LUISS

Department
of Economics and Finance

Course of Empirical Finance

Arbitrage and Trading in Cryptocurrency Markets

Prof. Paolo Santucci De Magistris
SUPERVISOR

Prof. Stefano Grassi
CO-SUPERVISOR

712221
CANDIDATE

Academic Year 2018/2019

Table of Contents

Introduction

Cryptocurrencies have experienced a noticeable rise and received a lot of attention over the past few years. These digital currencies are based on blockchain that is able to provide payments verification without a centralized custodian. Bitcoin came into existence in 2009, launched by anonymous person or group of persons, called Nakamoto. Over the past 10 years since that time, the cryptocurrency market has changed a lot. A lot of new coins have appeared, some of them are quite different from Bitcoin from a technological point of view. The total number of cryptocurrencies at the moment is more than two thousand, according to the Coinmarketcap website. All these altcoins are traded on more than 200 digital exchanges across the world. The average daily trading volume of Bitcoin is more than $20 billion as of May 2019, and for the entire cryptocurrency market, this figure is $ 65 billion. However, according to some sources, these figures might be significantly exaggerated (SEC, 2019); this topic will be analyzed further in the paper. It is estimated that the amount of active traders exceeds 15 million, including both retail and institutional investors (such as DRW, Jump Trading, or Hehmeyer Trading).

While the cryptocurrency market is still young, it provides many opportunities for economic research. This paper touches the topics of cross-exchange arbitrage, the efficient market hypothesis, order book liquidity and wash-trading. First, in different geographical areas and jurisdictions, there are many non-integrated digital exchanges that operate in parallel. Most of them are almost unregulated and owned and managed privately. Most of these exchanges operate as ordinary stock markets, where traders place orders, and the exchange clears transactions based on a centralized order book. However, cryptocurrency exchanges also have many differences from stock markets. For example, there is no guarantee of the best price, as provided on traditional markets by the Securities Exchange Commission's NBBO rule. The National Best Bid and Offer rule helps retail investors who may not have the capacity to compare prices on several exchanges, providing them with the best price for the submitted order. The absence of such mechanisms implies that the comparison of prices on different markets and the subsequent selection of one of them for placing an order lies on the shoulders of market participants. Secondly, the stock exchanges are scattered around the world, and today the largest and the most liquid ones are located in Europe, Asia (Hong Kong, Japan, Korea) and the United States. But between many countries there are barriers to the free movement of capital, and the exchanges themselves in some cases do not allow foreign citizens to open an account. Such market segmentation creates opportunities for large players who have both a large amount of funds and opportunities for the movement of capital between countries.

This paper is structured as follows. The first chapter provides a number of stylized facts about cryptocurrencies, their classification and the underlying technology. After that the prospects of Bitcoin as a replacement for fiat money is discussed. First, we demonstrate that there are substantial opportunities for arbitrage trading in bitcoin across exchanges, that open up repeatedly and often exist for long periods of time, hours, days and sometimes even months. The second chapter discusses the practical aspects of cryptocurrency trading. The results demonstrate that most of the trading volumes are fictitious, and also reveals some ways of manipulating the trading volumes used by dishonest exchanges. After assessing the real liquidity and market depth of cryptocurrency exchanges, which do not inflate their volumes, an analysis of the arbitrage opportunities was performed. The obtained results show that the price dispersion across exchanges is present even on the largest and the most liquid exchanges. We build an arbitrage index to demonstrate that during some periods of time, the price difference between particular exchanges has been more than 50%, which is absolutely a violation to the law of one price. After that there is a discussion of why arbitrage opportunities exist and do not close for a long time (sometimes up to 2 months), and possible strategies for arbitrage traders.

Chapter 1. General overview of cryptocurrencies and underlying technology

1.1 Technological aspects of blockchain and bitcoin

In 2008, the pseudonymous "Satoshi Nakamoto" posted a white paper describing an implementation of a digital currency called bitcoin that used blockchain technology. More than ten years later, hundreds of cryptocurrencies and innumerable other applications of blockchain technology are readily available. The rise of cryptocurrencies poses an existential threat to many traditional functions in finance. Cryptocurrencies embrace a peer-to-peer mechanism and effectively eliminate the "middleman", which could be a financial institution. Users don't need a bank account or credit card to make transactions in the world of cryptocurrencies. According to WorldBank 2017 and Global Findex 2017 reports, more than two billion people in the world do not use banking services. The widespread adoption of smartphones and the internet creates the capacity for a revolution in financial services world.

Moreover, the possible market of this technology is more than providing financial services to people who don't use banks. It holds the potential for cheap, secure, and near-instant transactions, allowing billions of people to join the world of internet commerce, paying, and being paid, for goods or services, outside of the traditional banking and credit card infrastructure. Cryptocurrencies transactions potentially enable near real-time micropayments. Credit cards are not designed to be used for a one-cent charge to download, for example, a product or service from the internet. Cryptocurrency systems promise to make micropayments seamless and allow businesses to offer real-time pay-per-use consumption of their products, such as video, audio, cell phone service, utilities, and so forth.

A cryptocurrency like bitcoin can be thought of as a decentralized autonomous organization (DAO), an open-source peer-to-peer digital network. In this DAO setting, the money supply is set by an algorithmic rule, and the integrity of the network replaces the need to trust the integrity of human participants. The growth of cryptocurrency technology therefore poses a challenge to traditional monetary authorities and central banks. Central banks understand this, and many banks have initiated their own national cryptocurrency initiatives (Bech and Garratt, 2017). As with any new technology, risks are present. In the nascent cryptocurrency market, one concern involves the anonymous nature of transactions in some cryptocurrencies, which could allow nefarious actors to conduct illegal business, or worse, to pose a broader threat to our society and institutions (Foley et al., 2018). The benefits, such as low transaction cost, security and the promise of quick processing, are readily measurable, but quantifying the risks is less straightforward.

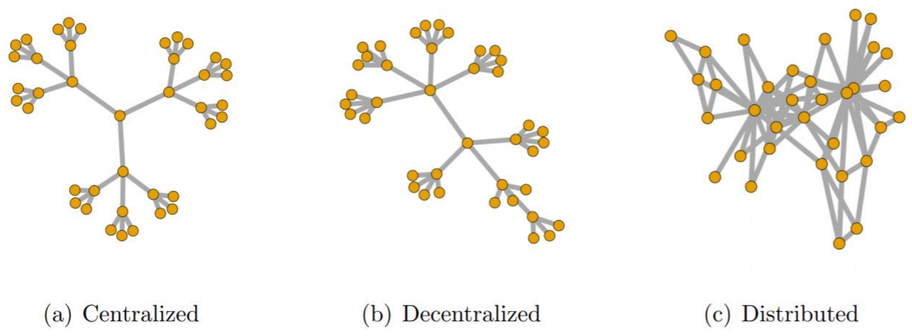
The concept of supplementary (Delmolino et al., 2016), alternative (Ametrano, 2016), or digital currencies (Chaum, 1982) is not new, but the concept of an open-source currency without a central point of trust, such as a central distribution agency or state lead control, is new (King and Nadal, 2012). A cryptocurrency is a digital asset designed to work as a medium of exchange using cryptography to secure transactions, to control the creation of additional value units, and to verify the transfer of assets. Many different cryptocurrencies exist, each with their own set of rules (see, for example, coinmarketcap.com; Iwamura et al., 2014; Park et al., 2015). Differences among the cryptocurrencies may involve, for example, the choice of the consensus mechanism, the latency, or the cryptographic hashing algorithms.

