that TWPQS and EWPQS are almost perfectly correlated. This is not surprising since the vast majority of our quotes have the same duration, which is linked to the fact that we have a low frequency of observations for the order-book datasets. In this situation, assigning time-weights is almost equivalent to assigning equal weights to our spreads. The same applies to SWPQS and EWPQS that are also highly correlated. As we can see from the individual distribution plots in Appendix I, most of our trades have the same depths and duration.

Summary statistics for cross-sectional data but conditioned on the volatility regimes can be found in tables 5.6, 5.7 and 5.8. Note at this stage, we removed intraday seasonality with the standardization of our proxies. Again, we still have right-skewness under every volatility regime. There are two main and opposite patterns here. On one hand the average value of EWDEPTH increases as daily volatility decreases. On the other hand, all other average spread measures rise in more volatile markets. As expected, the variance of all the liquidity proxies is the highest in the high-volatility regime. Besides, we see that autocorrelation tends to be lower in highly volatile days. On average, in low volatility regimes, all proxies but depth are negative. The opposite is true for high volatility regimes. In the medium volatility setup though, all proxies are negatives. When analysing the absolute value of size-weighted spread measures, we observe that the average quoted spread is systematically greater than the average traded spread which in turn is greater than the average effective spread. Absolute values of our quote-based spreads, we have that size-weighting yields the tightest average spreads whereas time-weighting and equal-weighting yield the largest average spreads in volatile and quiet markets respectively. In the quiet market setup, all proxies but EWDEPTH have negative medians and means and their standard deviations falls below 1. This is in line with the

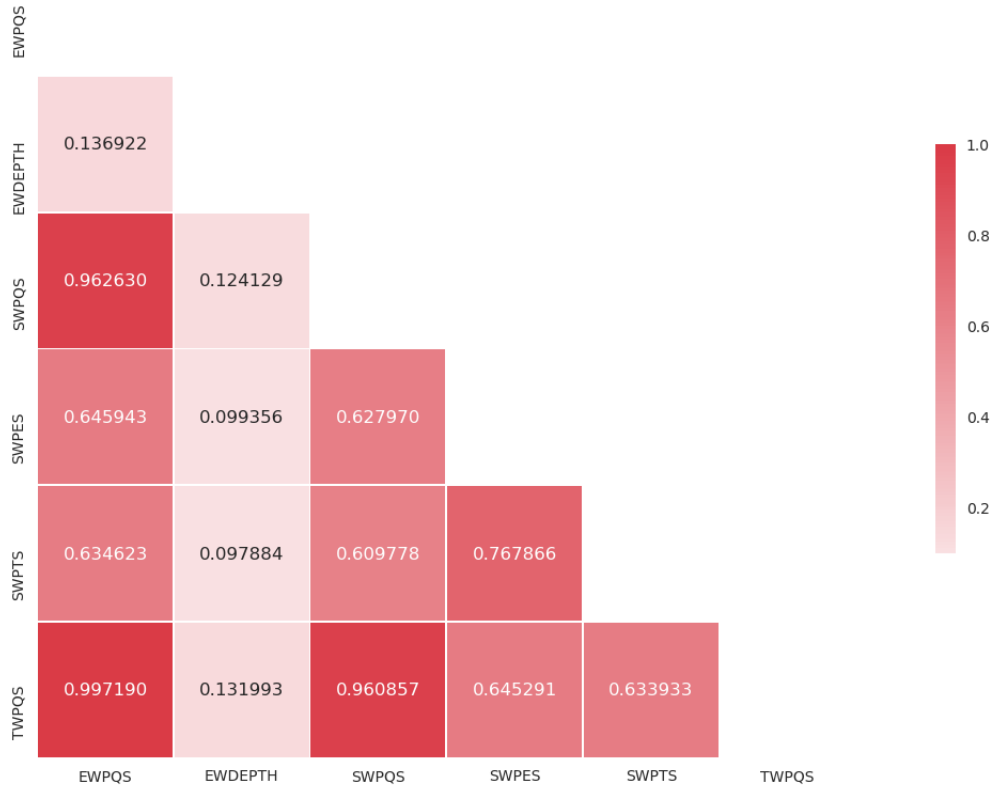


FIGURE 5.1: Average correlations between proxies

fact that the lower the volatility of the market the lower liquidity costs tend to be. On the opposite setup, we observe the exact reverse situation. Thus, we can say that spread-based and depth-based liquidity proxies depend on the volatility regime. The more volatile the market, the higher the spreads, which is consistent with Pástor and Stambaugh (2003). Reciprocally, the more volatile the market, the lower the depth. Those results show an inverse relationship between liquidity and volatility (as found in Beupain, Giot, and Petitjean (2010)) for our set of cryptocurrencies.

5.2.2 Market-model regressions disregarding volatility regimes

Tables 5.9 and 5.10 show the results of the regression presented in equation 3.8 for each liquidity proxy and for each cryptocurrency¹ and the results of the statistical tests. The results show evidence of common movements in liquidity. Indeed, at a 5% significance level all $\beta_{1,j}$ coefficients are statistically significant for all five cryptocurrencies. More than that, the coefficients $\beta_{1,j}$ up to $\beta_{3,j}$ representing the coefficients of the concurrent, lag and lead market liquidity proxies are all statistically greater from zero². Given that EWPQS, SWPQS and TWPQS are highly correlated, all their corresponding coefficients are very similar. These quote-based proxies display the strongest evidence of liquidity co-movements. In second position comes SWPES and SWPTS. Lastly, EWDEPTH shows the weakest evidence of co-movements in liquidity. This result is consistent with Chordia, Roll, and Subrahmanyam (2000). The best R^2 are obtained for the equally-weighted proportional

¹We do not show the coefficients for all the variables as the table are voluminous

²all t -stats are greater than the upper-tail critical value of the t -test, 1.645 at 5% confidence level

Liq. Proxy	Mean	Median	Stdev	AC(1)
EW PQS	-0.1423	-0.4335	0.8768	0.6565
EW DEPTH	0.1663	-0.1070	1.0709	0.3047
SW PQS	-0.1375	-0.4545	0.8660	0.5801
SW PES	-0.0670	-0.2923	0.9293	0.4543
SW PTS	-0.0739	-0.2616	0.9247	0.4273
TW PQS	-0.1415	-0.4370	0.8784	0.6544

TABLE 5.6: Aggregated standardized liquidity measures in low volatility regimes : Summary Statistics

Liq. Proxy	Mean	Median	Stdev	AC(1)
EW PQS	-0.1707	-0.4970	0.9036	0.6666
EW DEPTH	-0.0096	-0.1966	0.9245	0.2615
SW PQS	-0.1585	-0.4885	0.8980	0.5962
SW PES	-0.1434	-0.4045	0.9180	0.4881
SW PTS	-0.1458	-0.3600	0.9234	0.4108
TW PQS	-0.1723	-0.5007	0.9020	0.6628

TABLE 5.7: Aggregated standardized liquidity measures in medium volatility regimes : Summary Statistics

Liq. Proxy	Mean	Median	Stdev	AC(1)
EW PQS	0.4498	0.1343	1.1282	0.5708
EW DEPTH	-0.1444	-0.3360	1.0307	0.3007
SW PQS	0.4228	0.0680	1.1582	0.5016
SW PES	0.3266	0.0690	1.122147	0.4628
SW PTS	0.3377	0.1202	1.1136	0.4037
TW PQS	0.4517	0.1334	1.1282	0.5642

TABLE 5.8: Aggregated standardized liquidity measures in high volatility regimes : Summary Statistics

quoted spreads, ranging between 29% and 52% and are followed closely by the other two quote-based spread measures.

With respect to the results of the first Wald test, tables 5.9 and 5.10 show that the null hypothesis is rejected every time. This means that the common movements are not instantaneous, and we have to take the pervasiveness of the commonality into account. Out of 30 regressions, the null hypothesis is only rejected on three occasions for the second test (of which two rejections happened for the EWDEPTH). This shows that at our confidence level, we cannot reject the hypothesis that the external variables in the regressions have an influence on our liquidity proxies. For the last F-test, the null hypothesis is rejected every time, so we can conclude that there is a significant difference between our model and the intercept-only model at our confidence level.

5.2.3 Market-model regressions conditioned on volatility regimes

Tables K.2 and 5.12 show the results of the regression presented in equation 3.9, for each liquidity proxy and for each cryptocurrency. Like in Beaupain, Giot, and Petitjean (2010) we choose to show only the comparison between low and high volatility regimes because of the important volumes of data. For these regressions, it is important to note that we report our results in a different manner compared to what we did previously.

We observe that all our $\beta_{j,3}$ are statistically greater than 0 at a significance level of 5%, meaning there is evidence of positively correlated liquidity co-movements in high volatility setups for all cryptocurrencies. When analysing all the spread-based measures of liquidity, there is a statistical decrease in commonality in going from high volatility to low volatility in the case of Ripple. It constitutes the only case we can ever systematically reject the null in favour of the one-sided alternative that co-movements are higher in higher volatility regimes. For quote-based measures, when analysing the t -stats of $\beta_{j,1}$ s, we have evidence that, for all cryptocurrencies but Ripple (XRP) and Bitcoin Cash (BCH), there is a statistical increase in liquidity co-movements in going from high volatility to low volatility regimes. But here, in the case of BCH we cannot say that liquidity commonality varies with respect to these volatility regimes, as we never reject $H_0 : \beta_1 = \beta_3$. For the other three cryptocurrencies (BTC, EOS, ETH) we systematically reject the null in favour of the one-sided alternative that co-movements are higher in low volatility regimes. For effective or traded measures, only BCH's commonality is statistically higher in quiet markets. Finally, we notice that when conditioning upon the volatility regime, our R^2 improved significantly overall. For example, for the EWPQS proxy we have again the highest R^2 , ranging from 40% to 64%. Once again, EWDEPTH yields the worst fit for the regressions.

5.3 Results : evolution of liquidity commonality

We report the average values of the adjusted- R^2 of the daily and weekly regressions for each proxy by period in Appendix J and made plots of the evolution of the daily adjusted- R^2 in our Dropbox under the folder "Daily Adj. R^2 " ³. The difference between the average values of the daily and weekly adjusted- R^2 on the

³<https://www.dropbox.com/sh/y7ulbhnp2sr08hs/AADfvt4nZur-zGzhsSlGEsHKA?dl=0>

three periods and the value of the adjusted- R^2 for the regression on the whole period is striking. The adjusted- R^2 for all the three periods are much weaker compared to the values when the regressions is fitted on the entire period, even without conditioning upon volatility. We note that when moving from daily regressions to weekly regressions, the average R^2 seems to improve for every proxy but are still very low compared to the results we presented in tables 5.9 and 5.10. Furthermore, these first results seem to contradict our economic intuition for almost every cryptocurrencies and proxies of the Pre-Futures/Post-Futures comparison and Beginning/End of the sample. With respect to the graphs of the evolution of daily adjusted- R^2 , the graphical analysis does not reveal much as we can see that a large majority of the values oscillates at a very low level (between 0% and 10%). We reflected on the potential causes for this gap and our intuition is that the cryptocurrencies market is still too immature to be able to observe a consistent trend across time. When the frequency of trades for an asset are very high, those assets are more sensitive to idiosyncratic liquidity shocks. A small shock in the liquidity of an asset is enough to provoke a change in the price of the asset without observing common movements within the market when the frequency increases.

We report all the p-values of the tests of means discussed in subsection 3.2.3 in Appendix K. With respect to the comparison between the beginning and the end of the period and the introduction of Bitcoin futures, the results are not very conclusive. Indeed, for the daily regressions we cannot reject the null hypothesis for any of the tests. Furthermore, in the case of the comparison between the beginning and the end of the period, the average daily adjusted- R^2 and the fact that the p-values are very high for every proxy of every cryptocurrency except BTC are hints that the commonality is higher in the last three months of the sample, which contradicts our economic intuition. The situation is more or less the same for the introduction of futures but in this case the results are more mixed (the tests have small p-values for BTC and BCH and high p-values for EOS, ETH and XRP). When we move on to the weekly regression, we can reject the null hypothesis for four of BTC's proxies (SWPS, TWPQS, SWPES and SWPTS) at confidence level of 95% (RH_0), implying that for those proxies, the explanatory power of the model was better at during the three first months of the period in average compared to the three last months. We also notice that the p-values fluctuates a lot between the daily and weekly reports. Results are also difficult to interpret for the comparison between the month preceding the end of the BTC bubble and the month following it. For the daily regressions and at a confidence level of 95%, we can reject the null hypothesis (RH_0) for 6 proxies (EWPQS of BCH ; SWPES of EOS and XRP ; SWPTS of EOS ; ETH and XRP) and say that in average, the explanatory power of the model was better for those proxies during the bubble. With daily regressions, we can reject the null hypothesis for 16 proxies, 5 of them being already rejected at the daily frequency (EWPQS for EOS and XRP ; SWPQS for EOS ; TWPQS for EOS and XRP ; SWPES and SWPTS for every cryptocurrency ; EWDEPTH for BTC). For the rest of the proxies, we cannot reject the null hypothesis. We cannot even observe a clear pattern in the p-values so it is difficult to draw further conclusions on the (non)improvement of our models based on those results.

TABLE 5.9: Liquidity co-movements - Unconditioned upon volatility regime

EWPQS	BCH	BTC	EOS	ETH	XRP
$L_{M,t}$	0.2440	0.1788	0.2412	0.1761	0.2588
p -value	***	***	***	***	***
$L_{M,t-1}$	0.1960	0.1621	0.2416	0.1692	0.2570
p -value	***	***	***	***	***
$L_{M,t+1}$	0.2047	0.1563	0.2303	0.1609	0.2558
p -value	***	***	***	***	***
Adj. R²	0.3860	0.3212	0.4528	0.2945	0.5157
$H_0 : \beta_{1,j} + \beta_{2,j} + \beta_{3,j} = 0$					
p -value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
$H_0 : \gamma_j = \delta_{-1,j} = \delta_{0,j}$ $= \delta_{1,j} = 0$					
p -value	0.0054	<0.0001	0.0014	<0.0001	<0.0001
F -test					
p -value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
SWPQS	BCH	BTC	EOS	ETH	XRP
$L_{M,t}$	0.2166	0.1675	0.2256	0.1631	0.2425
p -value	***	***	***	***	***
$L_{M,t-1}$	0.1823	0.0336	0.2313	0.1475	0.2514
p -value	***	***	***	***	***
$L_{M,t+1}$	0.2006	0.1315	0.2235	0.1430	0.2436
p -value	***	***	***	***	***
Adj. R²	0.3255	0.2625	0.3953	0.2359	0.4489
$H_0 : \beta_{1,j} + \beta_{2,j} + \beta_{3,j} = 0$					
p -value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
$H_0 : \gamma_j = \delta_{-1,j} = \delta_{0,j}$ $= \delta_{1,j} = 0$					
p -value	0.0006	<0.0001	0.0186	<0.0001	<0.0001
F -test					
p -value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
TWPQS	BCH	BTC	EOS	ETH	XRP
$L_{M,t}$	0.2398	0.1816	0.2405	0.1738	0.2585
p -value	***	***	***	***	***
$L_{M,t-1}$	0.1981	0.1593	0.2360	0.1711	0.2538
p -value	***	***	***	***	***
$L_{M,t+1}$	0.2057	0.1530	0.2270	0.1608	0.2512
p -value	***	***	***	***	***
Adj. R²	0.3834	0.3195	0.4405	0.2942	0.5042
$H_0 : \beta_{1,j} + \beta_{2,j} + \beta_{3,j} = 0$					
p -value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
$H_0 : \gamma_j = \delta_{-1,j} = \delta_{0,j}$ $= \delta_{1,j} = 0$					
p -value	0.0033	<0.0001	0.0006	<0.0001	<0.0001
F -test					
p -value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

Newey and West's (1994) heteroskedasticity and autocorrelation robust standard errors are calculated using automatic lag truncation parameter selection. Significance levels: *** ≤ 0.001 , ** ≤ 0.01 , * ≤ 0.05 and ns > 0.05

TABLE 5.10: Liquidity co-movements - Unconditioned upon volatility regime

SWPES	BCH	BTC	EOS	ETH	XRP
$L_{M,t}$	0.1705	0.2038	0.1861	0.2165	0.2377
p -value	***	***	***	***	***
$L_{M,t-1}$	0.1460	0.0147	0.1705	0.1275	0.1754
p -value	***	ns	***	***	***
$L_{M,t+1}$	0.1169	0.0745	0.1463	0.1658	0.1905
p -value	***	***	***	***	***
Adj. R²	0.2841	0.2343	0.2614	0.2929	0.3923
$H_0 : \beta_{1,j} + \beta_{2,j} + \beta_{3,j} = 0$					
p -value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
$H_0 : \gamma_j = \delta_{-1,j} = \delta_{0,j}$ $= \delta_{1,j} = 0$					
p -value	0.0272	0.0001	<0.0001	0.1136	0.0001
F -test					
p -value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
SWPTS	BCH	BTC	EOS	ETH	XRP
$L_{M,t}$	0.1705	0.1664	0.2035	0.2055	0.2329
p -value	***	***	***	***	***
$L_{M,t-1}$	0.1424	0.0879	0.1519	0.1464	0.1827
p -value	***	***	***	***	***
$L_{M,t+1}$	0.1295	0.0866	0.1292	0.1581	0.1742
p -value	***	***	***	***	***
Adj. R²	0.2507	0.2117	0.2217	0.2556	0.3385
$H_0 : \beta_{1,j} + \beta_{2,j} + \beta_{3,j} = 0$					
p -value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
$H_0 : \gamma_j = \delta_{-1,j} = \delta_{0,j}$ $= \delta_{1,j} = 0$					
p -value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
F -test					
p -value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
EWDEPTH	BCH	BTC	EOS	ETH	XRP
$L_{M,t}$	0.0383	0.0158	0.0462	0.0143	0.0520
p -value	***	**	***	ns	***
$L_{M,t-1}$	0.0353	0.0036	0.0605	-0.0035	0.0722
p -value	***	ns	***	ns	***
$L_{M,t+1}$	0.0439	0.0101	0.0697	0.0014	0.0630
p -value	***	ns	***	ns	***
Adj. R²	0.0072	0.0074	0.0170	0.0043	0.0172
$H_0 : \beta_{1,j} + \beta_{2,j} + \beta_{3,j} = 0$					
p -value	<0.0001	0.0034	<0.0001	0.2718	<0.0001
$H_0 : \gamma_j = \delta_{-1,j} = \delta_{0,j}$ $= \delta_{1,j} = 0$					
p -value	0.5248	<0.0001	0.0008	<0.0001	0.3946
F -test					
p -value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

Newey and West's (1994) heteroskedasticity and autocorrelation robust standard errors are calculated using automatic lag truncation parameter selection. Significance levels: *** ≤ 0.001 , ** ≤ 0.01 , * ≤ 0.05 and ns > 0.05

TABLE 5.11: Liquidity co-movements - Conditioned upon volatility regime

EWPQS	BCH	BTC	EOS	ETH	XRP
$L_{M,1}$ "high volatility"	0.0090	0.0877	0.1274	0.2223	-0.0770
t -stat	(0.32)	(3.60)	(3.08)	(11.06)	(-1.96)
p -value	ns	***	**	***	*
$L_{M,3}$ "low volatility"	0.2654	0.1946	0.3171	0.0819	0.3053
t -stat	(11.14)	(11.50)	(9.70)	(8.215)	(8.09)
p -value	***	***	***	***	***
Adj. R^2	0.4908	0.4029	0.5300	0.5260	0.6392
reject H_0					
$H_0 : \beta_1 < \beta_3$ vs $H_1 : \beta_1 > \beta_3$	no	yes	yes	yes	no
$H_0 : \beta_1 > \beta_3$ vs $H_1 : \beta_1 < \beta_3$	no	no	no	no	yes
SWPQS	BCH	BTC	EOS	ETH	XRP
$L_{M,1}$ "high volatility"	0.3565	0.0993	0.1145	0.2311	-0.0864
t -stat	(1.28)	(4.06)	(3.00)	(10.67)	(-2.35)
p -value	ns	***	**	***	*
$L_{M,3}$ "low volatility"	0.2419	0.1574	0.3100	0.0723	0.3018
t -stat	(10.52)	(10.08)	(10.66)	(6.89)	(8.58)
p -value	***	***	***	***	***
Adj. R^2	0.4067	0.3383	0.4507	0.4279	0.5544
reject H_0					
$H_0 : \beta_1 < \beta_3$ vs $H_1 : \beta_1 > \beta_3$	no	yes	yes	yes	no
$H_0 : \beta_1 > \beta_3$ vs $H_1 : \beta_1 < \beta_3$	no	no	no	no	yes
TWPQS	BCH	BTC	EOS	ETH	XRP
$L_{M,1}$ "high volatility"	0.0203	0.0859	0.1304	0.2122	-0.0770
t -stat	(0.74)	(3.55)	(3.21)	(10.54)	(-1.96)
p -value	ns	***	**	***	*
$L_{M,3}$ "low volatility"	0.2579	0.1912	0.3123	0.0834	0.3082
t -stat	(10.99)	(11.38)	(9.76)	(8.27)	(8.22)
p -value	***	***	***	***	***
Adj. R^2	0.4865	0.4040	0.5217	0.5250	0.6270
reject H_0					
$H_0 : \beta_1 < \beta_3$ vs $H_1 : \beta_1 > \beta_3$	no	yes	yes	yes	no
$H_0 : \beta_1 > \beta_3$ vs $H_1 : \beta_1 < \beta_3$	no	no	no	no	yes

Newey and West's (1994) heteroskedasticity and autocorrelation robust standard errors are calculated using automatic lag truncation parameter selection. Significance levels: *** ≤ 0.001 , ** ≤ 0.01 , * ≤ 0.05 and ns > 0.05 .

TABLE 5.12: Liquidity co-movements - Conditioned upon volatility regime

SWPES	BCH	BTC	EOS	ETH	XRP
$L_{M,1}$ "high volatility"	0.0436	0.0138	0.0563	0.0006	-0.0568
t -stat	(1.77)	(0.52)	(1.58)	(10.55)	(-2.17)
p -value	ns	ns	ns	***	*
$L_{M,3}$ "low volatility"	0.1954	0.1917	0.2216	2.0e-5	0.2529
t -stat	(9.99)	(9.41)	(9.85)	(8.29)	(10.99)
p -value	***	***	***	***	***
Adj. R²	0.3741	0.2936	0.3191	0.522	0.5043
reject H_0					
$H_0 : \beta_1 < \beta_3$ vs $H_1 : \beta_1 > \beta_3$	<i>yes</i>	<i>no</i>	<i>no</i>	<i>yes</i>	<i>no</i>
$H_0 : \beta_1 > \beta_3$ vs $H_1 : \beta_1 < \beta_3$	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>yes</i>
SWPTS	BCH	BTC	EOS	ETH	XRP
$L_{M,1}$ "high volatility"	0.0734	0.0166	0.006	0.0792	-0.0220
t -stat s	(3.17)	(0.76)	(0.22)	(2.87)	(-1.01)
p -value	**	ns	ns	**	ns
$L_{M,3}$ "low volatility"	0.1842	0.1952	0.2700	0.2080	0.2440
t -stat	(11.87)	(12.48)	(17.66)	(12.67)	(14.33)
p -value	***	***	***	***	***
Adj. R²	0.3063	0.2620	0.2522	0.2852	0.4232
reject H_0					
$H_0 : \beta_1 < \beta_3$ vs $H_1 : \beta_1 > \beta_3$	<i>yes</i>	<i>no</i>	<i>no</i>	<i>yes</i>	<i>no</i>
$H_0 : \beta_1 > \beta_3$ vs $H_1 : \beta_1 < \beta_3$	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>yes</i>
EWDEPTH	BCH	BTC	EOS	ETH	XRP
$L_{M,1}$ "high volatility"	-0.0639	-0.0676	- 0.0089	-0.0417	0.0085
t -stat	(-3.25)	(-5.47)	(-0.49)	(-2.12)	(0.50)
p -value	**	***	ns	*	ns
$L_{M,3}$ "low volatility"	0.0775	0.0320	0.0707	0.0356	0.0554
t -stat	(4.20)	(4.00)	(4.85)	(3.11)	(4.84)
p -value	***	***	***	**	***
Adj. R²	0.0824	0.1236	0.0839	0.1043	0.1148
reject H_0					
$H_0 : \beta_1 < \beta_3$ vs $H_1 : \beta_1 > \beta_3$	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>
$H_0 : \beta_1 > \beta_3$ vs $H_1 : \beta_1 < \beta_3$	<i>yes</i>	<i>yes</i>	<i>no</i>	<i>yes</i>	<i>no</i>

Newey and West's (1994) heteroskedasticity and autocorrelation robust standard errors are calculated using automatic lag truncation parameter selection. Significance levels: *** ≤ 0.001 , ** ≤ 0.01 , * ≤ 0.05 and ns > 0.05 .

Chapter 6

Conclusion

Market quality and cryptocurrencies are two subjects that have known a growing interest in recent academic literature, but the topic of market quality of cryptocurrencies was barely researched before and only by very recent studies. Our motivation with this master's thesis was to extend those two bodies of literature by updating and verifying if findings of market quality-related academic literature were also verified with recent cryptocurrency data. In this process, we tried to answer to the problematic : "What findings of the academic literature about the market quality co-movements of traditional assets can be transferred to cryptocurrencies in 2017/2018 ?". Using intraday transactions and order book data of five cryptocurrencies on the Bitfinex exchange, we investigated some of the main findings of related academic literature within the limits of our datasets. Specifically, we researched for the presence of a common factor in the market quality movements of our set of cryptocurrencies using Principal Component Analysis (one of the two main methodologies used in the literature to assess commonality and popularized as such by Hasbrouck and Seppi (2001)) using a set of proxies proposed by Marsh and Mazza (2017) to capture the main dimensions of market quality. We used two important events of the recent Bitcoin's history, the introduction of trades on futures product and the end of Bitcoin bubble as dated by Liu et al. (2019) to define several sub-periods and compare the results of the Principal Component Analysis across those periods. We also compare a part of our results with Liu et al. (2019) and Christoffersen, Lunde, and Olesen (2019), two recent works studying the common factor structure in cryptocurrencies and commodity universe respectively and using the same method to assess commonality. In a second time, we used Chordia, Roll, and Subrahmanyam (2000)'s market model (the second main methodology in academic literature to assess commonality) to test for the presence of liquidity commonality for our set of cryptocurrencies in our period and study the impact of volatility regimes on commonality by reproducing the approach of Beaupain, Giot, and Petitjean (2010). Finally, we made statistical tests of means to assess the evolution of the explanatory power of the market model across our period.

The results of the PCA on the whole periods presented high values for the first principal components of every proxies of the dimensions of market quality, suggesting (1) the presence of a common factor in market quality movements across our set of cryptocurrencies between August 2017 and July 2018. Our temporal comparisons of the PCA results between periods highlighted higher principal components during the three last months of our period compared to the three first months, higher principal components for the three months following the introduction of Bitcoin futures compared to the three months preceding it and higher principal components

in the month following the end of the Bitcoin bubble dated by Liu et al. (2019) compared to the month preceding it. Those results suggests that (2) the common movements in the market quality of our set of cryptocurrencies were higher during the last three months of our period compared to the first three months (which is a counter-intuitive behaviour given that literature has found that the market quality of cryptocurrencies was improving with time), (3) that common movements in the market quality of our set of cryptocurrencies were higher during the three months following the introduction of futures product on Bitcoin compared to the three months preceding it (which is also counter-intuitive as literature has shown that option listing improved market quality) and (4) that the common movements in the market quality of our set of cryptocurrencies were higher during the month following the end of Bitcoin bubble compared to the month preceding it. Those results seem to be robust to a change in aggregation rule, but not to a removal of the standardization procedures. As expected, the standardization procedure (and thus intraday effects) seems to have an important impact on commonality, and this impact appears to vary between proxies. The second part of our work leads to three main findings. First, (5) at a confidence level of 95%, we conclude to the presence of liquidity commonality for our five cryptocurrencies during our period of interest in both quiet and volatile markets. Our second result (6) is that like Beaupain, Giot, and Petitjean (2010), we found differences in the magnitude of the commonality between high and low volatility regimes, with significant evidence of a positive relationship between the commonality in liquidity and volatility at a confidence level of 95%. Finally, (7) our results show that the common variations depend on the type of liquidity proxy chosen, quote-based measures leading to contrasting results compared to effective or traded spread measures. The third part of our research was less conclusive. We showed that the results of the average adjusted- R^2 of the regressions varied importantly based on the time-frame of the fit of the regressions, with daily regressions having very small explanatory power in average compared to regressions fitted on the whole periods (weekly fits improve the average explanatory power compared to daily fits but they are still smaller compared to the regressions fitted on the whole period). We attribute this gap to the immaturity of the cryptocurrencies market. Our results are mixed for the comparison between the beginning and the end of the sample and between the periods following and preceding the introduction of futures on Bitcoin, but we found (8) statistical evidence that the explanatory power of the model for four of the Bitcoin's liquidity proxies was significantly better in average during the three first months of our period compared to the last three months at a confidence level of 95%. However, based on the p-values of the other tests and the average adjusted- R^2 during the different periods, the results hints that the behaviour of the explanatory power of the models for the other four cryptocurrencies was opposite to Bitcoin during this period. Finally, we found (9) that the explanatory power of the model for 16 liquidity proxies was significantly better in average during the month preceding the end of Bitcoin bubble compared to the month following based on weekly fits and 6 proxies based on daily fits. Those results tend to corroborate those of Liu et al. (2019), who showed that the explanatory power of their model was better during the Bitcoin bubble.

On retrospect, there are four main limits to the approach we followed. First is that we work on a restricted set of cryptocurrencies. In order to be more consistent, we could have included more than the top five cryptocurrencies in terms of market cap. Furthermore, we could have extended our analyses beyond Bitfinex exchange and include data from other exchanges. Second is that we do treat completely the

problem of missing data for the order books, mainly at the end of January. In their research, Liu et al. (2019) are faced with the same problem and they use linear interpolations to estimate the values for small missing intervals and a modified approach of Hansen and Lunde (2005) to account for larger breaks. This should be an interesting complement to our master's thesis. Third is that we used data of cryptocurrencies against USD and did not take inflation into account in our analyses. Ideally when studying common movements, cryptocurrencies against Bitcoin for example would be preferable as we get rid of the commonality linked to the dollar. Lastly, we highlighted the presence of commonality but did not investigate the sources driving those common movements, which would be a natural complement to our research given academic literature. We also would have liked to be able to conduct more robustness checks for the results of liquidity commonality but were unable to do so due to a lack of time.

In general, the market quality of cryptocurrencies being barely studied before, they are a lot of areas where it could be interesting to conduct further researches. All along our work we highlighted a series of those topics. First is a thorough research in the intraday patterns of the market quality of cryptocurrencies. Comparison between our results before and after standardization lets us think that intraday movements in this area have an important impact on commonality. Secondly, we think that literature is missing a broad comparative study in the commonality of various type of assets including cryptocurrencies, commodities and other types of assets on the same period. Finally, we think that investigating the sources driving common movements in market quality could have a lot of interesting practical applications and would also shed further lights on the impact of several events like the introduction of futures or the end of the Bitcoin bubble on common movements.

Appendix A

Proxies used by Marsh and Mazza (2017) to capture market quality

We take the six main proxies of market quality used by Marsh and Mazza (2017) in their research. The authors computed other proxies meant as robustness checks, but we only use the six main proxies that they use in their analysis. For each dimension of market quality in the following sections, we give the definitions and formulas used to compute the proxies designed to measure the dimensions.

A.1 Cost of immediacy

The authors use several similar proxies for spread among which Effective spread, Real Spread and Realized Spread but they focus on the last one in their text. *"The relative spread is a measure of the current bid-ask spread and captures the direct cost of immediacy for small quantities"* Marsh and Mazza (2017). This proxy is computed based on quotes and is related to ex-ante market quality. Here is the formula used by the authors that we reproduced to compute the Realized Spread :

$$RS_{j,t} = (a_{j,t} - b_{j,t}) / m_{j,t} \quad (\text{A.1})$$

Where $RS_{j,t}$ is the Realized Spread at time t for the cryptocurrency j and $a_{j,t}$, $b_{j,t}$, $m_{j,t}$ are respectively the ask price, bid price and mid-price of the quote prevailing at time t . The mid-price is computed as below by taking the bid and ask prices at the best limit :

$$m_{j,t} = (a_{j,t} + b_{j,t}) / 2 \quad (\text{A.2})$$

A.2 Adverse selection

According to Hendershott and Moulton (2011), the effective spread can be divided into a transitory component and a permanent component. Marsh and

Mazza (2017) use the Realized Spread to estimate the transitory component and the Percentage Price Impact for the permanent component with the following formula :

$$PPI_{i,t} = 2q_{i,t}(m_{i,j,t+5} - m_{i,j,t})/m_{i,j,t} \quad (\text{A.3})$$

Where $PPI_{i,t}$ is the Percentage Price Impact of the trade i at time t , $q_{i,t}$ is a variable that reflects trade direction (it is equal to 1 if the trade was initiated by a buyer and -1 if the trade was initiated by a seller), $m_{i,j,t}$ is the mid-price of the asset j prevailing at time t for trade i and $m_{i,j,t+5}$ denotes the mid-price of the fifth reported record after trade i . To compute $q_{i,t}$, Marsh and Mazza (2017) follow Lee and Ready (1991)'s tick test and the correction of Ellis, Michaely, and O'Hara (2000) but we don't need to use this procedure as the information about the trade direction is already known in our datasets with certainty. On the other hand, to compute the Percentage Price Impact, the authors first merge the trade and order book datasets by matching a trade with the quote prevailing at the time of the trade. This is something that would be problematic to reproduce in our case as there is a discrepancy in the frequency of observations between our two datasets as explained in section 3.2. Consequentially, we are unable to compute the Percentage Price Impact of each trade as proposed by the authors because we need data both from the trade and order book datasets at any time. To solve this problem, we use the price of the trade instead of the mid-price of the quote, which gives us the following formula :

$$PPI_{i,t} = 2q_{i,t}(p_{i,j,t+5} - p_{i,j,t})/p_{i,j,t} \quad (\text{A.4})$$

Where $p_{i,j,t}$ is the price of trade i of asset j at time t and $p_{i,j,t+5}$ is the price of the fifth trade recorded after trade i . This workaround is not ideal as we do not take into account the bid-ask bounces, but it is close to the other proxy.