1.1.1 Blockchain, mining and hash function

Abadi and Brunnermeier (2018) describe a blockchain trilemma, i.e. that no ledger can satisfy all ideal qualities of any recordkeeping system — correctness, decentralization, and cost efficiency — simultaneously. Yet, a blockchain is more efficient than a centrally managed traditional ledger (Babich and Hilary, 2018a). A blockchain can be implemented in many ways, but most share several common features. Blockchain can be thought of as a very special database. Its structure is shared, or distributed, rather than centralized, and thus is often referred to as distributed ledger technology (DLT). Figure 1-1 shows a distributed network. As it will be discussed later, the distributed network provides some level of security, because it is unlikely an attack can be launched on every copy of the database. Distributed databases are not new, and most distributed databases are not blockchains. The key difference between a regular distributed database and one set on a blockchain is the structure (Babich and Hillary, 2018b).

A blockchain is divided into sub-sheets of data, each one called a block. At the end of each block there is a digest that summarizes the contents of the block. The digest is repeated as the first line of the next block. If any change is made in the content of a historical block, the digest changes for that block and it will not match the first line of the next block. When the network detects such an inconsistency, it throws out the corrupted block and replaces the block with the original. In this sense, the database is immutable. Given this structure (i.e., data organized in blocks with updates to the blockchain being append-only, based on the respective consensus mechanism), it is extremely unlikely that history can be rewritten. The digest at the end of a block and at the beginning of the next is generated by a cryptographic hashing function.

Figure 1-1. Types of network



(Source: Popescu et al., "Surveying Position Based Routing Protocols for Wireless Sensor and Ad-hoc Networks")

Bitcoin and similar digital currencies are called cryptocurrencies because the underlying algorithms and security are intimately related to digital cryptographic algorithms. Unlike fiat currency issued by governments, a publicly available database records every trade of currency. Every bitcoin is associated with an address, and a transaction is a trade of bitcoins from one address to another. This database is called the blockchain. A transaction in bitcoin is not final until it is included in the blockchains available from many sources. No bitcoins exist or are held independently of the blockchain. To keep the blockchains, scattered around the peer-to-peer network the same, there is a rule that the correct blockchain is the longest one. Additions to the blockchain are made as part of the process of mining bitcoins.

Miners add to the blockchain by solving a computational problem and adding new transactions. Finding the solution to the problem provides "proof-of-work" which verifies that the miner did the work. Other miners can verify at low cost that the solution has been found, although reproducing the work is not low cost. Miners compete to add the next chain to the blockchain, which includes the record of the miner's acquisition of the new bitcoins and recent transactions. Transactions fees provide an incentive for miners to include recent transactions. While bitcoins are being produced, miners also receive new bitcoins, and this currently is the major payoff from adding to the blockchain. In order to add to the blockchain, a miner starts from a hash of certain information in specific fields. The information in each increment of the blockchain includes information about new transactions including bitcoins received by the miner, a hash referencing the previous increment to the blockchain, the hash of the transactions in this increment and identifying information for the block.

A hash is a transformation of the original information. Bitcoin relies extensively on hash functions. A hash function takes a message M with arbitrary length and produces the hash value h, that is $h = H(M)$. For the blockchain, obviously the hash is much shorter than the message length. Bitcoin uses one-way hash functions, which are a subset of hash functions. One-way hash functions are not invertible except at high, preferably prohibitive, marginal cost. A one-way hash function has the following characteristics (Schneier 1996, p. 429):

1. Given M, it is easy to compute h;
2. Given h, it is hard to compute M such that $H(M) = h$
3. Given M, it is hard to find another message M' such that $H(M) = H(M')$.

Miners' difficulty in solving the computational problem is not computing the hash, which is easy. The difficulty in solving the computational problem posed for miners arises because the hash value h is restricted to be less than or equal to some value.[1] The problem is solved by searching for a hash value that is less than or equal to h^*, and miners change open fields in the message space to alter the hash and achieve h^*. There is a target for Bitcoin of having an increment to the blockchain roughly every ten minutes and, as the amount of mining increases, the difficulty is increased by reducing h^*.

Miners can increase the probability of finding a small enough hash value by using faster computers and more computers. Specialized devices are sold to mine bitcoins. In addition, miners form pools to work on finding a small enough hash value, effectively pooling their computers. Miners participate in some of these pools on a piece-rate basis. Miners also participate in some of these pools as employees, who receive a fixed payoff whether or not the pool finds a small enough hash first. Mining is a contest. Multiple miners and mining pools are working simultaneously on finding a small enough hash. Because there is no guarantee of being the first to find a small enough hash, the actual outlay of resources by a miner or pool of miners is unlikely to be as high as the value of the bitcoins received from being successful. If miners are maximizing expected earnings, resource use by any pool will be as high as the expected value of bitcoins received on finding a hash less than or equal to h^*.

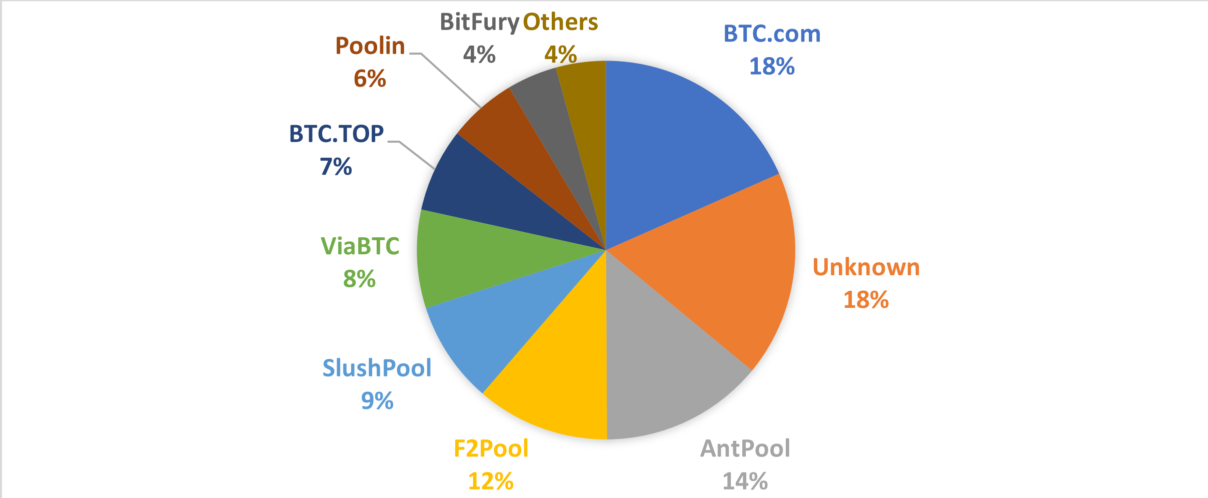
In order to maintain its reliance on competition in mining, it is important that mining be distributed across mining pools. There is an underlying reason for such combinations of miners. A miner participating in the mining contest faces the idiosyncratic risk of losing the contest. By pooling resources with others, the miner can reduce their idiosyncratic risk. In the limit, if all miners participate in one pool, there is no idiosyncratic risk of losing the contest. This pooling of risk creates an incentive for miners to cooperate in the largest pool. While it does not necessarily suggest that mining eventually will be dominated by one mining pool, it is a tendency contrary to mining being competitive.[2]

---

[1] This is described as requiring leading zeroes in the hash because the overall hash has a maximum value.