A.3 Liquidity

To capture liquidity, Marsh and Mazza (2017) use the modification of Amihud (2002)'s ratio proposed by Karolyi, Lee, and Van Dijk (2012). Amihud (2002)'s ratio of illiquidity is the response of the price to a unitary volume change and is computed as follow :

$$IL_t = \frac{|r_t|}{v_t} \quad (\text{A.5})$$

Where IL_t is the illiquidity ratio on interval t , r_t is the return over interval t and v_t is the aggregated volume on interval t . Karolyi, Lee, and Van Dijk (2012) modify this ratio by applying the natural logarithmic transformation to reduce outliers' impact and multiplying the ratio by -1 to make the relation between the ratio and liquidity positive. This gives us the following formula :

$$Liq_t = -\ln(1 + \frac{|r_t|}{v_t}) \quad (\text{A.6})$$

Where Liq_t is the modified liquidity ratio at time t .

A.4 Price efficiency

To estimate Price Efficiency, the authors use the absolute value of the mid-prices return autocorrelation. They use quotes mid-points in order to avoid bid-ask bounces and compute returns on 5 minutes intervals. The return autocorrelation is then computed on 30-minutes intervals. They take the absolute value to measure the extent of the deviation from a random walk (in a random walk, return autocorrelation are equal to zero). We use a lag of 1 to compute autocorrelation.

A.5 Volatility

The authors use two proxies to assess volatility. First is the Realized Volatility, which is computed by aggregating movements tick by tick over 30-minutes intervals then using the following formula :

$$RV_T = \sqrt{\sum_{i=1}^N r_t^2} \quad (\text{A.7})$$

Where RV_T is the Realized Volatility associated with the 30-minute interval T , r_t is the tick by tick return at time t and N is the ensemble of trades that occurred during T . The authors also use a second proxy, the quote return volatility for completeness, computed as the standard deviation over 30-minutes intervals of the mid-points returns computed over 5 minutes intervals.

Appendix B

Description of the variables of the datasets

B.1 Trade datasets

This part of our datasets contains the information of every trade of our five cryptocurrencies paired with US dollar recorded on the Bitfinex exchange (the timeframe of the observations for each cryptocurrency vary as discussed in section 4.1). Below are the names of the variables available as stated in the trade datasets and a short description of what they represent :

- "datetime" : datetime of the trade
- "id" : unique id of the trade
- "exchange" : exchange on which the trade was recorded (here Bitfinex for all trades)
- "symbol" : symbol of the asset traded
- "price" : price of the trade for one unit of the asset
- "amount" : amount of the asset traded
- "sell" : binary variable for the trade direction (sell if true, buy if false)
- "day" : day of the trade
- "hour" : hour of the trade
- "volume" : volume of the trade

B.2 Order Book

Here are the names of the variables as stated in the order book datasets and a short description of what they represent :

- "datetime" : datetime of the observation

- "PB01" : price of the best bid
- "QB01" : quantity of the best bid
- "PB02" : price of the bid at the second limit
- "QB02" : quantity of the bid at the second limit
- "PB03" : price of the bid at the third limit
- "QB03" : quantity of the bid at the third limit
- "PB04" : price of the bid at the fourth limit
- "QB04" : quantity of the bid at the fourth limit
- "PB05" : price of the bid at the fifth limit
- "QB05" : quantity of the bid at the fifth limit
- "PB06" : price of the bid at the sixth limit
- "QB06" : quantity of the bid at the sixth limit
- "PB07" : price of the bid at the seventh limit
- "QB07" : quantity of the bid at the seventh limit
- "PB08" : price of the bid at the eighth limit
- "QB08" : quantity of the bid at the eighth limit
- "PB09" : price of the bid at the ninth limit
- "QB09" : quantity of the bid at the ninth limit
- "PB10" : price of the bid at the tenth limit
- "QB10" : quantity of the bid at the tenth limit
- "PA01" : price of the best ask
- "QA01" : quantity of the best ask
- "PA02" : price of the ask at the second limit
- "QA02" : quantity of the ask at the second limit
- "PA03" : price of the ask at the third limit
- "QA03" : quantity of the ask at the third limit
- "PA04" : price of the ask the fourth limit
- "QA04" : quantity of the ask at the fourth limit
- "PA05" : price of the ask at the fifth limit
- "QA05" : quantity of the ask at the fifth limit
- "PA06" : price of the ask at the sixth limit

- "QA06" : quantity of the ask at the sixth limit
- "PA07" : price of the ask at the seventh limit
- "QA07" : quantity of the ask at the seventh limit
- "PA08" : price of the ask at the eighth limit
- "QA08" : quantity of the ask at the eighth limit
- "PA09" : price of the ask at the ninth limit
- "QA09" : quantity of the ask at the ninth limit
- "PA10" : price of the ask at the tenth limit
- "QA10" : quantity of the ask at the tenth limit

Appendix C

Description of our set of cryptocurrencies

In this Appendix, we briefly describe the five cryptocurrencies we used in our research and give insights on their history and market cap evolution. The Appendix is organized in five sections, one for each cryptocurrency.

C.1 Bitcoin (BTC)

Bitcoin was developed by Satoshi Nakamoto (pseudonym). In 2008, the anonymous author(s) presented the currency in a white paper titled *Bitcoin: A Peer-to-Peer Electronic Cash System*¹ which later became an open-source project. By January 2009, the Blockchain was born and the bitcoins were used by third parties. All documentation related to the Blockchain, transactions, mining, contracts, code and other repositories are made available online for free^{2 3}. Only two years after its implementation, alternative cryptocurrencies or altcoins started appearing progressively. In figure C.1, we show the evolution of the market capitalization of BTC (in hundreds of billions of USD) for our time scope using an online source⁴.

We can see that the market capitalization has varied a lot over our period. From late September 2017 to mid-December 2017 the capitalization of BTC has more than tripled, reaching well over \$300 billion. From January 2018 to November 2018, the capitalization falls back below the \$150 billion mark. In the background, the number of BTC has been increasing (even though more slowly as time passes) over the entire period.

Another interesting feature is that Bitcoin's block size of 1MB suffers of scalability issues. The number of transactions that can be handled is rather limited in time (under 7 transactions per second). For the sake of comparison, Visa (the payment network), could handle up to 47,000 transactions in 2013 (Poon and Dryja, 2016). For a transaction size of 300 bytes, this would take a processing rate of 8GB per block every ten minutes. This would generate hundreds of TB of data and thus centralize the network to use only the nodes with such storage capacities, which

¹<http://bitcoin.org/bitcoin.pdf>

²<http://bitcoin.org>

³<http://github.com/bitcoin>

⁴<http://www.blockchain.com/charts>

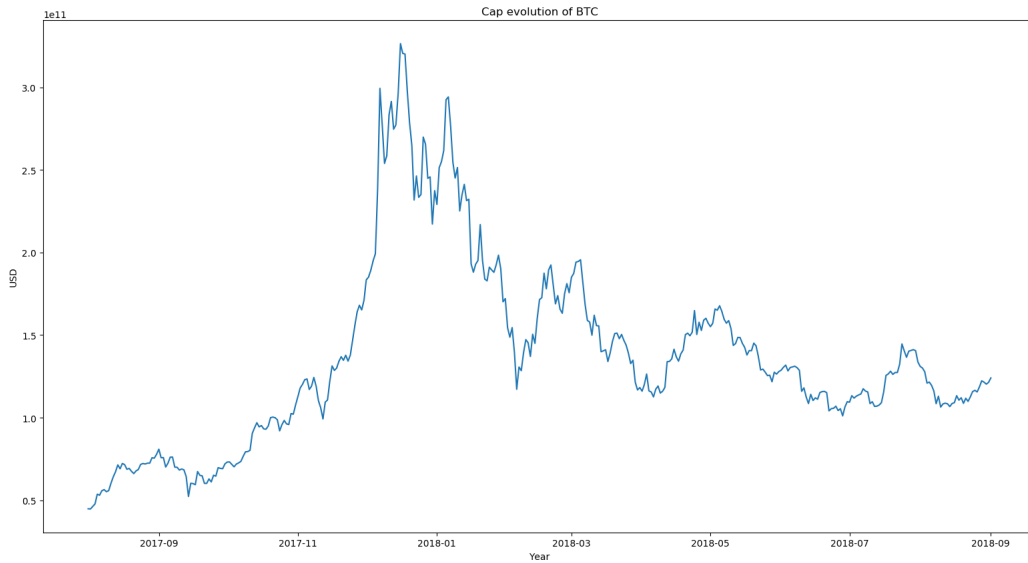


FIGURE C.1: BTC market capitalization

would completely defeat the purpose of BTC and the Blockchain (Vujičić, Jagodić, and Randić, 2018)

C.2 Bitcoin Cash (BCH)

Bitcoin Cash is a recent spin-off of Bitcoin, as it was created in 2017. A divergent branch appeared in the Bitcoin community as some members wanted to reduce transaction fees by increasing the block size. They achieved their purpose, as the fees for BCH are lower than those of credit cards and there are no chargebacks nor any other extra fees for transactions. The block size was changed to 8MB, resulting in much faster confirmation times. Since it is a branch that derived from Bitcoin, both chains share the same information up to the date of the separation.

As we can see on Figure C.2, the market capitalization of BCH increased tenfold from November 2017 to December 2017, reaching over \$ 60 billion. Starting from 2018, the market capitalization plummets and the capitalization was back at \$10 billion by the first quarter of the year. It then recovered to the \$30 billion mark by May, before falling slowly back to \$10 billion by the end of November 2018.

C.3 Eos (EOS)

EOS is another recent altcoin. In 2018, "block.one", a private company provider of blockchain solutions released a white paper developing a new blockchain protocol⁵. EOS differs from the Bitcoin root in various aspects. Mainly, EOS cannot be mined whereas bitcoins can be mined by those who put the resources for all the

⁵<http://eos.io/about-us>

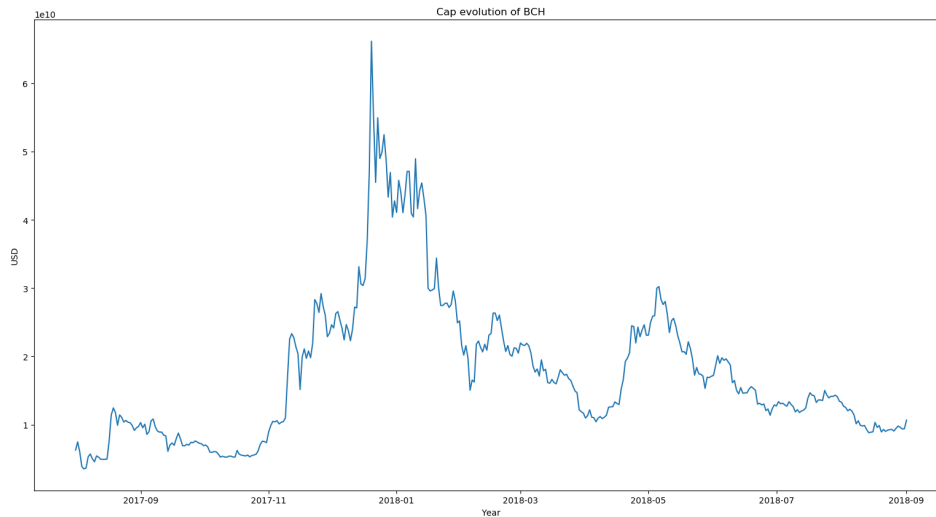


FIGURE C.2: BCH market capitalization

computations on the blockchain. EOS also uses a different consensus algorithm, the "delegated proof of stake" whereas Bitcoin uses the "proof of work".

As we can see on Figure C.3, we observe two peaks within our timeframe for the market capitalization of EOS. By January 2018, EOS reached \$10 billion before dropping by 50% within a month. Then from April 2018 to May 2018, the capitalization nearly tripled in value reaching \$17 billion. From that point, the cap went back to the \$10 billion mark in four months.

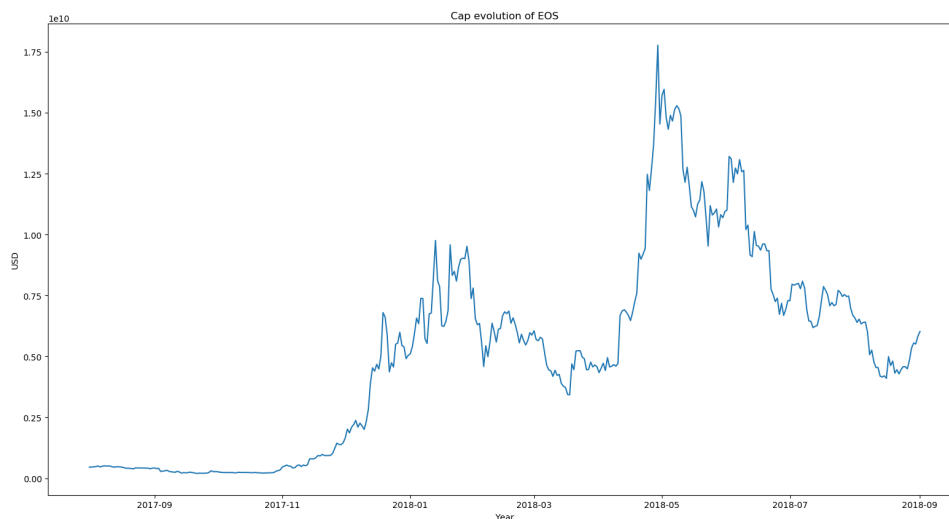


FIGURE C.3: EOS market capitalization

C.4 Ethereum (ETH)

Vitalik Buterin presented Ethereum in his white paper in 2015⁶. This paper tackled multiple limitations of Bitcoin. Although there are some issues regarding scalability, the Ethereum network has managed to reach 11 transactions per second (Vujičić, Jagodić, and Randić, 2018). Nowadays, Bitcoin and Ethereum are the most capitalized cryptocurrencies.

As we can see on Figure C.4, the capitalization of ETH tripled from November 2017 to late January 2018 reaching over \$130 billion. By April of the same year, it had already come back below the \$40 billion mark. Then, it recovered to a little over \$80 billion before crashing back again, below the \$40 billion mark.

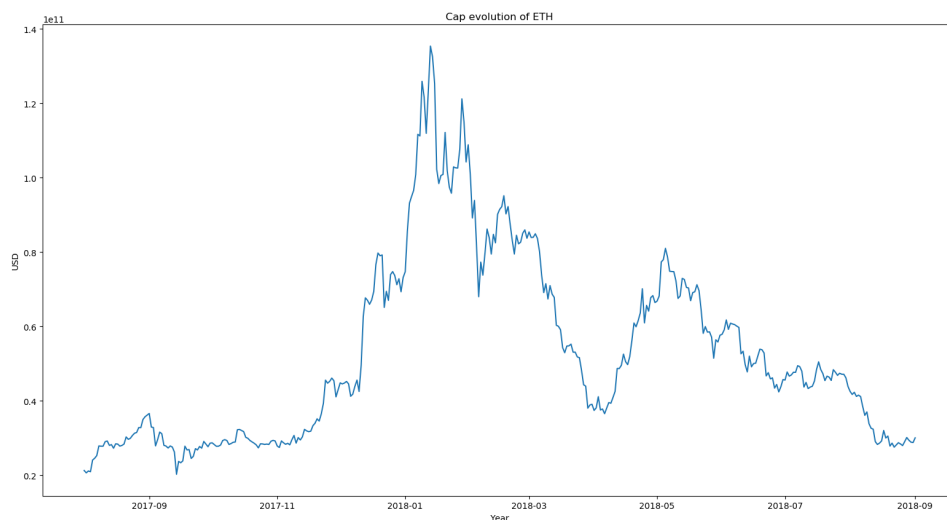


FIGURE C.4: ETH market capitalization

C.5 Ripple (XRP)

Ripple was released in 2012, by "Ripple Labs Inc". The main difference with the other cryptocurrencies is that Ripple is mainly known as a digital means of payment using the blockchain, that works on distributed consensus mechanism, that uses negligible power and that is mainly used by financial institutions⁷.

As we can see on Figure C.5, the capitalization of XRP remained stable at \$1 billion, from September to November 2017, before skyrocketing up to \$13 billion in January 2018. One month later, it had crashed back to \$2 billion. Other than some short-lasting come backs in late February and May, the cap remained overall downward sloping almost hitting the \$1.3 billion in November 2018.

⁶http://blockchainlab.com/pdf/Ethereum_white_paper-a_next_generation_smart_contract_and_decentralized_application_platform-vitalik-buterin.pdf

⁷<https://www.ripple.com/company/>

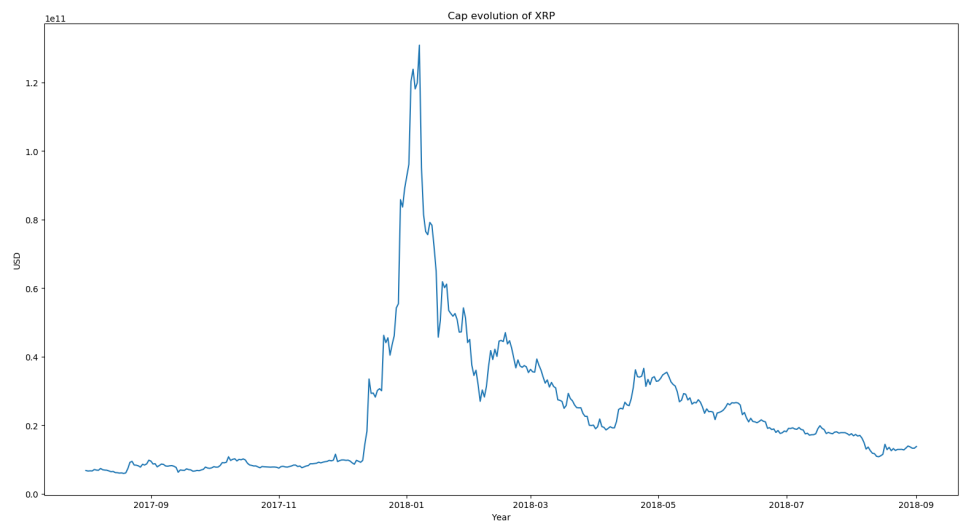


FIGURE C.5: XRP market capitalization

Appendix D

Most significant time skips in the order book datasets

	datetime	PB01	QB01	PB02	QB02
8887	26/01/2018 04:13:09	11525	2.12809	11520	0.09999
8888	26/01/2018 04:14:09	11522	0.886679	11521	0.93
8889	26/01/2018 04:15:09	11521	0.961038	11520	10
8890	26/01/2018 04:16:09	11519	8	11512	0.669062
8891	26/01/2018 04:17:09	11511	2.87579	11510	1.41041
8892	26/01/2018 04:18:09	11496	0.420787	11495	0.166
8893	26/01/2018 04:19:09	11467	1.9209	11466	2.658
8894	26/01/2018 04:20:09	11419	1.45405	11416	0.538477
8895	26/01/2018 04:21:09	11439	2.88285	11435	0.00871...
8896	26/01/2018 04:22:03	11441	0.345726	11440	4.0978
8897	26/01/2018 04:25:04	11453	0.783464	11451	10.0467
8898	26/01/2018 04:27:09	11437	0.53	11434	0.19
8899	26/01/2018 04:28:09	11455	0.80273	11453	0.3
8900	28/01/2018 12:57:12	11750	1.27903	11749	0.004
8901	28/01/2018 12:58:15	11748	1.00468	11746	6.74276
8902	28/01/2018 12:59:15	11745	0.291629	11744	0.06
8903	28/01/2018 13:00:15	11726	0.05177...	11725	2.64889
8904	28/01/2018 13:01:15	11722	0.7	11720	1
8905	28/01/2018 13:02:16	11740	2.1497	11739	4.90886

FIGURE D.1: Significant time skip in BTC order book

	datetime	PB01	QB01	PB02	QB02	PB03
5091	28/01/2018 12:53:19	1725.2	17.502	1724.6	0.883676	1724.3
5092	28/01/2018 12:54:19	1724.1	0.99	1724	16.5999	1722.2
5093	28/01/2018 12:55:19	1722.3	4.789	1720.7	1.46009	1720.2
5094	28/01/2018 12:56:19	1721	2.86	1720.2	3.2173	1720
5095	28/01/2018 12:58:12	1725	4	1724.9	19	1724.1
5096	28/01/2018 12:59:12	1728.3	20.8952	1727.1	0.2132	1726
5097	28/01/2018 13:00:12	1727.2	0.385368	1725.9	0.2118	1725.8
5098	28/01/2018 13:01:12	1721.7	3	1721.6	0.1865	1719.1
5099	28/01/2018 13:02:12	1721.8	4.7204	1721.6	10	1719.3
5100	30/01/2018 06:20:00	1608	3.548	1607.8	10	1607.2
5101	30/01/2018 06:21:00	1609.1	10	1609	8.3856	1608.7
5102	30/01/2018 06:22:00	1609.1	0.1721	1609	1.7322	1608.9
5103	30/01/2018 06:23:00	1608.2	0.0002	1608	0.73	1605.6
5104	30/01/2018 06:24:00	1606.9	0.248924	1606.3	0.68595	1605.7
5105	30/01/2018 06:25:00	1608	7.87621	1607.7	0.234997	1607
5106	30/01/2018 06:26:00	1609.5	5.178	1608.2	0.413822	1608.1
5107	30/01/2018 06:27:00	1610.1	2.853	1609.6	0.1721	1609.5
5108	30/01/2018 06:28:00	1607.8	0.1675	1607.7	3	1607.6
5109	30/01/2018 06:29:00	1608.4	4	1608.3	8.3755	1607.7

FIGURE D.2: Significant time skip in BCH order book

	datetime	PB01	QB01	PB02	QB02
1662	21/01/2018 03:43:11	14.708	1340.15	14.7	86.208
1663	21/01/2018 03:44:11	14.703	138.33	14.647	108.577
1664	21/01/2018 03:45:11	14.636	421.963	14.635	18.2835
1665	21/01/2018 03:46:11	14.573	43.413	14.572	7
1666	21/01/2018 03:47:11	14.512	411.019	14.511	180.858
1667	21/01/2018 03:48:11	14.664	167.557	14.637	715.46
1668	21/01/2018 03:49:11	14.613	225.689	14.576	183.157
1669	21/01/2018 03:50:11	14.608	5.13389	14.574	50
1670	21/01/2018 03:51:05	14.546	523.76	14.545	523.76
1671	21/01/2018 03:52:02	14.491	200	14.48	747.52
1672	21/01/2018 03:52:03	14.491	200	14.48	747.52
1673	22/01/2018 13:38:00	13.061	7.7423	13.05	50
1674	22/01/2018 13:39:00	13.141	125.329	13.14	930.73
1675	22/01/2018 13:40:00	13.122	559.54	13.1	100
1676	25/01/2018 10:39:03	14.189	55.3327	14.179	360.82
1677	25/01/2018 10:40:03	14.184	200	14.18	11.21
1678	25/01/2018 10:41:03	14.186	1091	14.172	4
1679	25/01/2018 10:42:03	14.225	750.18	14.217	364.43
1680	25/01/2018 10:43:03	14.206	360.82	14.195	360.82

FIGURE D.3: Significant time skip in EOS order book

	datetime	PB01	QB01	PB02
7818	25/01/2018 10:17:03	1.3453	4500	1.345
7819	25/01/2018 10:18:03	1.348	3402.11	1.3467
7820	25/01/2018 10:19:03	1.3442	985.291	1.3441
7821	25/01/2018 10:20:03	1.3418	1500	1.3417
7822	25/01/2018 10:21:03	1.344	799.367	1.3439
7823	25/01/2018 10:22:03	1.341	384.559	1.3402
7824	25/01/2018 10:23:03	1.341	4325.24	1.3397
7825	25/01/2018 10:24:03	1.34	5980.84	1.3399
7826	25/01/2018 10:25:03	1.3414	4025.11	1.341
7827	25/01/2018 10:26:03	1.3414	3298.59	1.3405
7828	25/01/2018 10:27:03	1.3419	22262	1.3416
7829	25/01/2018 10:28:03	1.3438	900	1.3437
7830	26/01/2018 04:23:02	1.302	50	1.3005
7831	26/01/2018 04:24:04	1.3029	1778.8	1.3028
7832	26/01/2018 04:26:04	1.3005	1851.84	1.3004
7833	26/01/2018 04:26:09	1.3004	156.136	1.2993
7834	26/01/2018 04:27:07	1.3004	81.7833	1.2992
7835	26/01/2018 04:27:10	1.3004	81.7833	1.299
7836	26/01/2018 04:28:05	1.3004	2551.6	1.3003

FIGURE D.4: First significant time skip in XRP order book

	datetime	PB01	QB01	PB02
11222	28/01/2018 12:55:16	1.252	1511.02	1.2519
11223	28/01/2018 12:56:16	1.2503	8456.72	1.2501
11224	28/01/2018 12:57:14	1.2505	5000	1.2503
11225	28/01/2018 12:57:25	1.2519	1459.22	1.2518
11226	28/01/2018 12:58:25	1.2535	447.389	1.2531
11227	28/01/2018 12:59:25	1.252	100476	1.2519
11228	28/01/2018 13:00:25	1.2521	24143.4	1.252
11229	28/01/2018 13:01:25	1.2521	17726.9	1.252
11230	29/01/2018 12:58:00	1.2721	5000	1.272
11231	29/01/2018 12:59:00	1.267	25000	1.2666
11232	29/01/2018 13:00:00	1.268	1470	1.267
11233	29/01/2018 13:01:00	1.265	8520.82	1.2646
11234	29/01/2018 13:02:00	1.265	10.18	1.2644
11235	29/01/2018 13:03:00	1.2644	10	1.2643
11236	29/01/2018 13:04:00	1.2648	1922.73	1.2645
11237	29/01/2018 13:05:00	1.2672	5908.95	1.267
11238	29/01/2018 13:06:00	1.2672	1265.6	1.267
11239	29/01/2018 13:07:00	1.2672	798.962	1.267
11240	29/01/2018 13:08:00	1.2672	2298.93	1.2663

FIGURE D.5: Second significant time skip in XRP order book

	datetime	PB01	QB01
12263	30/01/2018 06:11:01	1.2185	100
12264	30/01/2018 06:12:01	1.2174	4500
12265	30/01/2018 06:13:01	1.2179	78.611
12266	30/01/2018 06:14:01	1.2173	10182.5
12267	30/01/2018 06:15:01	1.2173	12282.2
12268	30/01/2018 06:16:01	1.2201	3322.39
12269	30/01/2018 06:17:01	1.2216	400
12270	30/01/2018 06:18:01	1.2217	1556.91
12271	02/02/2018 05:03:04	0.8681	356.43
12272	02/02/2018 05:04:04	0.8692	81.12
12273	02/02/2018 05:05:04	0.8681	94710.7
12274	02/02/2018 05:06:04	0.8656	1772.42
12275	02/02/2018 05:07:04	0.86521	1531.54
12276	02/02/2018 05:08:04	0.8655	16018.8
12277	02/02/2018 05:09:04	0.86501	995.862
12278	02/02/2018 05:10:04	0.86501	3468.88
12279	02/02/2018 05:11:04	0.86017	100
12280	02/02/2018 05:12:04	0.86831	89.4085
12281	02/02/2018 05:13:04	0.8708	4251.29

FIGURE D.6: Third significant time skip in XRP order book

	datetime	PB01	QB01	PB02	QB02
3683	22/01/2018 13:30:11	976.51	8.00109	976	2.91327
3684	22/01/2018 13:31:11	976.11	28.0985	976.03	13
3685	22/01/2018 13:32:11	984.1	13.9703	984	100.098
3686	22/01/2018 13:33:11	986.1	1.8729	985.15	71.018
3687	22/01/2018 13:34:11	982	3.37558	980	0.051
3688	22/01/2018 13:35:11	986.38	29.9778	986.21	2.72633
3689	22/01/2018 13:36:11	989.9	96.5415	988.31	0.586465
3690	22/01/2018 13:37:11	980.21	10.1925	980	2
3691	22/01/2018 13:38:00	975.91	35.4	975.9	6.488
3692	22/01/2018 13:39:00	982.92	0.02	982.79	5.31958
3693	22/01/2018 13:40:00	979.42	0.02	979.41	0.02
3694	25/01/2018 10:30:07	1071.4	9.3467	1071	76.083
3695	25/01/2018 10:31:07	1067.6	0.1	1067.4	6.9
3696	25/01/2018 10:32:07	1067.6	3.42884	1067	8.37
3697	25/01/2018 10:33:07	1067.5	10.0937	1067	8.37
3698	25/01/2018 10:34:07	1064.3	6	1064	8.00595
3699	25/01/2018 10:35:07	1065.5	0.155216	1065.2	4.36512
3700	25/01/2018 10:36:07	1065.2	0.1	1065	11.1421
3701	25/01/2018 10:37:07	1065.1	0.4548	1065	2.87495

FIGURE D.7: First significant time skip in ETH order book

	datetime	PB01	QB01	PB02	QB02
4760	26/01/2018 04:17:03	1068.6	8.9238	1068.1	8.44848
4761	26/01/2018 04:18:03	1069.3	0.04208...	1069.2	0.814906
4762	26/01/2018 04:19:03	1067.3	1.9746	1066.7	0.340004
4763	26/01/2018 04:20:03	1063.4	6.34	1063.2	0.5
4764	26/01/2018 04:21:03	1061.8	6	1061.2	10.2693
4765	26/01/2018 04:22:03	1062.3	9.33994	1062.2	4
4766	26/01/2018 04:24:06	1063	0.00090...	1062	2.5605
4767	26/01/2018 04:26:10	1064.8	0.106909	1063.6	1.0324
4768	26/01/2018 04:27:15	1064	2.51	1063.7	6.99929
4769	26/01/2018 04:28:05	1065	15.45	1064.9	8.9656
4770	28/01/2018 13:03:14	1222.1	3.23639	1222	26.2047
4771	28/01/2018 13:04:14	1223	12.9	1222.6	30.4
4772	28/01/2018 13:05:14	1226.1	0.5	1226	32.3474
4773	28/01/2018 13:06:14	1228.3	10.721	1228.2	0.301711
4774	28/01/2018 13:07:14	1228.8	6.03472	1228.6	0.372018
4775	28/01/2018 13:08:14	1230.1	0.8	1230	12.0373
4776	28/01/2018 13:09:14	1229.9	0.449287	1229.5	2.1681
4777	28/01/2018 13:10:14	1230	31.7045	1229.1	11.1681
4778	28/01/2018 13:11:14	1226.5	0.2	1226.4	8.1562

FIGURE D.8: Second significant time skip in ETH order book

Appendix E

Summary statistics of the trade and order books datasets

Statistics	Bitcoin (BTC)	Bitcoin Cash (BCH)	Eos (EOS)	Ripple (XRP)	Ethereum (ETH)
Mean	0.4360	2.0038	219.7460	1 602.5399	3.8739
Standard deviation	2.0893	7.0553	842.9014	4 700.0935	14.9210
Min	0.0000	0.0000	0.0000	0.0000	0.0000
Trade Amount	0.0200	0.1133	11.2546	99.9050	0.2575
50% quantile	0.0940	0.5000	50.0000	424.1670	1.0000
75% quantile	0.3300	1.7176	175.2500	1 180.4000	3.4984
Max	1 817.0600	1 975.6100	50 000.0000	200 000.0000	4 999.2000
Mean	9 482.3155	1 354.4715	9.5924	1.0422	617.6807
Standard deviation	3 617.6048	709.5330	4.7373	0.6518	266.0743
Min	2 980.4000	270.0100	0.4760	0.1474	198.0000
Trade Price	6 926.7000	807.0000	6.8699	0.6318	406.3700
25% quantile	8 790.5000	1 263.3000	9.0735	0.8915	575.5800
50% quantile	1 1461.0000	1 630.2000	13.0320	1.3027	784.1000
75% quantile	19 891.0000	4 104.3000	23.0290	3.3000	1 424.3000
Max	3 712.2091	2 265.9407	1 716.1565	1 370.6599	2 191.0848
Mean	16 856.1427	7 611.5643	6 277.1948	4 336.9254	8 362.6128
Standard deviation	0.0000	0.0000	0.0000	0.0000	0.0000
Min	182.2940	149.4910	92.6669	78.9938	146.5370
Trade Volume	787.9280	587.8000	411.2055	356.0830	544.6140
25% quantile	2 889.0000	1 969.4200	1 493.0000	1 137.9200	1 983.2400
50% quantile	14 899 900.00	2 195 940.0000	999 641.0000	535 988.0000	3 802 400.0000
75% quantile					
Max					

TABLE E.1: Summary statistics of the trades for the five cryptocurrencies

Statistics	Bitcoin (BTC)	Bitcoin Cash (BCH)	Eos (EOS)	Ripple (XRP)	Ethereum (ETH)
Mean	8 098.2730	1 080.7647	7.8663	0.6261	561.7860
Standard deviation	2 987.6291	559.0468	5.0447	0.4420	228.6736
Min	2 983.9000	271.9800	0.4767	0.1480	198.8000
25% quantile	6 350.1000	693.6600	2.9096	0.2476	374.0000
50% quantile	7 589.7000	993.6900	8.1487	0.5487	519.4450
75% quantile	9 243.1000	1 368.2000	11.0740	0.8021	694.7800
Max	4 099.9000	22.9340	3.2924	1 423.2000	1 423.2000
Mean	5.3749	7.7767	952.1569	8 619.2061	42.9810
Standard deviation	19.0215	40.7575	3 718.1231	22043.5426	123.9355
Min	0.0000	0.0000	0.0000	0.0000	0.0000
25% quantile	0.4349	0.5098	52.9848	640.3533	5.7394
50% quantile	1.9480	2.1196	257.9890	3 000.0000	21.0000
75% quantile	5.8900	5.7230	782.6340	7 592.0000	44.2556
Max	7 000.2500	2 115.8800	399 938.0000	2 398 030.0000	10 430.8000