[2] Details are provided in Dwyer (2014). There is a contrary tendency. There is no known way to prove that a particular set of transactions will produce a small enough hash. If one pool were mining, it would have to have a rule for when

Figure 1-2. Current Bitcoin hash rate distribution as of May 2019



(Source: blockchain.com)

Note: the graph shows the market share of the most popular bitcoin mining pools. The numbers are estimated by the blockchain.com and should only be used as a rough estimate. The "unknown" portion of blocks doesn't mean an attack on the network, it simply means that the origin was undetermined

Cryptocurrency mining is a little similar to gold mining. Gold mining is expensive, cryptocurrency miners also spend computing power to find the hash as described above. A gold miner only gets rewarded if gold is found. Cryptocurrency miners only get rewarded if they are the first to find the winning hash. Like mining for gold, mining for cryptocurrency is risky. The continuous expenditure of resources such as for hardware and energy for a prolonged period without being rewarded is an inherent risk. The website blockchain.info presents information which suggests that mining has generated negative net revenue since February 2018.[3] This of course is possible if mining has positive nonpecuniary returns, for example if a mined bitcoin is worth more to a miner than a purchased bitcoin, or if miners can use others' resources to mine.

The announced limit on the number of bitcoins is 21 million. The increase is determined by a simple rule which halves the increase every four years (Nakamoto 2009) and generates a decreasing increase over time. This inelasticity of supply is viewed as an advantage by some economists and a disadvantage by others. An inelastic supply is roughly in line with Friedman's solution for the optimal quantity of money if the income elasticity of the demand for the money is one and loss of bitcoins is unimportant. From the viewpoint of a private currency such as Bitcoin,

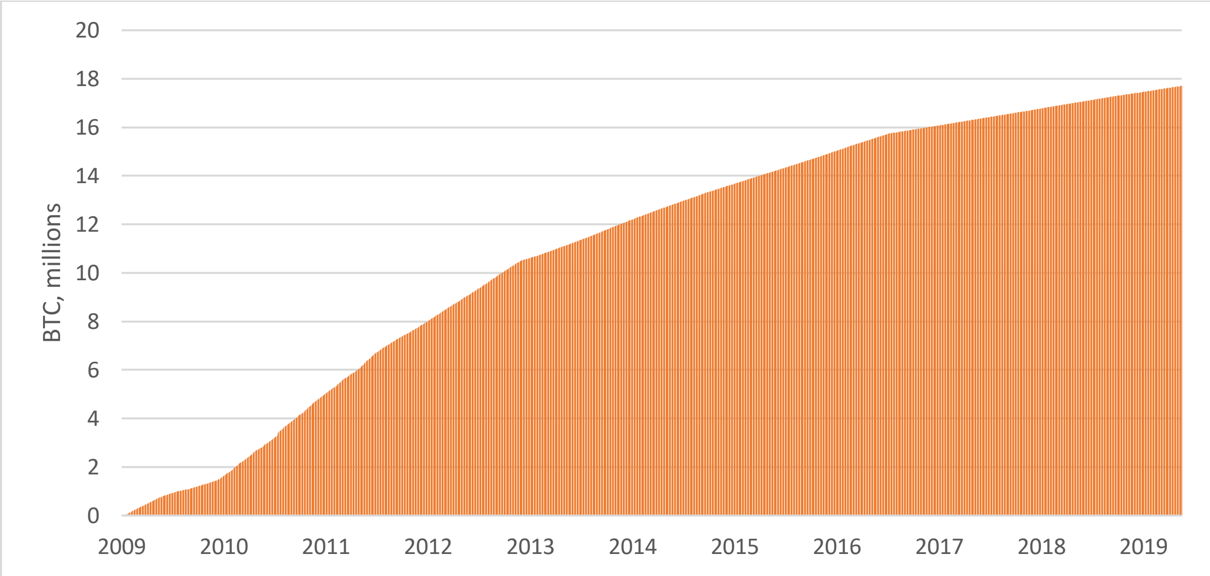
---

to change the set of transactions to search for a small enough hash. Two or more mining pools effectively solve this problem by searching with different transactions at the same time.

[3] See https://www.blockchain.com/charts/miners-revenue?timespan=2years and https://btc.com/tools/mining-calculator

an advantage is predictability of the quantity produced even if a different rule for the evolution of the stock of bitcoins would have advantages.

Figure 1-3. The total number of bitcoin that have already been mined (current supply of bitcoin network)



(Source: blockchain.info)

1.1.2 Types of cryptocurrencies

A currency without an intrinsic value, such as a Bitcoin, can only function if sufficient market acceptance is present and if the belief exists that the currency has the value attributed to it. With a conventional fiat system, money has value because people trust the central bank. For a cryptocurrency, additions to the public ledger are confirmed by a crowd of participants. There is no central bank, and participants do not need to trust each other — trust only applies to the algorithm and the network that defines the particular blockchain. A transaction is only valid if the output is equal to the input, that is, the transactor actually has the funds he/she wants to transfer. The only exceptions are new issues of the cryptocurrency, which are algorithmically predetermined.

Proof-of-work makes it unlikely that a historical block and all subsequent blocks can be altered, although such risk exists, and it is referred to as 51% attack risk. Nakamoto (2008) states that if a single entity gains 51% of the computing power, "it would be able to prevent new transactions from gaining confirmations, allowing them to halt payments between some or all users. The attackers would also be able to reverse transactions that were completed, while they were in control of the network, meaning they could double-spend coins." In fact, one mining pool has approached 50 percent of computing power at least twice.[4]

[4] Mining pool Ghash.io in 2014, see https://www.investopedia.com/terms/1/51-attack.asp

Proof-of-work is only one approach to consensus, many alternative mechanisms exist, and they may not entail the high equipment and energy costs that bitcoin miners face. The second leading cryptocurrency, Ethereum, used a similar proof-of-work mechanism. But in 2014 Ethereum switched to a proof-of-stake mechanism. Instead of allocating block mining proportionally to the relative hashing power, the proof-of-stake protocol allocates blocks proportionally to the current holdings (Buterin, 2014). As a result, the participants with the most cryptocurrency are particularly incented to do the right thing to keep the system running and healthy. Such a method holds the promise of improved latency and substantially less energy consumption. A participant who possesses 1% of the cryptocurrency could mine 1%, on average, of the proof-of-stake blocks. Ethereum has several other differences from bitcoin. Its blocks are added approximately every 14 seconds rather than every 10 minutes, and importantly, ethereum allows for smart contracts, or small computer programs, to be deployed in its blockchain. These smart contracts are run redundantly on each node.

Many other consensus mechanisms are currently available: STEEM's proof-of-brain rewards participants for creating and curating content in their social network (STEEM.io Bluepaper) and Slimcoin's proof-of-burn bootstraps one cryptocurrency off another by demonstrating proof of having “burnt” some units of value by sending a specific amount to a verifiable unspendable address (Slimcoin Whitepaper), or different implementations of the Byzantine fault tolerance, which was first described as the Byzantine Generals' Problem by Lamport et al. (1982), are used by systems such as NEO, Stellar and Hyperledger Fabric.

Cryptocurrencies can be divided into seven broad categories. Bitcoin falls into the first category; it was originally designed as a transaction mechanism. Litecoin is very similar to bitcoin and was one of the first alternatives to bitcoin. Litecoin's blocks are added every 2.5 minutes, on average, compared to every 10 minutes for bitcoin. Ethereum falls into the second class: a distributed computation token. Smart contracts permit trusted transactions and agreements to be carried out among disparate, anonymous parties without the need for a central authority, legal system, or external enforcement mechanism. The contracts are self-executing, with the terms of the agreement between buyer and seller being directly written into lines of code. The code and the agreements contained therein exist across a distributed, decentralized Ethereum network. Other examples in this class include Tezos, EOS and DFinity. The third class of cryptocurrency is called a utility token. A utility token is a programmable blockchain asset. One example is Golem, a currency that allows the user to buy computing power from a network of users or to sell excess capacity to others. Storj is similar and allows the user to rent out unused disk storage. Other examples in this class are Sia and FileCoin.

The fourth class of cryptocurrency is a security token, a token that represents stocks, bonds, derivatives, or other financial assets. New security token offerings are called STOs. This type of token could lead to substantial efficiency gains in both clearing and settlement. The fifth class is called fungible tokens. The most popular is called ERC-20 which is issued on the ethereum blockchain. A non-fungible token is the sixth classification. In this case, each token is unique and not interchangeable with the others. One popular protocol is Ethereum's ERC-721. Dhrama debt agreements fall into this classification. Two other examples of non-fungible tokens are Cryptokitties and Decentraland (LAND).

The final class of cryptocurrencies are called stablecoins. There are four categories. The first category is collateralized with fiat currency. This includes stablecoins such as Tether (USDT) and Circle's USDC. These cryptocurrencies are designed to be fully collaterized by US dollar deposits. LBXPeg is tied to pound sterling. An emerging market, Mongolia has a cryptocurrency called Candy tied to their currency. This class also includes national cryptofiats. As mentioned earlier, many central banks are investigating the potential Fedcoin (US Federal Reserve), Eurocoin (European Central Bank), CADCoin (Bank of Canada), for example. Venezuela already issued a national crypto called Petro. The second category of stablecoins are collateralized with real assets. Examples include currencies that are collateralized by gold (Digix Gold, DGX), a basket of seven precious metals used in technology (Tiberius coin, TCX) or even Swiss real estate (Swiss Real Coin, SRC). The third category of stablecoins are cryptocurrency collateralized. The leading example is the collateralized debt positions that MakerDAO offers, which enable its DAI coin to be pegged to the US dollar. The final category of stablecoins are uncollateralized. An example of this type is the Basis project and their Basecoin which has been put on hold given regulatory concerns.

This list of classifications is not exhaustive because many cryptocurrency concepts, such as Overlay, do not easily fit within this seven-category classification. The point is simple: cryptocurrencies have many uses and characteristics that extend beyond the traditional cryptocurrencies of Bitcoin and Ethereum.

1.1.3 Wallets

The evidence of ownership of bitcoins is entirely in the blockchain. Holders of bitcoins use "wallets" to keep track of their balances as well as to send and receive bitcoins. Despite the use of the word "wallet", this wallet does not contain bitcoins. It is more of a spreadsheet program which keeps track of a balance than a wallet full of currency. Every bitcoin is associated with an "address", which is the name for a public key in Bitcoin transactions. Public-key cryptography is essential for recording transactions and keeping track of the balance held by any individual. Public-key cryptography relies on private and public keys to encrypt and decrypt messages and this is crucial for verifying whether a transaction is valid.[5] The address to which bitcoins are sent is the recipient's public key. The sender's digital signature is an encryption using the private key, which can be unencrypted using the sender's public key. In this way, the sender is verified, and the address of the recipient is known.

The digital wallet keeps track of the public key, called the address, and the private key. If someone loses their private key, the bitcoins are lost because it is not possible to produce the digital signature to transfer the bitcoins to anyone else without that private key. If an intruder into a computer obtains access to someone else's private key, the intruder can send the bitcoins to an address using the private key, effectively stealing the bitcoins. There is no way for the victim to recover the bitcoins even though the victim knows the thief's address (which is a public key). The victim does not know the thief's private key and cannot reverse the transaction. By the name "address", it might seem that an address would identify the thief, but any user can create an arbitrary number of sets of private and public keys with no reason to identify a particular person or computer with any public key. Furthermore, the trail of transactions can be obscured by trades of bitcoins designed to obscure the trail.[6]

1.1.4 Transactions authentication

Bitcoin uses authentication by a peer-to-peer network to solve the double-spending problem, which is quite different from central authentication proposed by Chaum, Fiat and Naor (1990) for example.[7] Multiple websites maintain copies of the blockchain and update their copies by making copies from other nodes on the network. To understand which copy is the correct one, there is a rule that the longest valid chain available on the Internet is the correct version. Nodes obtain copies of the database from other nodes when the other nodes have longer chains. Transactions can occur in a matter of seconds, although the risk of double spending is not reduced to a low level for ten or more minutes when it is included in a block in the chain. The risk of double spending in fast transactions cannot be eliminated (Karame, Androulaki and Capkun, 2012).

[5] The recipient of a message has a private key known only to them and a public key which is widely known. The sender encrypts the message with the public key. The recipient then decrypts the message with the private key known only by the recipient.
[6] There are real limits to the ability to obscure the trail of bitcoins without giving up ownership of the bitcoins to an anonymous party for a while and possibly forever if the anonymous party does not return bitcoins. Meiklejohn at al. (2013) provide a very informative tracking of bitcoins.
[7] The most obvious way to authenticate transactions is to have a trusted central authority inform a recipient of the currency that the currency is indeed owned by the other party to the transaction. The central authority then updates the database on the ownership of the currency and the transaction occurs. The novelty in the solution proposed by Chaum et al. (1990) was anonymity of the exchange partners.