TABLE E.2: Summary statistics of the first limit ask price and quantity for the five cryptocurrencies

Statistics	Bitcoin (BTC)	Bitcoin Cash (BCH)	Eos (EOS)	Ripple (XRP)	Ethereum (ETH)
Mean	8 097.1762	1 079.6603	7.8596	0.6256	561.6492
Standard deviation	2 986.7722	558.4814	5.0423	0.4416	228.5940
Min	2 981.1000	270.4600	0.4760	0.1477	198.7900
25% quantile	6 350.0000	693.3300	2.9020	0.2472	373.9500
50% quantile	7 589.2000	992.7100	8.1410	0.5484	519.3900
75% quantile	9 242.6000	1 366.9000	11.0660	0.8018	694.6000
Max	19 890.0000	4 085.1000	22.9310	3.2920	1 422.7000
Mean	5.8006	7.5464	1090.7730	9 111.7019	39.7089
Standard deviation	19.2089	34.1217	20 049.0538	29 534.6715	127.3119
Min	0.0000	0.0000	0.0000	0.0000	0.0000
25% quantile	0.5270	0.5542	52.4090	686.7653	6.8400
50% quantile	2.1702	2.4277	287.2750	3 291.8000	21.4175
75% quantile	6.2814	6.2263	817.5220	8 383.9150	42.6328
Max	2 683.0000	4 561.3100	15 111 700.00	14 713 400.00	15 027.6000

TABLE E.3: Summary statistics of the first limit bid price and quantity for the five cryptocurrencies

Appendix F

Descriptive statistics of the market quality proxies on 30 minutes intervals

Proxy	Statistics	Bitcoin (BTC)	Bitcoin Cash (BCH)	Eos (EOS)	Ripple (XRP)	Ethereum (ETH)
Realized Spread	Mean	0.0001	0.0011	0.0017	0.0012	0.0002
	Standard deviation	0.0001	0.0009	0.0015	0.0013	0.0003
	Median	0.0000	0.0010	0.0012	0.0008	0.0001
	Skewness	2.9623	-7.3392	1.5746	1.8809	2.0930
PPI	Kurtosis	15.4265	425.7802	3.0203	3.7980	6.5779
	Mean	-0.0000	-0.0000	-0.0001	-0.0001	-0.0000
	Standard deviation	0.0000	0.0002	0.0006	0.0005	0.0001
	Median	0.0000	0.0000	0.0000	0.0000	0.0000
Liquidity	Skewness	-2.4956	-4.2112	-8.6514	-6.5701	-2.6953
	Kurtosis	16.5008	54.4137	159.1822	89.1478	10.7891
	Mean	0.0000	0.0000	0.0000	0.0000	0.0000
	Standard deviation	0.0000	0.0000	0.0000	0.0000	0.0000
Kurtosis	Median	0.0000	0.0000	0.0000	0.0000	0.0000
	Skewness	-2.9751	-121.5343	-74.7312	-65.4255	-4.9917
	Kurtosis	16.6128	15 070.7389	7 060.7427	6 072.6325	86.8532

TABLE F.1: Summary statistics of the proxies for market quality on 30 minutes intervals (1)

Proxy	Statistics	Bitcoin (BTC)	Bitcoin Cash (BCH)	Eos (EOS)	Ripple (XRP)	Ethereum (ETH)
Return Autocorrelation	Mean	0.3229	0.3281	0.3327	0.3344	0.3358
	Standard deviation	0.2321	0.2359	0.2359	0.2361	0.2382
	Median	0.2750	0.2829	0.2897	0.2900	0.2909
	Skewness	0.6798	0.6343	0.6216	0.6216	0.6094
Realized Volatility	Kurtosis	-0.3766	-0.4745	-0.4892	-0.4734	-0.5079
	Mean	0.0102	0.0268	0.03288	0.0281	0.0191
	Standard deviation	0.0124	0.0399	0.0427	0.05190	0.0225
	Median	0.0070	0.0169	0.0232	0.01652	0.0127
Quote Return	Skewness	6.7751	12.4604	25.3614	36.40	6.2227
	Kurtosis	99.9886	331.7928	1 331.4871	2 602.63	98.5072
	Mean	0.0022	0.0033	0.0037	0.0031	0.0023
	Standard deviation	0.0019	0.0034	0.0033	0.0034	0.0021
Kurtosis	Median	0.0016	0.0024	0.0028	0.0021	0.0017
	Skewness	3.2831	5.6278	4.5603	5.0691	4.5259
	Kurtosis	20.7114	76.1167	59.5088	1.5535	59.5262

TABLE F.2: Summary statistics of the proxies for market quality on 30 minutes intervals (2)

Appendix G

Robustness check : non-standardized proxies

Measure	Components				
	1	2	3	4	5
Realized Spread	76.555	13.303	8.873	1.144	0.125
Percentage Price Impact	50.043	40.396	9.123	0.343	0.094
Realized Volatility	56.554	20.922	16.305	5.160	1.059
New Amihud	82.957	12.625	4.417	0.001	0.000
Return Autocorrelation	30.881	18.453	18.117	17.647	14.902
Quote Return	61.869	16.547	13.446	5.744	2.394

TABLE G.1: Explanatory power (in %) of the five first components of the PCA for the whole period and with non-standardized proxies on 30 minutes intervals

Period	Measure	Components				
		1	2	3	4	5
Beginning	Realized Spread	54.973	27.726	16.441	0.695	0.165
	Percentage Price Impact	52.609	41.511	5.715	0.140	0.025
	Realized Volatility	48.258	30.297	13.530	6.833	1.083
	New Amihud	83.068	12.445	4.487	0.000	0.000
	Return Autocorrelation	24.992	20.229	19.649	19.185	15.946
	Quote Return	47.106	23.516	18.753	8.385	2.240
End	Realized Spread	51.799	26.288	21.851	0.056	0.006
	Percentage Price Impact	61.527	24.865	12.632	0.840	0.136
	Realized Volatility	80.169	12.385	4.075	3.280	0.091
	New Amihud	70.323	26.482	2.144	0.947	0.104
	Return Autocorrelation	33.884	17.323	17.266	16.736	14.790
	Quote Return	73.337	12.929	6.736	4.646	2.353

TABLE G.2: Explanatory power (in %) of the five first components of the PCA for the three first and three last months of the period and with non-standardized proxies on 30 minutes intervals

Period	Measure	Components				
		1	2	3	4	5
Before futures	Realized Spread	61.595	25.923	10.818	1.430	0.233
	Percentage Price Impact	60.699	34.588	4.496	0.177	0.039
	Realized Volatility	61.400	25.520	7.511	4.017	1.552
	New Amihud	91.173	4.924	3.902	0.000	0.000
	Return Autocorrelation	25.808	20.022	19.468	19.000	15.702
	Quote Return	57.354	23.058	11.253	6.127	2.208
After futures	Realized Spread	60.395	18.642	15.779	4.710	0.474
	Percentage Price Impact	42.542	26.847	22.603	5.823	2.185
	Realized Volatility	77.038	12.271	5.951	3.974	0.766
	New Amihud	59.171	33.396	6.892	0.520	0.021
	Return Autocorrelation	32.323	18.964	17.434	17.039	14.240
	Quote Return	72.209	12.323	8.813	4.492	2.163

TABLE G.3: Explanatory power (in %) of the five first components of the PCA for the three months before and after the introduction of futures on Bitcoin and with non-standardized proxies on 30 minutes intervals

Period	Measure	Components				
		1	2	3	4	5
Before BTC bubble	Realized Spread	63.926	19.437	11.572	4.678	0.387
	Percentage Price Impact	37.116	26.157	19.911	13.796	3.020
	Realized Volatility	73.393	14.323	8.257	3.426	0.601
	New Amihud	91.643	7.051	0.923	0.361	0.023
	Return Autocorrelation	29.902	20.231	18.843	16.958	14.065
	Quote Return	69.281	17.190	7.116	4.282	2.131
After BTC bubble	Realized Spread	56.396	23.780	13.516	5.706	0.603
	Percentage Price Impact	68.895	17.395	13.278	0.402	0.031
	Realized Volatility	70.663	13.725	8.926	4.897	1.789
	New Amihud	62.860	32.905	3.878	0.333	0.024
	Return Autocorrelation	38.165	17.194	16.132	15.004	13.506
	Quote Return	72.601	9.847	7.998	6.934	2.620

TABLE G.4: Explanatory power (in %) of the five first components of the PCA for the month preceding and following the end of BTC bubble and with non-standardized proxies on 30 minutes intervals

Appendix H

Robustness check : mean as aggregation rule

Measure	Components				
	1	2	3	4	5
Realized Spread	69.590	13.955	6.798	5.082	4.575
Percentage Price Impact	57.532	14.729	12.128	10.878	4.733
Realized Volatility	59.019	13.839	11.197	9.075	6.870
New Amihud	39.344	18.054	16.285	14.611	11.704
Return Autocorrelation	32.583	18.004	17.823	16.882	14.707
Quote Return	65.286	12.151	10.288	7.877	4.398

TABLE H.1: Explanatory power (in %) of the five first components of the PCA for the whole period, with standardized proxies on 30 minutes intervals and with the mean as aggregation rule

Period	Measure	Components				
		1	2	3	4	5
Beginning	Realized Spread	53.309	16.240	13.402	11.041	6.008
	Percentage Price Impact	34.379	29.858	19.966	11.179	4.618
	Realized Volatility	46.029	22.908	15.223	9.342	6.498
	New Amihud	32.342	25.316	18.501	14.373	9.468
	Return Autocorrelation	26.106	20.036	19.560	18.618	15.681
	Quote Return	49.139	19.074	15.627	11.758	4.401
End	Realized Spread	60.956	14.224	10.762	8.315	5.744
	Percentage Price Impact	45.886	20.279	14.053	11.931	7.852
	Realized Volatility	55.705	22.163	11.264	8.276	2.593
	New Amihud	62.619	24.528	11.756	0.808	0.289
	Return Autocorrelation	35.674	17.343	16.519	15.754	14.711
	Quote Return	73.641	11.204	6.793	4.658	3.705

TABLE H.2: Explanatory power (in %) of the five first components of the PCA for the three first and three last months of the period, with non standardized proxies on 30 minutes intervals and with the mean as aggregation rule

Period	Measure	Components				
		1	2	3	4	5
Before futures	Realized Spread	48.898	21.692	11.828	9.996	7.587
	Percentage Price Impact	29.653	20.585	19.080	17.797	12.885
	Realized Volatility	52.986	19.465	14.275	8.693	4.581
	New Amihud	32.656	23.474	17.693	15.507	10.670
	Return Autocorrelation	22.491	20.422	20.058	19.335	17.695
	Quote Return	57.421	18.173	13.647	5.768	4.990
After futures	Realized Spread	73.354	11.869	7.161	4.699	2.916
	Percentage Price Impact	57.626	34.187	5.366	1.766	1.055
	Realized Volatility	63.666	12.980	10.757	7.773	4.824
	New Amihud	65.345	22.018	7.803	2.834	2.000
	Return Autocorrelation	26.077	19.796	18.870	18.010	17.247
	Quote Return	67.234	12.777	10.222	5.764	4.003

TABLE H.3: Explanatory power (in %) of the five first components of the PCA for the three months before and after the introduction of futures on Bitcoin, with standardized proxies on 30 minutes intervals and with the mean as aggregation rule

Period	Measure	Components				
		1	2	3	4	5
Before BTC bubble	Realized Spread	64.082	15.999	10.593	6.185	3.142
	Percentage Price Impact	49.761	17.089	14.755	12.533	5.861
	Realized Volatility	70.404	12.612	9.851	4.621	2.512
	New Amihud	66.837	21.722	7.353	2.993	1.096
	Return Autocorrelation	31.872	20.018	18.381	16.109	13.620
	Quote Return	72.095	12.421	7.334	4.883	3.267
After BTC bubble	Realized Spread	75.353	12.479	8.866	2.075	1.227
	Percentage Price Impact	51.860	25.572	10.209	8.875	3.484
	Realized Volatility	74.617	9.199	7.341	5.978	2.864
	New Amihud	63.583	21.345	13.622	0.743	0.707
	Return Autocorrelation	40.899	16.686	15.588	14.193	12.634
	Quote Return	79.050	7.060	5.480	4.757	3.652

TABLE H.4: Explanatory power (in %) of the five first components of the PCA for the month preceding and following the end of BTC bubble, with standardized proxies on 30 minutes intervals and mean as aggregation rule