Copies of the database are maintained because miners maintain copies as part of mining. Miners must have a copy and be linked to other sites in order to post their solution to the computational problem in the database. In addition, if someone else solves the cryptographic problem first and this information is likely to be reasonably widely known, miners' optimal strategy is to move onto the next block. Hence, miners have an incentive to update frequently and stay informed about the state of the blockchain. Furthermore, they have an incentive to make this information available to others.[8]

By design, the determination of valid transactions is one CPU, one vote. Otherwise, someone could become a controlling force for determining blocks by using multiple email or network addresses, which are much cheaper to acquire than acquiring more than 50 percent of the computational power on the Bitcoin network. On occasion, more than one new block is added to a set of previous blocks. Which block is correct? The rule is to use the longest block. While there can be more than one longest block at any one time, accretions soon result in one block becoming the longest block and being used.

What is to prevent a node from substituting a solution for a prior block, adding solutions for later blocks and creating the largest block? This is an example of a "Sybil attack": an attack by creating clones of valid nodes. The authentication by the longest chain could be subject to such an attack. In this context, such an attack would involve creating earlier apparently valid transactions and the longest chain, thereby appropriating coins earned by other miners. This attack requires that the attacker have more than 50 percent of the computing power among miners.

1.1.5 Equilibria with positive values for Bitcoin

Why would anyone use digital currency? As with physical currency, the most obvious reason is a low cost of transfer from person to person. Digital deposits can be used in many transactions and no doubt will be used in more transactions in the future given plausible technological developments. Still, digital deposits are not transferable without the intervention, in general, of two banks and possibly a clearing institution. The payer's bank and the payee's bank both must be involved in a transfer of funds. Another aspect of currency transfers is their anonymity. Transfers of physical currency are anonymous in the sense that no agent has a central database with all transfers of currency stored.[9] While no institution has a central database of all transfers of bank deposits, aggregation of information across banks would make this possible.

[8] Each block includes the previous hash value in the newly encrypted block, which makes the blocks a chain.

[9] The U.S. government does require selected institutions including banks to report cash transactions of $10,000 or more.

Nonetheless, transfers of physical currency self-verify that an agent has receipts from one or more sources sufficient to transfer purchasing power in exchange for something else.

Bitcoin is not anonymous, and anonymity was not included as a design goal (Nakamoto 2008). While a user of bitcoins can take steps to make his identity and a sequence of counterparties less obvious, the evidence available so far does not support the proposition that it is particularly simple to hide one's sequence of transactions (Reid and Harrigan 2013). It may well be impossible. If someone desires anonymous transactions, physical currency has the advantage if there is a possibility of direct (physical) transfer. Loss of the associated private key associated with an address and its balance of bitcoins has the same consequence as the loss of paper currency: it is gone. Similarly, theft of a private key results in loss of the associated bitcoins just as does theft of paper currency.

Current fiat currencies are associated with particular countries or sets of countries, but digital currency does not need to be associated with a particular country. Hence, the common strategy of defining the real quantity of money as the nominal quantity divided by a price level for an economy identified as a country does not work for a private digital currency. Because people can only be in one place at one time and there are nontrivial time and other costs of travel, households generally are concerned with the level of prices in a particular area. In general, there seems no reason to think the demand for money is different in this respect with or without digital currency.

Prices of digital currencies including Bitcoin in various fiat currencies are readily available. Starting from price levels in terms of the prices of goods and services in a fiat money in a particular locale, the real quantity of money demanded could be determined using the exchange rate of digital currency for the currency in which local goods and services are priced. While local goods and services could be priced in terms of the digital currency, it is not necessary. If there are multiple digital currencies, at this level of generality, there is even less reason to expect prices to be denominated in any particular digital currency. There is virtually no data to decide how many bitcoins to allocate to what country and therefore there is no obvious way to compute a real quantity of bitcoins.

Because bitcoins are not redeemable in anything else from some particular agent or set of agents, bitcoins are not an immediate store of value. A full-bodied metallic coin requires resources to produce it but much of the value of the resources can be recovered by melting the currency down.[10] The valuable resources used to produce bitcoins are electricity and computer plus a small amount of related labor. All of these resources are services consumed in production and are not

---

[10] Full-bodied coins are ones for which the metal in the coin has a face value equal to the face value or close to it. A token coin Is one for which the metal is a small fraction of the face value.

available to anyone after a bitcoin is produced. They are sunk costs. It would make no difference if existing bitcoins were produced at zero marginal cost other than the relationships between mining, maintaining the blockchain and distributing new bitcoins.

Irredeemable currency raises issues not raised by redeemable currency. Redeemable currency includes a promise that the currency can be turned into something else. The value of bitcoins is determined by the demand for bitcoins in conjunction with the rules governing supply. While possibly undesirable in some respects, the rule limiting the number of bitcoins combined with the use of a peer-to-peer network for bitcoins created makes it relatively easy to determine whether additional bitcoins are being added to the stock other than those promised.

Even if bitcoins were costless to produce, there would be equilibria in which bitcoins are valued. It might seem that available theoretical results are not applicable because the theoretical literature has focused on private currency created with zero marginal cost. The production cost is irrelevant, though, once bitcoins have been produced because those costs are sunk. Hence, theoretical results are applicable. Results in Marimon, Nicolini and Teles (2012) for currency created with zero marginal cost indicate that an equilibrium with private currency held by consumers exists with commitment.[11] And knowledge of the quantity produced is a commitment device in their setup. The possibility of entry is not addressed by Marimon et al. (2012). It is possible to create a digital currency with a positive marginal cost of production as for Bitcoin, but it is possible to create other digital currencies with zero marginal cost of production. If the marginal cost of production is zero and holders of digital currency are largely indifferent between various currencies, the value of digital currency will go to zero in equilibrium.

Marimon et al. do consider the possibility of multiple currencies but as in the early paper by Klein (1974), the existence of an equilibrium with positive values for private currencies requires there be a reputational equilibrium in which the currencies are distinguishable. There must be something that distinguishes between the currencies and prevents them from being perfect substitutes. The digital representation of these currencies means that physical differences are uninteresting, although characteristics associated with finality of transactions and other characteristics may come into play. For example, Litecoin updates its blockchain more frequently than Bitcoin. Some other currencies have rules for continued creation of new coins forever. The liquidity of exchanges of a digital currency for goods and services, physical currencies and other digital currencies is a plausible differentiating factor. As for stocks in which exchanges become dominant due to liquidity on the exchange, the liquidity of the currencies is likely to be a very

[11] See also Berentsen (2006) and Martin and Schreft (2008).

important factor in determining their relative use. This characteristic suggests that a solution with the value of digital currency positive is possible although not certain.