Appendix I

Weight variables - Statistics

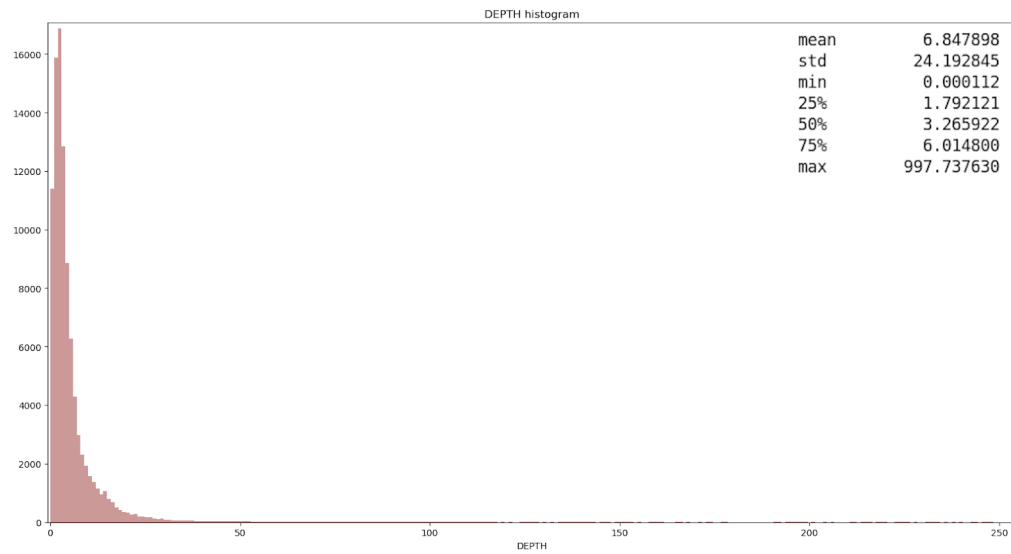


FIGURE I.1: BCH 's DEPTH - main statistics

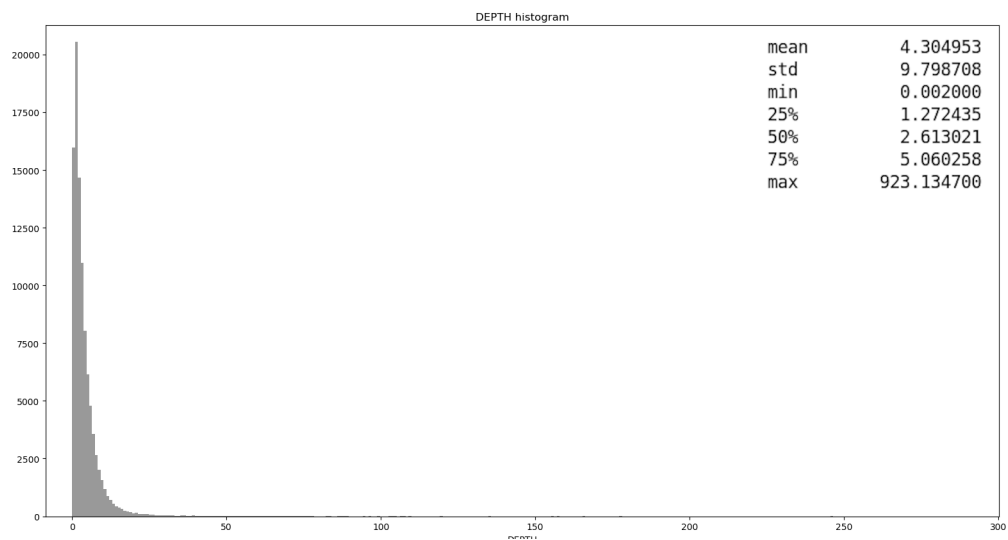


FIGURE I.2: BTC 's DEPTH - main statistics

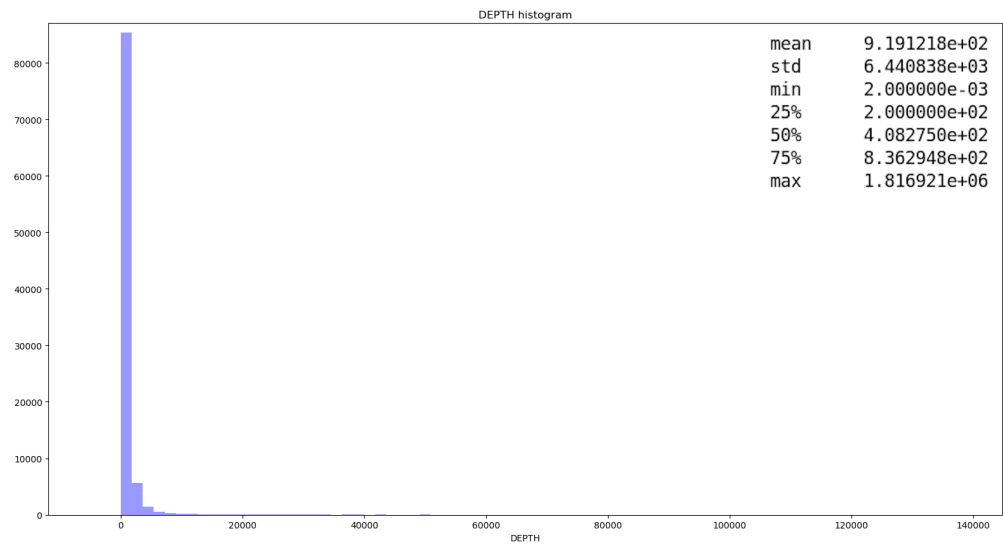


FIGURE I.3: EOS ' DEPTH - main statistics

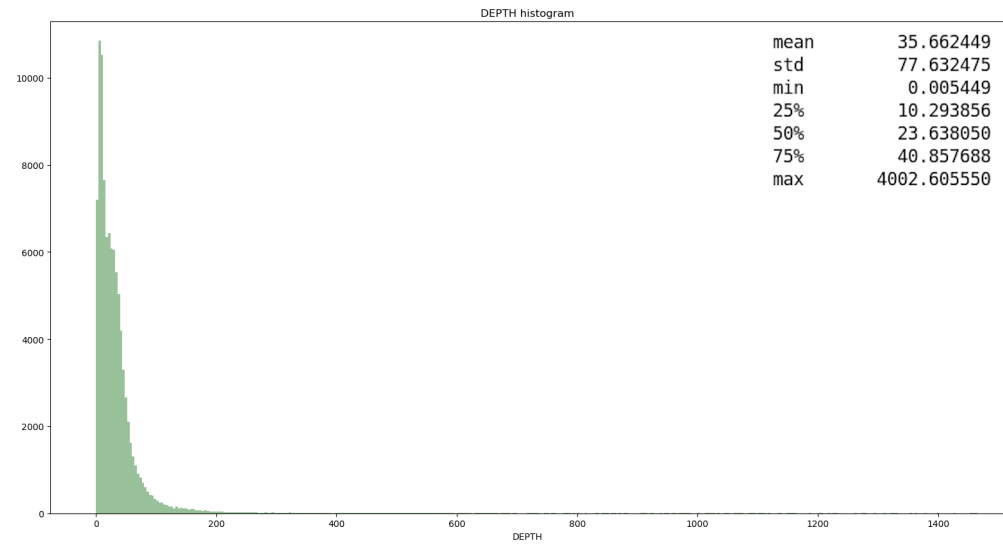


FIGURE I.4: ETH 's DEPTH - main statistics

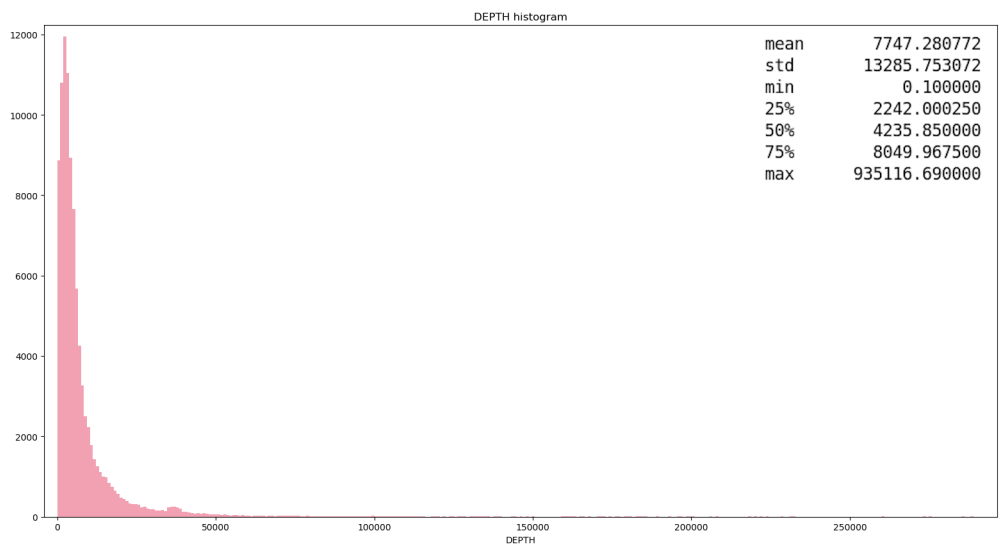


FIGURE I.5: XRP 's DEPTH - main statistics

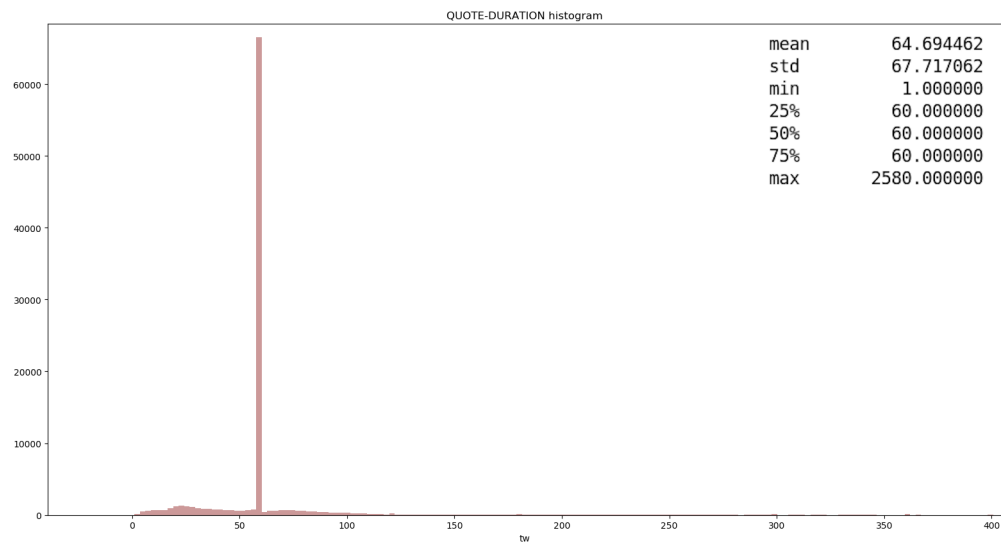


FIGURE I.6: BCH 's quote-duration - main statistics

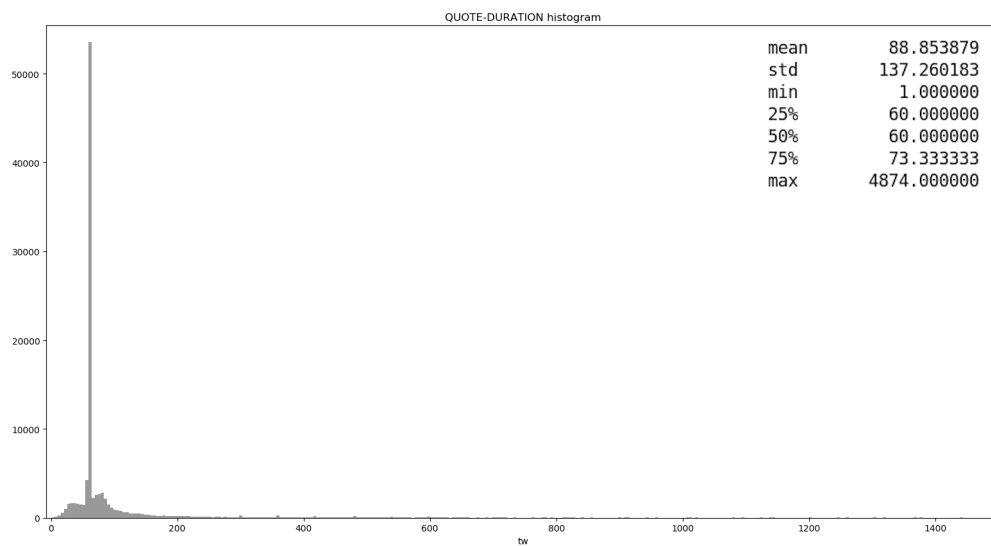


FIGURE I.7: BTC 's quote-duration - main statistics

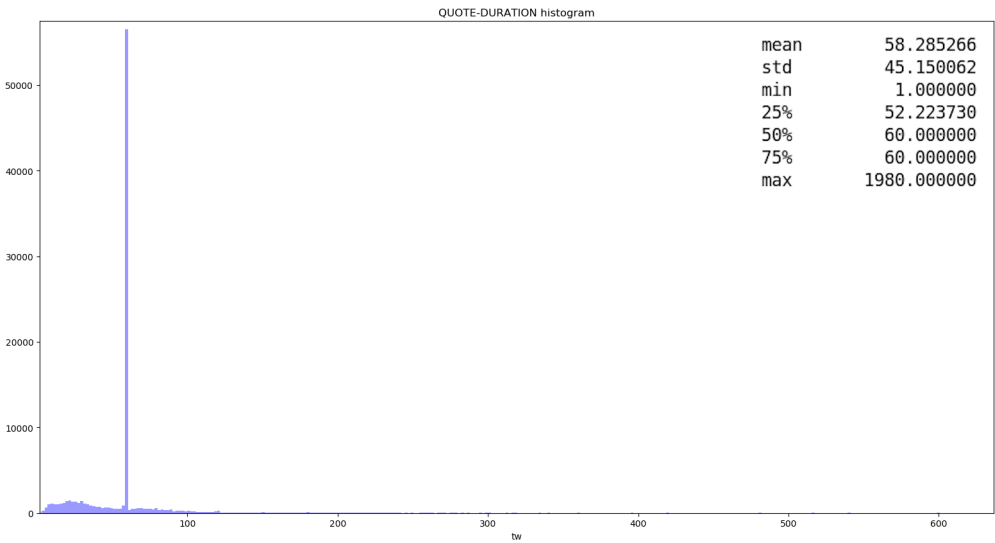


FIGURE I.8: EOS ' quote-duration - main statistics

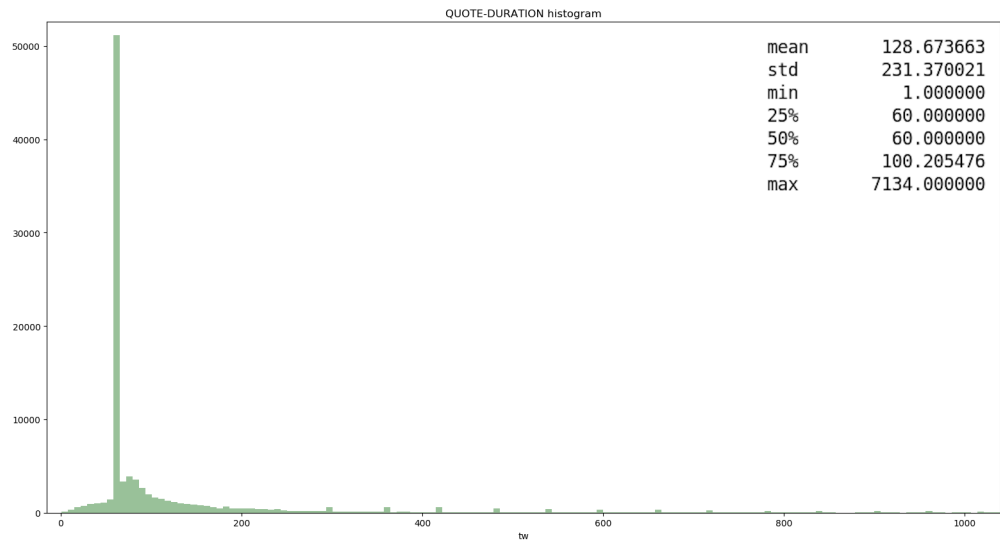


FIGURE I.9: ETH 's quote-duration - main statistics

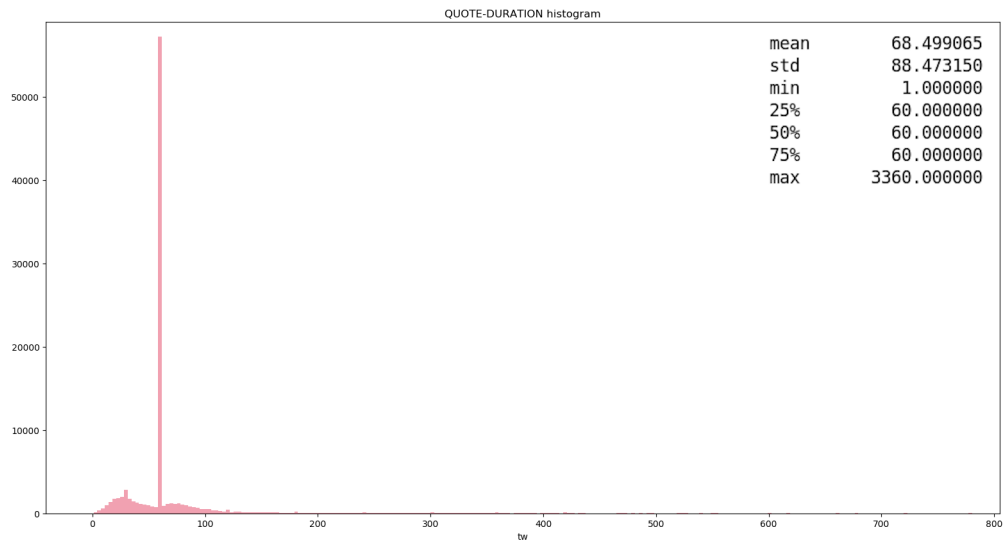


FIGURE I.10: XRP 's quote-duration - main statistics

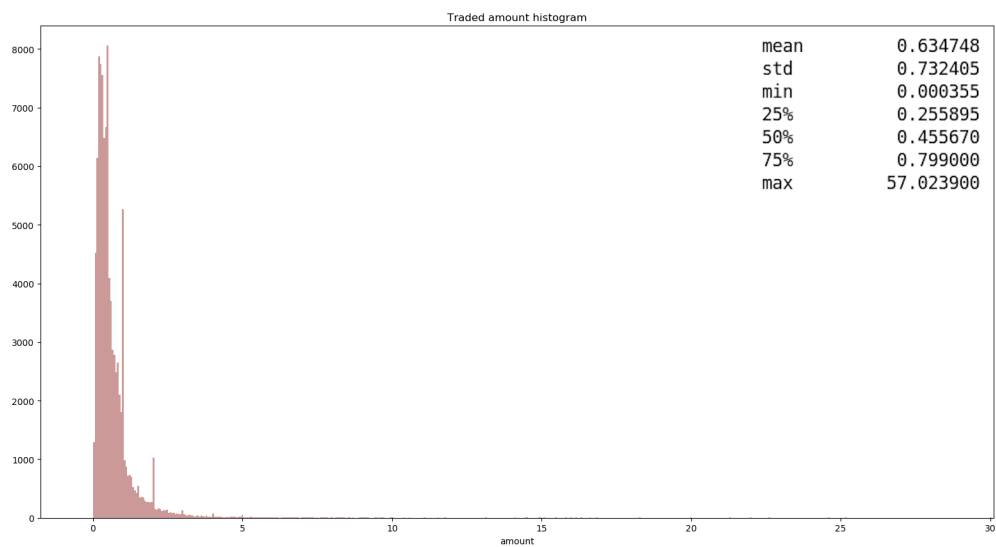


FIGURE I.11: BCH 's traded amount - main statistics

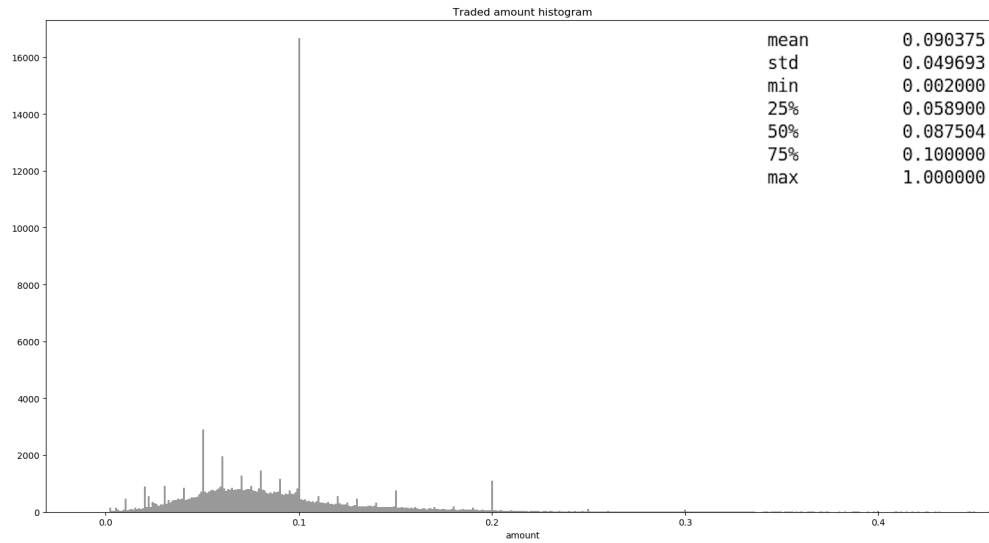


FIGURE I.12: BTC 's traded amount - main statistics

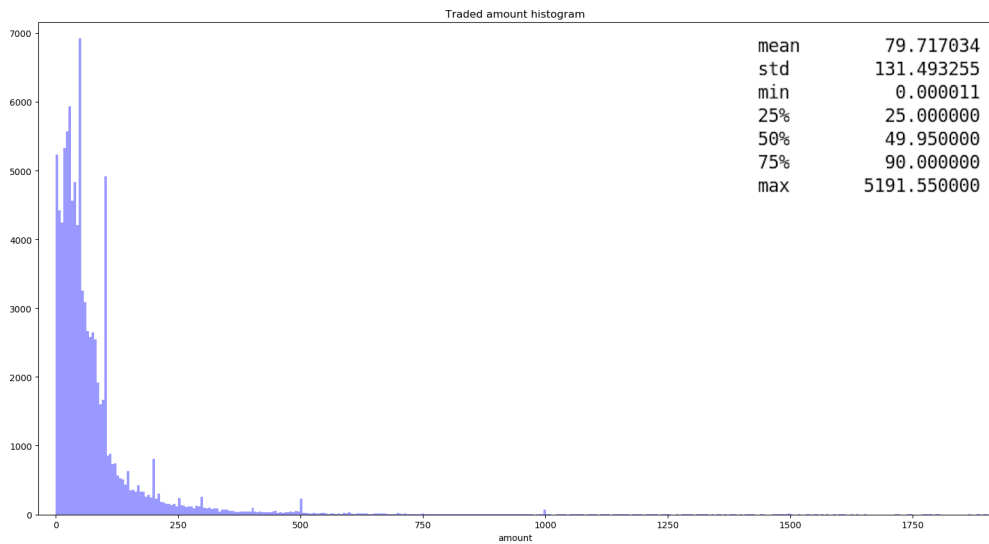


FIGURE I.13: EOS ' traded amount - main statistics

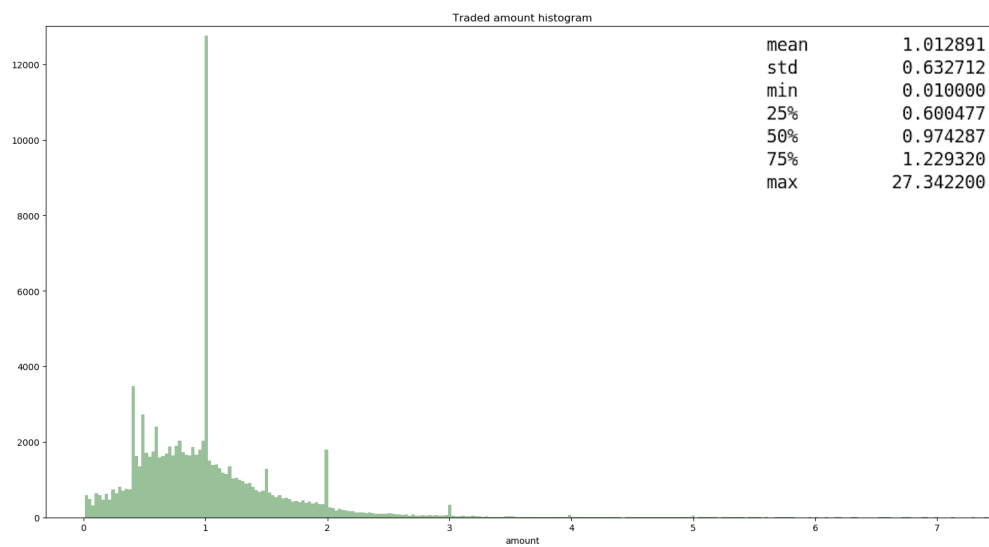


FIGURE I.14: ETH 's traded amount - main statistics

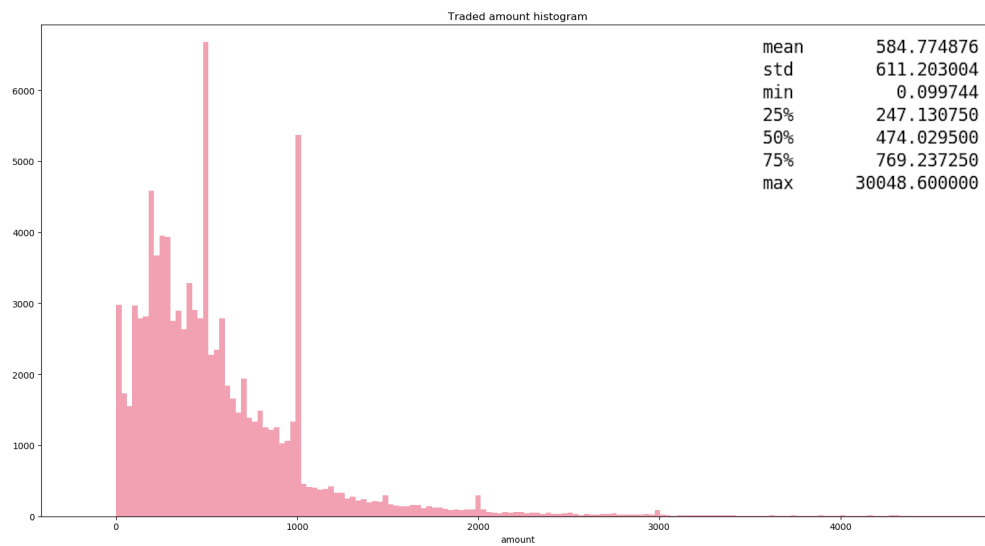


FIGURE I.15: XRP 's traded amount - main statistics

Appendix J

Evolution of daily and weekly Adj. R-squared by periods

TABLE J.1: Comparison of the average daily adjusted R^2 by periods (1)

	Pre-bubble	Post-bubble	Pre-futures	Post-futures	Start of series	End of series
BCH						
EWPOQs	0.0302	0.0017	0.0477	0.0353	0.0294	0.0433
SWPOQs	0.0223	0.0040	0.0401	0.0298	0.0221	0.0367
TWPQs	0.0239	0.0009	0.0481	0.0326	0.0284	0.0457
SWPES	0.1299	0.0960	0.1175	0.1365	0.0814	0.1423
SWPTS	0.0702	0.0402	0.0834	0.0666	0.0538	0.0887
EWDEPTH	0.0067	0.0571	0.0231	0.0280	0.0224	0.0048
BTC						
EWPOQs	0.0474	-0.0232	0.0737	0.0484	0.0686	0.0375
SWPOQs	0.0362	0.0464	0.0645	0.0615	0.0568	0.0261
TWPQs	0.0446	-0.0215	0.0769	0.0487	0.0714	0.0328
SWPES	0.2049	0.1869	0.2464	0.2319	0.2232	0.2364
SWPTS	0.1503	0.1175	0.1905	0.1482	0.1581	0.1317
EWDEPTH	-0.0191	0.0070	0.0172	0.0038	0.0222	0.0168
EOS						
EWPOQs	0.0484	0.0633	0.0185	0.0489	0.0214	0.0463
SWPOQs	0.0207	0.0541	0.0074	0.0360	0.0142	0.0400
TWPQs	0.0530	0.0668	0.0189	0.0522	0.0196	0.0481
SWPES	0.1541	0.0927	0.0853	0.1348	0.0574	0.2232
SWPTS	0.1128	0.0416	0.0445	0.0829	0.0236	0.1121
EWDEPTH	0.0012	0.0275	0.0048	0.0113	0.0114	0.0076

Pre-bubble and post-bubble periods refer to one month prior and one month after the end of Bitcoin bubble the 28th of January 2018. Pre-futures and post-futures refer to three months prior and three months after futures introduction on Bitcoin the 7th of December 2017. The start and end of the time-series refer to the first and last three months of our time frame, that is from August 2017 to November 2017 and from April 2018 to July 2018. The adjusted- R^2 comes from the regression in equation 3.8 computed on a daily basis.

TABLE J.2: Comparison of the average daily adjusted R^2 by periods (2)

	Pre-bubble	Post-bubble	Pre-futures	Post-futures	Start of series	End of series
ETH						
EWPOQs	0.0759	0.0595	0.0518	0.0727	0.0409	0.0755
SWPOQs	0.0700	0.0279	0.0334	0.0509	0.0295	0.0722
TWPOQs	0.0796	0.0696	0.0520	0.0777	0.0420	0.0827
SWPES	0.1651	0.1617	0.1429	0.1832	0.1217	0.1809
SWPTS	0.1214	0.0594	0.0782	0.0997	0.0707	0.0819
EWDEPTH	-0.0046	-0.0175	0.0204	-0.00183	0.0066	<0.0001
XRP						
EWPOQs	0.0356	-0.0146	0.0291	0.0305	0.0270	0.0318
SWPOQs	0.0091	-0.0328	0.0231	0.0133	0.023	0.0303
TWPOQs	0.0358	-0.0182	0.0299	0.0305	0.0272	0.0392
SWPES	0.2128	0.1198	0.0900	0.2009	0.0666	0.1624
SWPTS	0.1736	0.0552	0.0435	0.1333	0.0253	0.0883
EWDEPTH	-0.0113	0.0257	0.0025	0.0085	0.0041	0.0108

Pre-bubble and post-bubble periods refer to one month prior and one month after the end of Bitcoin bubble the 28th of January 2018. Pre-futures and post-futures refer to three months prior and three months after futures introduction on Bitcoin the 7th of December 2017. The start and end of the time-series refer to the first and last three months of our time frame, that is from August 2017 to November 2017 and from April 2018 to July 2018. The adjusted- R^2 comes from the regression in equation 3.8 computed on a daily basis.

TABLE J.3: Comparison of the average weekly adjusted R^2 by periods (1)

	Pre-bubble	Post-bubble	Pre-futures	Post-futures	Start of series	End of series
BCH						
EWPOQs	0.1791	0.0411	0.0925	0.1187	0.0339	0.0639
SWPOQs	0.1435	0.0300	0.0773	0.0898	0.0267	0.0521
TWPQs	0.1787	0.0394	0.0924	0.1191	0.0333	0.0663
SWPES	0.2376	0.0581	0.1873	0.1641	0.1132	0.1479
SWPTS	0.1608	0.0322	0.1137	0.1064	0.0504	0.0895
EWDEPTH	0.0077	-0.0052	0.0080	0.0004	0.0099	0.0039
BTC						
EWPOQs	0.1654	0.0538	0.1089	0.1140	0.0599	0.0192
SWPOQs	0.1488	0.0380	0.0948	0.096	0.0520	0.0130
TWPQs	0.1659	0.0552	0.1111	0.115	0.0609	0.0178
SWPES	0.2952	0.0913	0.2773	0.2227	0.2447	0.1749
SWPTS	0.2486	0.0445	0.2031	0.1651	0.1531	0.0793
EWDEPTH	0.0029	0.0001	0.0071	0.003	0.0071	0.0079
EOS						
EWPOQs	0.1833	0.0329	0.0387	0.1066	0.0452	0.0442
SWPOQs	0.1509	0.0201	0.0262	0.0784	0.034	0.0370
TWPQs	0.1803	0.0339	0.038	0.106	0.0433	0.0484
SWPES	0.2774	0.0405	0.0871	0.1573	0.0500	0.210
SWPTS	0.1548	0.0209	0.0458	0.0967	0.0207	0.1240
EWDEPTH	-0.0005	0.0220	0.0029	0.0090	0.0035	0.0058

Pre-bubble and post-bubble periods refer to one month prior and one month after the end of Bitcoin bubble the 28th of January 2018. Pre-futures and post-futures refer to three months prior and three months after futures introduction on Bitcoin the 7th of December 2017. The start and end of the time-series refer to the first and last three months of our time frame, that is from August 2017 to November 2017 and from April 2018 to July 2018. The adjusted- R^2 comes from the regression in equation 3.8 computed on a weekly basis.

TABLE J.4: Comparison of the average weekly adjusted R^2 by periods (2)

	Pre-bubble	Post-bubble	Pre-futures	Post-futures	Start of series	End of series
ETH						
EWPOQs	0.1403	0.0663	0.1024	0.1152	0.0615	0.0766
SWPOQs	0.1193	0.0523	0.0735	0.0918	0.0448	0.0701
TWPOQs	0.1410	0.0688	0.1039	0.1178	0.0627	0.0889
SWPES	0.2378	0.1160	0.1790	0.2150	0.1356	0.1534
SWPTS	0.1841	0.0618	0.1288	0.1334	0.0846	0.0667
EWDEPTH	0.0023	-0.0025	0.0056	0.0012	0.0046	0.0078
XRP						
EWPOQs	0.2293	0.0273	0.0712	0.1318	0.0607	0.0516
SWPOQs	0.1939	0.0190	0.0482	0.1074	0.044	0.0427
TWPOQs	0.2293	0.0282	0.0710	0.132	0.0589	0.0554
SWPES	0.4047	0.1247	0.1262	0.2923	0.0866	0.1680
SWPTS	0.3513	0.1142	0.0651	0.2336	0.0293	0.0924
EWDEPTH	0.0037	0.0007	0.0039	0.0027	0.0016	0.0062

Pre-bubble and post-bubble periods refer to one month prior and one month after the end of Bitcoin bubble the 28th of January 2018. Pre-futures and post-futures refer to three months prior and three months after futures introduction on Bitcoin the 7th of December 2017. The start and end of the time-series refer to the first and last three months of our time frame, that is from August 2017 to November 2017 and from April 2018 to July 2018. The adjusted- R^2 comes from the regression in equation 3.8 computed on a weekly basis.

Appendix K

Results of the statistical tests of means

TABLE K.1: Results of the tests of means between the average daily adjusted- R^2 by periods

	Bubble	Futures	Start-end
$H_0 :$	$\mu_5 < \mu_6$	$\mu_3 < \mu_4$	$\mu_1 < \mu_2$
$H_1 :$	$\mu_5 > \mu_6$	$\mu_3 > \mu_4$	$\mu_1 > \mu_2$
BCH			
EWPQS	0.0343	0.1801	0.9264
SWPQS	0.0861	0.2028	0.9468
TWPQS	0.0620	0.1273	0.9595
SWPES	0.2059	0.7844	0.9997
SWPTS	0.1173	0.1719	0.9961
EWDEPTH	0.9088	0.6302	0.0574
BTC			
EWPQS	0.1773	0.1819	0.0355
SWPQS	0.6350	0.4267	0.0188
TWPQS	0.1906	0.1550	0.0101
SWPES	0.3201	0.2757	0.7233
SWPTS	0.2337	0.0328	0.0855
EWDEPTH	0.8009	0.1702	0.320
EOS			
EWPQS	0.6487	0.9773	0.9938
SWPQS	0.8113	0.9774	0.995
TWPQS	0.6337	0.9843	0.9977
SWPES	0.0409	0.9958	0.9999
SWPTS	0.0103	0.9955	0.9999
EWDEPTH	0.7321	0.6678	0.3541
ETH			
EWPQS	0.3006	0.9151	0.9882
SWPQS	0.1721	0.8509	0.9973
TWPQS	0.3804	0.9530	0.994
SWPES	0.4704	0.9605	0.9982
SWPTS	0.0013	0.9210	0.8044
EWDEPTH	0.3697	0.0799	0.209
XRP			
EWPQS	0.1739	0.5273	0.6868
SWPQS	0.2683	0.3406	0.7624
TWPQS	0.1690	0.5101	0.8842
SWPES	0.0425	0.9999	0.9999
SWPTS	0.0016	0.9999	0.9999
EWDEPTH	0.8896	0.7045	0.7691

μ_i stands for the mean value of the adjusted R^2 resulting from regression 3.8 over time-period i . First, we compare the adj. R^2 one month prior (μ_5) and one month after (μ_6) the end of Bitcoin bubble the 28th of January 2018. Second, we compare three months prior (μ_3) and three months after (μ_4) the introduction of Futures on Bitcoin. Third, we compare the first three months of our time-series (μ_1) (August to November 2017) to the three last months of our time-series (μ_2) (April to July 2018).