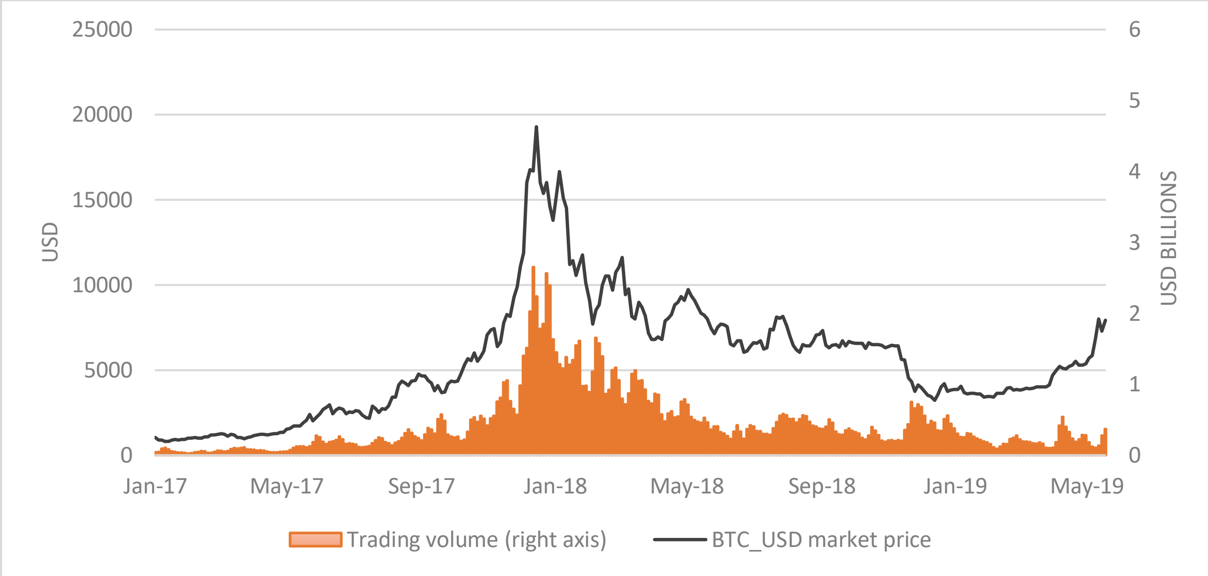
While mining new bitcoins is ongoing, miners update the record of valid transactions, because mining is impossible without making the record of valid transactions available to the network. But at some point, mining will end. The final number of bitcoins will be determined by the marginal cost of mining and the marginal return in terms of bitcoins, with an upper limit of 21 million. If mining produces a number of bitcoins falling by half every four years (Nakamoto 2008), 20.7 million bitcoins will be produced by 2041 given the algorithm. Who will maintain the blockchain of valid transactions when there is no mining? Nakamoto (2008) makes the supposition that transactions fees will support those who make the record available and update it. Such fees currently are collected but they are small relative to the new bitcoins received for completing a block. While a block would be created without transactions fees, competition among transactions to be included quickly in the blockchain results in positive fees even today because there is no incentive to include a particular transaction in a new block without a transaction fee. Babaioff, Dobzinski, Oren and Zohar (2012) point out that the structure of those fees will be more important for creating an equilibrium in which bitcoins are useful when there is no payoff in terms of new bitcoins.[12]

1.2 Bitcoin price

Figure 1-4 shows the dynamics of Bitcoin to U.S. dollars market price, and the volume traded since the beginning of 2017. This short time period was chosen because prior to this, cryptocurrency market was substantially illiquid. It is obvious that both prices and volumes have increased substantially. The first trade on Mt. Gox on July 7, 2010 was a trade of 20 bitcoins for \$0.04951.

[12] It would of course make a difference in terms of efficiency. If there is a less costly mechanism for distributing new digital currency, this is a quite inefficient mechanism for creating bitcoins. One obvious alternative would be to distribute new currency to existing holders, which has its own advantages and disadvantages.

Figure 1-4. Dynamics of market price and trading volume of Bitcoin



(Source: blockchain.info)

Is a price of USD 8000 for a bitcoin high or low? This question is even harder to answer than for governments' fiat currencies. There is no reason to use Purchasing Power Parity for bitcoins to assess the price even if it were feasible. A simple and somewhat informative way to look at the question is to examine the aggregate purchasing power in dollars represented by the quantity of bitcoins. There were about 17.5 million bitcoins already mined in May 2019 (see Figure 1-3). At a price of USD 8000 per bitcoin, this indicates an approximate value of bitcoins of approximately USD 140 billion. While not trivial, this is small compared to the value of U.S. M2 of USD 14.5 trillion for May 2019.[13] Of course, a comparison of the worldwide holdings of bitcoins to U.S. dollars of 0.97% may seem out of line. It is obvious that the U.S. dollar is in no danger of being replaced by bitcoins in terms of value. It also is obvious that the value of bitcoins in dollars outstanding today is not particularly large.

Another way of looking at the aggregate value of bitcoins is to compare their value to the value of reserves in the banking system. This comparison is suggested by the possibility that bitcoins will be useful in finalizing transactions between other fiat currencies. This is somewhat similar to the use of banking reserves to clear transactions between deposits in various banks. Before the Financial Crisis of 2007-2008, reserves in the U.S. banking system alone were USD 8.75 billion. Mostly these were clearing balances held by banks. The value of all bitcoins in March 2014 of USD 140 billion is much higher than this value of reserves in the Federal Reserve. Current reserve balances as of May 2019 are about USD 1.5 trillion[14], which is more than 10 times higher than current bitcoin market capitalization. Given the early stage of development of bitcoin, this

[13] https://fred.stlouisfed.org/series/M2
[14] https://fred.stlouisfed.org/series/WRESBAL

seems quite large if the only role of bitcoins is for finalizing transactions in U.S. dollars. On the other hand, bitcoins are not useful only in the United States and USD 140 billion of bitcoins may be small relative to reserves weighted by holdings of bitcoins and possible future use.

1.3 Prospects of Bitcoin as a currency

The most successful currencies in the world have been convertible into fixed amounts of gold or other valuable metals for the most part of the 19^{th} and 20^{th} centuries, and many currencies were physically made of gold or silver for thousands of years preceding that. This straightforward link between cash and gold, guaranteed by public inventories like the U.S. Fort Knox depository, established government trust in the worth of a currency. Between the 1920s and 1970s, the gold standard crashed in most countries, partially owing to the stresses of funding two World Wars, but likely even more so because global gold output did not maintain pace with financial development. Nearly every significant country has since released paper fiat currency, the value of which is based on public faith that the state or main bank of a nation will not raise the availability of fresh banknotes too quickly. At comparable circumstances, multinational consortia published currency like the euro. Fiat currencies have been circulating for thousands of years, and faster or slower almost all of them were inflated by governments faced with tense government finances to the point of worthlessness.

Bitcoin tries to solve both fiat and gold-based cash faults, working as an algorithmic currency with a deterministic supply and development pace linked to mathematics rigor. No state or other key regulator can manipulate bitcoin supplies. Rather, the currency is controlled by cryptographic laws which are implemented in a distributed way by transparent software code. While some people have indicated a link between the algorithmic growth process of Bitcoin and the financial standard supported by Milton Friedman (Brito, J. et al., 2015), the Bitcoin protocol seems to pay little or no consideration to any appropriate pace of monetary growth. Rather, it offers for the bitcoin emission level to decrease asymptotically to zero by the year 2140, when the last bitcoin is planned to be released and the ultimate sum is set at 21 million units.

While the constrained supply of bitcoins appears to be one of the crucial components of the currency's attraction, it has become a question of some dispute whether the supply is actually limited. In one of the interviews with Goldman Sachs (2014), the case against this perspective is indicated by bitcoin entrepreneur: “If you needed to create more, you could. That would require 51% of the computing power of the network to switch their software to adopt the change. Changes to the software have occurred a couple of times in the past. There are developer forums where such types of changes are typically discussed, and a consensus is ultimately reached across the mining community that maintains the network.” If his understanding is correct, it means that for financial policy purposes the "emission" of bitcoins could be adjusted, with these macroeconomic policy choices being governed by an internet debate forum or website rather than by an official organization such as the Federal Reserve.

Bitcoin emerged from a model described in Nakamoto (2008), a nine-page proposition for a "peer-to-peer digital currency system." The writer or writers of this paper were not recognized[15], but their system was intended so as not to grant them credits or remaining ownership privileges in order to profit from the implementation of Bitcoin. New bitcoins are developed and given to participants who find the answer of pre-specified mathematical problems according to Nakamoto's suggested algorithms. Over time, these tasks become more difficult technically, and their appearance becomes less regular. An open decentralized database records each bitcoin's possession and its later movements after its original holders "mine" it. The public knows with certainty all the quantities and growth rates of bitcoins, so its circulation cannot be impacted by financial regulation in the manner, in which the Federal Reserve handles the growth rate of U.S. dollar public supply.

Wallace (2011) describes Bitcoin's early history and reports that Nakamoto mined and put the first 50 units into circulation in 2009, primarily to show the technique to a community of internet users Initially, Bitcoin's flow arose among computer world volunteers and enthusiasts Interest expanded to the extent that in 2010 Bitcoin started trading on an internet exchange based in Japan, Mt. Gox, which was initially developed as a platform to exchange fantasy game magic trading cards. During the first day of trading on Mt. Gox, 20 bitcoins changed hands for a total volume of slightly less than one U.S. dollar at a price of 4,951 cents.

Wallace (2011) and other sources claim that the first purchase of goods using bitcoin took place in 2009, someone bought two pizzas at a price of 10,000 bitcoins. The pizza shop did not recognize Bitcoins straight, and instead a third-party dealer who decided to purchase the pizzas using a credit card (with a standard currency) and accepted the Bitcoins, worth nearly $8 million at latest rates, as compensation. Most of the bitcoin trade still takes place with intermediaries who facilitate the immediate conversation of bitcoins into fiat currencies.

The marketplace of Silk Road, an online portal for the purchase of illegal drugs that recognized only bitcoins for purchase, was sometimes stated to account for even more than 50% of the early quantity of bitcoin transactions although this figure is subject to significant controversy. The Silk Road organization contributed to creation of a criminal reputation for bitcoin from the very beginning, and this outlaw cachet may not have damaged its attraction. Bitcoin use expanded to the bricks-and-mortar economies and U.S. officials shut down Silk Road after their

[15] A controversial Newsweek story on March 6, 2014, claimed to have located and identified Nakomoto, but the magazine’s claim was denied by the subject of the story and has been subject to continuing uncertainty. See http://mag.newsweek.com/2014/03/14/bitcoin-satoshi-nakamoto.html.

operator was imprisoned in October 2013 in San Francisco. Far from destroying bitcoin, this incident created advertising that could have even increased its popularity. The first bitcoin ATM came into use in a Vancouver coffee store shortly in the same month.

Bitcoin trading on the Mt. Gox exchange and other platforms rose quickly. Many digital exchanges have started trading bitcoin along with other cryptocurrencies that have emerged as rivals[16], although some of these markets are quite illiquid and trading may at best seem episodic. With a growth in the value of bitcoin, digital exchanges became hackers' objectives. In April 2013, Mt. Gox confirmed three denial of service attacks that dramatically lowered trading volume on different dates, however the exchange appeared to restore in a couple of hours in each situation. A range of investment funds have been launched to accommodate for Bitcoin traders, including the one that tried to register with the Securities and Exchange Commission in July 2013 under the patronage of the Winklevoss brothers, who became popular in the online entrepreneurship universe because of their judicial conflicts with Mark Zuckerberg over Facebook possession.

Bitcoin attracts two different type of people. One party is made up of innovation and technology lovers embracing bitcoin for digital commerce. As more and more regular company operations tend to move online, these people think that the importance (and value, as a consequence) of bitcoin must rise because of the growing demand for payments, as well as citing its cost benefits over credit cards and other payment solutions for regular bricks-and-mortar retail shopping. The other group, to which bitcoin is appealing is a distinct community with pseudo-libertarian political views. And the reason they find bitcoin perspective as a currency is its absence of association with any state authority. Some of these people are completely distrustful of the global financial mechanism and the timing of the implementation of bitcoin, happened to coincide with the very bottom of the 2008-2009 global financial crisis, has likely caused their ranks to grow.

Passion for cryptocurrencies unites both the technology and liberal groups, but not much else. Debates regarding cryptocurrency merits can lure strange combinations of businessmen, scholars, and polemicists. An instance was given at a panel debate moderated by a journalist from the New York Times in March 2014. Bitcoin activist Andreas Antonopoulos, in reaction to a comment that up to 10 percent of the bitcoin supply has already been purloined by internet hackers, called the 10 percent robbery “a huge enhancement over the remainder of our society, where 80 percent is in the hands of criminals - and that's banks”.[17] The acclaimed Stanford finance professor

[16] See http://coinmarketcap.com/ for a list of over 2000 new cryptocurrencies, produced nearly by the day. According to accounting data on this webpage, bitcoin currently holds about 60% of the overall market value of this asset category.
[17] See http://www.coindesk.com/despite-challenges-bitcoin-technology-stay.

Susan Athey listened on from the other end of the screen with a mixture of frustration and bafflement as the crowd applauded in recognition.

Bitcoins regular trading volume structure indicates that the lion's share of global supply occurs in two regions, North America and Asia. While China has adopted several measures to prohibit bitcoin use, U.S. authorities have been more tolerant. American regulators' comparative tolerance of bitcoin may result from their acknowledgement that there is a universal internet inspection path for all bitcoin operations. Although there are facilities such as "tumblers" on the Web that tend to anonymize or hide bitcoin transactions, trust in the safety of these procedures seems insignificant. The detention of Silk Road's owner in October 2013, which occurred in the midst of widespread Internet data monitoring advertising by the U.S. National Security Agency, disappointed many regarding the chance of retaining any information pertaining to bitcoin anonymous. Tax avoidance, money laundering, contraband transactions and other unlawful operations using internet payments become much riskier when countries with appropriate technical capabilities are able to reconstruct the uses of a cryptocurrency such as bitcoin.

This section provides analysis of how bitcoin does not adhere to a currency's classic characteristics. A real currency usually works as an exchange medium, an account unit, and a store of value. In fulfilling all three of these requirements, Bitcoin has difficulties.