TABLE K.2: Results of the tests of means between the average weekly adjusted- R^2 by periods

	bubble	futures	start-end
$H_0 :$	$\mu_5 < \mu_6$	$\mu_3 < \mu_4$	$\mu_1 < \mu_2$
$H_1 :$	$\mu_5 > \mu_6$	$\mu_3 > \mu_4$	$\mu_1 > \mu_2$
BCH			
EWPQS	0.0582	0.7048	0.9511
SWPQS	0.0773	0.6166	0.9755
TWPQS	0.0579	0.7074	0.9609
SWPES	0.0080	0.3412	0.8539
SWPTS	0.0204	0.4391	0.9760
EWDEPTH	0.1378	0.1039	0.2307
BTC			
EWPQS	0.1259	0.5407	0.0273
SWPQS	0.1048	0.5133	0.0160
TWPQS	0.1283	0.5302	0.0234
SWPES	0.0219	0.1723	0.0344
SWPTS	0.0039	0.1997	0.0022
EWDEPTH	0.0747	0.1713	0.5590
EOS			
EWPQS	0.0412	0.9603	0.4808
SWPQS	0.0410	0.9440	0.5521
TWPQS	0.0451	0.9605	0.6063
SWPES	0.0064	0.9232	0.9999
SWPTS	0.0191	0.9364	0.9999
EWDEPTH	0.8087	0.7305	0.6941
ETH			
EWPQS	0.1057	0.6159	0.7204
SWPQS	0.0937	0.7059	0.8933
TWPQS	0.1092	0.6262	0.8126
SWPES	0.0345	0.7359	0.6764
SWPTS	0.0051	0.5399	0.2532
EWDEPTH	0.2197	0.1439	0.7417
XRP			
EWPQS	0.0473	0.8686	0.3541
SWPQS	0.0539	0.8953	0.4568
TWPQS	0.0481	0.8726	0.4418
SWPES	0.0062	0.9967	0.9998
SWPTS	0.0126	0.9984	0.9999
EWDEPTH	0.3776	0.3967	0.7933

μ_i stands for the mean value of the adjusted R^2 resulting from regression 3.8 over time-period i . First, we compare the adj. R^2 one month prior (μ_5) and one month after (μ_6) the end of Bitcoin bubble the 28th of January 2018. Second, we compare three months prior (μ_3) and three months after (μ_4) the introduction of Futures on Bitcoin. Third, we compare the first three months of our time-series (μ_1) (August to November 2017) to the three last months of our time-series (μ_2) (April to July 2018)

Bibliography

- Acharya, Viral V and Lasse Heje Pedersen (2005). "Asset pricing with liquidity risk". In: *Journal of financial Economics* 77.2, pp. 375–410.
- Adjaoute, Kpate and Jean-Pierre Danthine (2001). "EMU and portfolio diversification opportunities". In:
- Admati, Anat R and Paul Pfleiderer (1988). "A theory of intraday patterns: Volume and price variability". In: *The Review of Financial Studies* 1.1, pp. 3–40.
- Aggarwal, Priyanka et al. (2011). "Map matching and heuristic elimination of gyro drift for personal navigation systems in GPS-denied conditions". In: *Measurement Science and Technology* 22.2, p. 025205.
- Amihud, Yakov (2002). "Illiquidity and stock returns: cross-section and time-series effects". In: *Journal of financial markets* 5.1, pp. 31–56.
- Amihud, Yakov and Haim Mendelson (1986). "Asset pricing and the bid-ask spread". In: *Journal of financial Economics* 17.2, pp. 223–249.
- Anagnostidis, Panagiotis and Patrice C Fontaine (2018). "Liquidity provision, commonality and high frequency trading". In:
- Androulaki, Elli et al. (2013). "Evaluating user privacy in bitcoin". In: *International Conference on Financial Cryptography and Data Security*. Springer, pp. 34–51.
- Bariviera, Aurelio F (2017). "The inefficiency of Bitcoin revisited: A dynamic approach". In: *Economics Letters* 161, pp. 1–4.
- Bartels, Robert (1982). "The rank version of von Neumann's ratio test for randomness". In: *Journal of the American Statistical Association* 77.377, pp. 40–46.
- Baur, Dirk G, Thomas Dimpfl, and Konstantin Kuck (2018). "Bitcoin, gold and the US dollar—A replication and extension". In: *Finance Research Letters* 25, pp. 103–110.
- Beaupain, Renaud, Pierre Giot, and Mikael Petitjean (2010). "Volatility regimes and liquidity co-movements in cap-based portfolios". In: *Finance* 31.1, pp. 55–79.
- Benos, Evangelos et al. (2017). "Interactions among high-frequency traders". In: *Journal of Financial and Quantitative Analysis* 52.4, pp. 1375–1402.
- Biryukov, Alex and Ivan Pustogarov (2015). "Bitcoin over Tor isn't a Good Idea". In: *2015 IEEE Symposium on Security and Privacy*. IEEE, pp. 122–134.
- Bouri, Elie, Georges Azzi, and Anne Haubo Dyhrberg (2016). "On the return-volatility relationship in the Bitcoin market around the price crash of 2013". In: *Available at SSRN* 2869855.
- Breckenfelder, Johannes (2019). "Competition among high-frequency traders, and market quality". In:
- Brockman, Paul and Dennis Y Chung (2002). "Commonality in liquidity: Evidence from an order-driven market structure". In: *Journal of Financial Research* 25.4, pp. 521–539.
- Brockman, Paul, Dennis Y Chung, and Christophe Pérignon (2009). "Commonality in liquidity: A global perspective". In: *Journal of Financial and Quantitative Analysis* 44.4, pp. 851–882.

- Brogaard, Jonathan et al. (2010). "High frequency trading and its impact on market quality". In: *Northwestern University Kellogg School of Management Working Paper* 66.
- Brunnermeier, Markus K and Lasse H Pedersen (2009). "Funding liquidity and market liquidity". In: *Review of Financial Studies* 22.2201-2238, p. 6.
- Cheah, Eng-Tuck et al. (2018). "Long memory interdependency and inefficiency in Bitcoin markets". In: *Economics Letters* 167, pp. 18–25.
- Cheung, Adrian, Eduardo Roca, and Jen-Je Su (2015). "Crypto-currency bubbles: an application of the Phillips–Shi–Yu (2013) methodology on Mt. Gox bitcoin prices". In: *Applied Economics* 47.23, pp. 2348–2358.
- Chiu, Jonathan and Thorsten V Koeppl (2016). "Trading dynamics with adverse selection and search: Market freeze, intervention and recovery". In: *The Review of Economic Studies* 83.3, pp. 969–1000.
- Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam (2000). "Commonality in liquidity". In: *Journal of financial economics* 56.1, pp. 3–28.
- Christoffersen, Peter, Asger Lunde, and Kasper V Olesen (2019). "Factor structure in commodity futures return and volatility". In: *Journal of Financial and Quantitative Analysis* 54.3, pp. 1083–1115.
- Clements, Ryan (2018). "Assessing the Evolution of Cryptocurrency: Demand Factors, Latent Value, and Regulatory Developments". In: *Mich. Bus. & Entrepreneurial L. Rev.* 8, p. 73.
- Corbet, Shaen, Brian Lucey, and Larisa Yarovaya (2018). "Datestamping the Bitcoin and Ethereum bubbles". In: *Finance Research Letters* 26, pp. 81–88.
- Corbet, Shaen et al. (2019). "Cryptocurrencies as a financial asset: A systematic analysis". In: *International Review of Financial Analysis* 62, pp. 182–199.
- Corwin, Shane A and Marc L Lipson (2011). "Order characteristics and the sources of commonality in prices and liquidity". In: *Journal of Financial Markets* 14.1, pp. 47–81.
- Coughenour, Jay F and Mohsen M Saad (2004). "Common market makers and commonality in liquidity". In: *Journal of Financial economics* 73.1, pp. 37–69.
- Daniel, Lawrence, Suk-Joong Kim, and Michael D McKenzie (2014). "The efficiency of the information processing in the Australian dollar market: Price discovery following scheduled and unscheduled news". In: *International Review of Financial Analysis* 32, pp. 159–178.
- Dick-Nielsen, Jens and Marco Rossi (2018). "The cost of immediacy for corporate bonds". In: *The Review of Financial Studies* 32.1, pp. 1–41.
- Dodd, Olga and Aaron Gilbert (2016). "The Impact of Cross-Listing on the Home Market's Information Environment and Stock Price Efficiency". In: *Financial Review* 51.3, pp. 299–328.
- Domowitz, Ian, Oliver Hansch, and Xiaoxin Wang (2005). "Liquidity commonality and return co-movement". In: *Journal of Financial Markets* 8.4, pp. 351–376.
- Ellis, Katrina, Roni Michaely, and Maureen O'Hara (2000). "The accuracy of trade classification rules: Evidence from Nasdaq". In: *Journal of Financial and Quantitative Analysis* 35.4, pp. 529–551.
- Engle, Robert (2002). "Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models". In: *Journal of Business & Economic Statistics* 20.3, pp. 339–350.
- EY (2017). *EY research: initial coin offerings (ICOs)*. Tech. rep.
- Fernando, Chitru S (2003). "Commonality in liquidity: Transmission of liquidity shocks across investors and securities". In: *Journal of Financial Intermediation* 12.3, pp. 233–254.

- Fishman, Michael J and Kathleen M Hagerty (1989). "Disclosure decisions by firms and the competition for price efficiency". In: *The Journal of Finance* 44.3, pp. 633–646.
- Gibson, Rajna and Nicolas Mougeot (2004). "The pricing of systematic liquidity risk: Empirical evidence from the US stock market". In: *Journal of banking & finance* 28.1, pp. 157–178.
- Gonzalo, Jesus and Clive Granger (1995). "Estimation of common long-memory components in cointegrated systems". In: *Journal of Business & Economic Statistics* 13.1, pp. 27–35.
- Gürkaynak, Refet S (2008). "Econometric tests of asset price bubbles: taking stock". In: *Journal of Economic surveys* 22.1, pp. 166–186.
- Hameed, Allaudeen, Wenjin Kang, and Shivesh Viswanathan (2010). "Stock market declines and liquidity". In: *The Journal of Finance* 65.1, pp. 257–293.
- Hansen, Peter R and Asger Lunde (2005). "A forecast comparison of volatility models: does anything beat a GARCH (1, 1)?" In: *Journal of applied econometrics* 20.7, pp. 873–889.
- Hasbrouck, Joel (1995). "One security, many markets: Determining the contributions to price discovery". In: *The journal of Finance* 50.4, pp. 1175–1199.
- (2002). "Stalking the "efficient price" in market microstructure specifications: an overview". In: *Journal of Financial Markets* 5.3, pp. 329–339.
- Hasbrouck, Joel and Duane J Seppi (2001). "Common factors in prices, order flows, and liquidity". In: *Journal of financial Economics* 59.3, pp. 383–411.
- Hendershott, Terrence and Pamela C Moulton (2011). "Automation, speed, and stock market quality: The NYSE's hybrid". In: *Journal of Financial Markets* 14.4, pp. 568–604.
- Heston, Steven L and K Geert Rouwenhorst (1994). "Does industrial structure explain the benefits of international diversification?" In: *Journal of Financial Economics* 36.1, pp. 3–27.
- (1995). "Industry and country effects in international stock returns". In: *Journal of Portfolio Management* 21, pp. 53–53.
- Hotchkiss, Edith S and Tavy Ronen (2002). "The informational efficiency of the corporate bond market: An intraday analysis". In: *The Review of Financial Studies* 15.5, pp. 1325–1354.
- Huberman, Gur and Dominika Halka (2001). "Systematic liquidity". In: *Journal of Financial Research* 24.2, pp. 161–178.
- Jacoby, Gady, David J Fowler, and Aron A Gottesman (2000). "The capital asset pricing model and the liquidity effect: A theoretical approach". In: *Journal of Financial Markets* 3.1, pp. 69–81.
- Jain, Pankaj K, Pawan Jain, and Thomas H McInish (2016). "Does high-frequency trading increase systemic risk?" In: *Journal of Financial Markets* 31, pp. 1–24.
- Ji, Qiang et al. (2019). "Information interdependence among energy, cryptocurrency and major commodity markets". In: *Energy Economics* 81, pp. 1042–1055.
- Johansen, Søren and Morten Ørregaard Nielsen (2012). "Likelihood inference for a fractionally cointegrated vector autoregressive model". In: *Econometrica* 80.6, pp. 2667–2732.
- Kamara, Avraham, Xiaoxia Lou, and Ronnie Sadka (2008). "The divergence of liquidity commonality in the cross-section of stocks". In: *Journal of Financial Economics* 89.3, pp. 444–466.
- Karolyi, G Andrew, Kuan-Hui Lee, and Mathijs A Van Dijk (2012). "Understanding commonality in liquidity around the world". In: *Journal of Financial Economics* 105.1, pp. 82–112.

- Klein, Olga and Shiyun Song (2017). "Multimarket High-Frequency Trading and Commonality in Liquidity". In: *Available at SSRN* 2984887.
- Koch, Andrew, Stefan Ruenzi, and Laura Starks (2016). "Commonality in liquidity: a demand-side explanation". In: *The Review of Financial Studies* 29.8, pp. 1943–1974.
- Kumar, Raman, Atulya Sarin, and Kuldeep Shastri (1995). "The impact of index options on the underlying stocks: The evidence from the listing of Nikkei stock average options". In: *Pacific-Basin Finance Journal* 3.2-3, pp. 303–317.
- Kyle, Albert S (1984). "Market structure, information, futures markets, and price formation." In:
- Lee, Charles MC and Mark J Ready (1991). "Inferring trade direction from intraday data". In: *The Journal of Finance* 46.2, pp. 733–746.
- Lee, Hsiu-Chuan, Yung-Ching Tseng, and Chung-Jen Yang (2014). "Commonality in liquidity, liquidity distribution, and financial crisis: Evidence from country ETFs". In: *Pacific-Basin Finance Journal* 29, pp. 35–58.
- Liu, Jiatao et al. (2019). "Factor Structure in Cryptocurrency Returns and Volatility". In: *Available at SSRN* 3389152.
- Ljung, Greta M and George EP Box (1978). "On a measure of lack of fit in time series models". In: *Biometrika* 65.2, pp. 297–303.
- Lo, Andrew W and A Craig MacKinlay (1988). "Stock market prices do not follow random walks: Evidence from a simple specification test". In: *The review of financial studies* 1.1, pp. 41–66.
- Marsh, Ian W. and Paolo Mazza (2017). "Common Factors in Market Quality: Evidence from European Stocks". Working Paper.
- Marshall, Ben R, Nhut Nick Hoang Nguyen, and Nuttawat Visaltanachoti (2019). "Bitcoin Liquidity". In: *Nuttawat, Bitcoin Liquidity (May 31, 2019)*.
- Moshirian, Fariborz et al. (2017). "The determinants and pricing of liquidity commonality around the world". In: *Journal of Financial Markets* 33, pp. 22–41.
- Nicole, Bullock (2017). "DRW leads high frequency trading charge into cryptocurrencies". In: *Financial Times*.
- Norden, Lars and Martin Weber (2004). "Informational efficiency of credit default swap and stock markets: The impact of credit rating announcements". In: *Journal of Banking & Finance* 28.11, pp. 2813–2843.
- O'Hara, Maureen and Mao Ye (2011). "Is market fragmentation harming market quality?" In: *Journal of Financial Economics* 100.3, pp. 459–474.
- Pástor, L'uboš and Robert F Stambaugh (2003). "Liquidity risk and expected stock returns". In: *Journal of Political economy* 111.3, pp. 642–685.
- Phillips, Peter CB, Shuping Shi, and Jun Yu (2015). "Testing for multiple bubbles: Historical episodes of exuberance and collapse in the S&P 500". In: *International economic review* 56.4, pp. 1043–1078.
- Phillips, Peter CB, Yangru Wu, and Jun Yu (2011). "Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values?" In: *International economic review* 52.1, pp. 201–226.
- Phillips, Peter CB and Jun Yu (2011). "Dating the timeline of financial bubbles during the subprime crisis". In: *Quantitative Economics* 2.3, pp. 455–491.
- Pieters, Gina and Sofia Vivanco (2017). "Financial regulations and price inconsistencies across Bitcoin markets". In: *Information Economics and Policy* 39, pp. 1–14.
- Poon, Joseph and Thaddeus Dryja (2016). *The bitcoin lightning network: Scalable off-chain instant payments*.
- Reid, Fergal and Martin Harrigan (2013). "An analysis of anonymity in the bitcoin system". In: *Security and privacy in social networks*. Springer, pp. 197–223.

- Roll, Richard (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.
- Ryznar, Margaret (2018). "The Future of Bitcoin Futures". In: *Hous. L. Rev.* 56, p. 539.
- Sensoy, Ahmet (2019). "The inefficiency of Bitcoin revisited: A high-frequency analysis with alternative currencies". In: *Finance Research Letters* 28, pp. 68–73.
- Sockin, Michael and Wei Xiong (2018). "A model of cryptocurrencies". In: *Unpublished manuscript, Princeton University*.
- Su, Chi-Wei et al. (2018). "Testing for multiple bubbles in bitcoin markets: A generalized sup ADF test". In: *Japan and the World Economy* 46, pp. 56–63.
- Subrahmanyam, Avanidhar (1991). "Risk aversion, market liquidity, and price efficiency". In: *The Review of Financial Studies* 4.3, pp. 417–441.
- Tiwari, Aviral Kumar et al. (2018). "Informational efficiency of Bitcoin—An extension". In: *Economics Letters* 163, pp. 106–109.
- Urquhart, Andrew (2016). "The inefficiency of Bitcoin". In: *Economics Letters* 148, pp. 80–82.
- Vidal-Tomás, David and Ana Ibañez (2018). "Semi-strong efficiency of Bitcoin". In: *Finance Research Letters* 27, pp. 259–265.
- Vujičić, Dejan, Dijana Jagodić, and Siniša Randić (2018). "Blockchain technology, bitcoin, and Ethereum: A brief overview". In: *2018 17th International Symposium INFOTEH-JAHORINA (INFOTEH)*. IEEE, pp. 1–6.
- Wei, Wang Chun (2018). "Liquidity and market efficiency in cryptocurrencies". In: *Economics Letters* 168, pp. 21–24.