1.3.1 Medium of exchange.

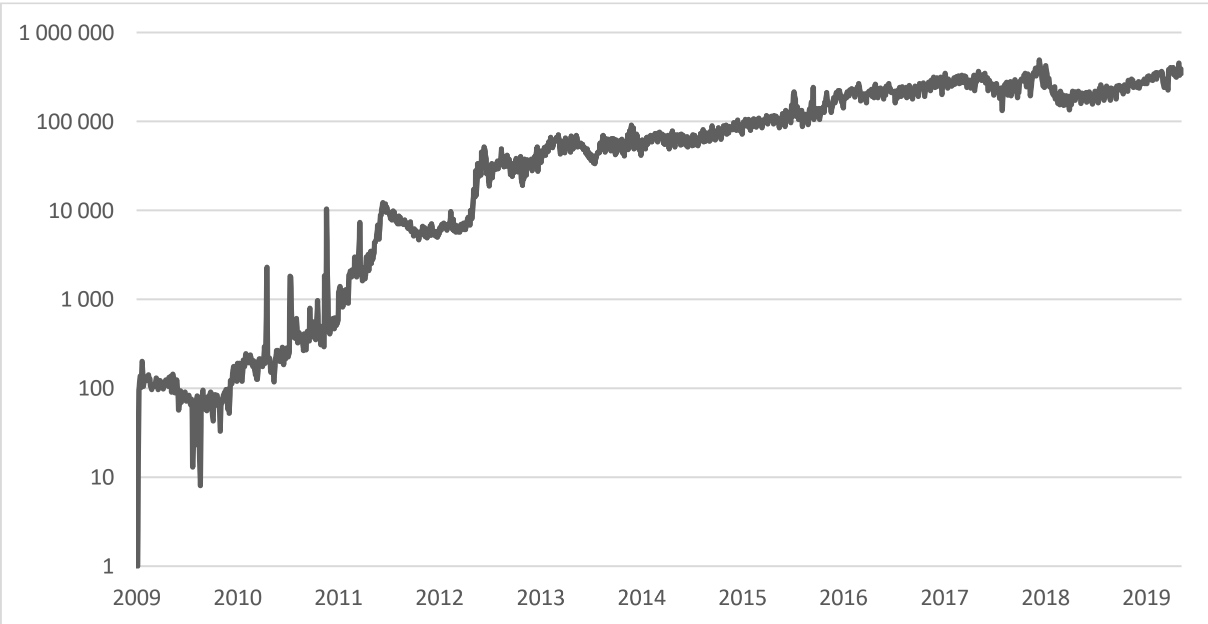
Because bitcoin does not have any fundamental value its significance eventually depends on its effectiveness as a consumer economy currency. Indication of the footprint of Bitcoin in ordinary business is mostly anecdotal, composed of journal articles about individuals residing only through bitcoin expenditure or estimates of big amounts of companies ready to embrace Bitcoin. It is difficult to calculate the precise amount of companies that receive Bitcoin because Bitcoin payment servers are very confidential about their customer stats. CoinGate, a Bitcoin payment processor of medium size, said that 4,500 dealers are actively using its software[18]. Since Bitpay, Coinbase Commerce, and BTCPay are much more common than CoinGate, it is reasonable to suppose that there are tens of thousands of dealers adopting Bitcoin. Among the companies that accept Bitcoin one can find Microsoft, Overstock, KFC Canada, Shopify, Badoo, etc.[19]

It is possible to get some insight about the implementation of bitcoin from information gathered from the global Bitcoin transaction ledger. According to information accessible on countless websites, the current number of daily Bitcoin transactions has grown at

[18] See https://coindiligent.com/who-accepts-bitcoin
[19] See https://www.lifewire.com/big-sites-that-accept-bitcoin-payments-3485965

roughly 400,000.[20](See Figure 1-5) However, it is commonly recognized that plenty of these operations are actually transfers between speculators and that only a minority are used to buy products. For example, in a March 2014 interview, Fred Ersham, Coinbase's co-founder, the major cryptocurrency wallet company, reported that 80% of his site's usage was linked to speculation, down from perhaps 95% a year previously (Goldman Sachs, 2017). If we accept this evaluation as precise then buying a good or a service from a dealer may encompass 75,000 bitcoin transfers per day. Bitcoin seems to have a negligible market share in the world with 7 billion customers, most of whom create numerous financial operations every day. In his interview, Ersham further says that Coinbase registers 24,000 companies.[21] If all the world's bitcoin trade took place within this category (almost definitely an overstatement), these companies would average well below three transaction per day per merchant. In other words, Bitcoin operations seem to be rare, even for the few companies who support them.

Figure 1-5. Number of daily confirmed bitcoin transactions (logarithmic scale)



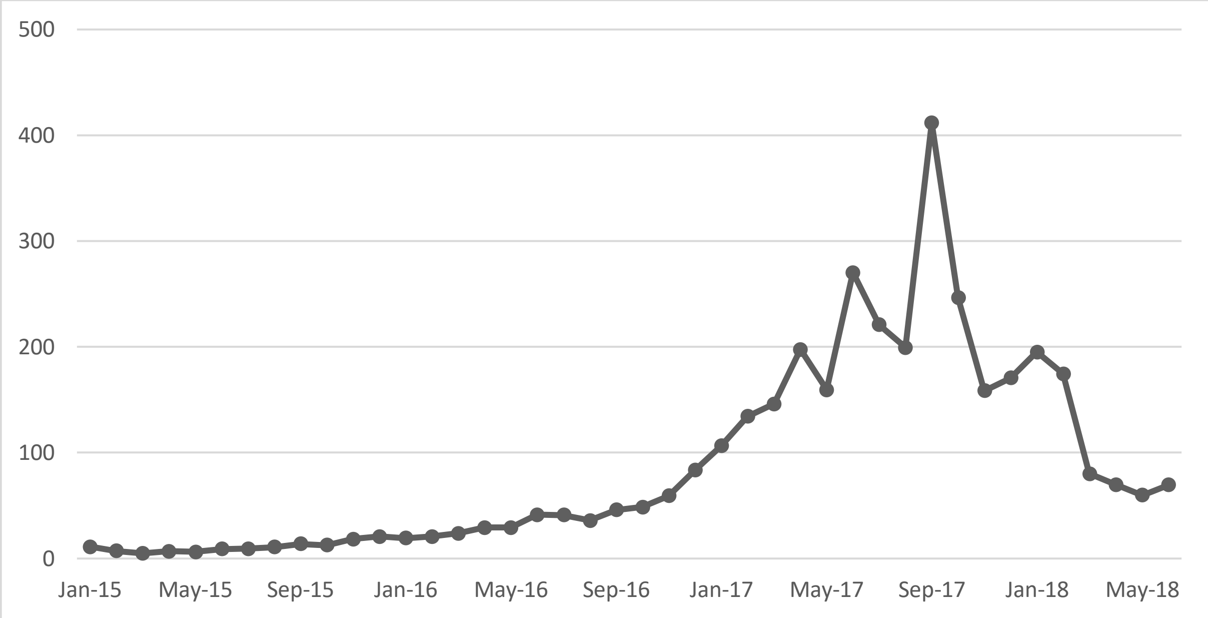
(Source: https://www.blockchain.com/charts/n-transactions)

After reaching USD 411 million in September 2017, the revenue earned in the most widespread cryptocurrency by the biggest 17 crypto merchant-processing facilities was steadily declining, reaching a latest peak of $ 60 million in May 2018 (see Figure 1-6 below), according to studies undertaken for Bloomberg News by startup Chainalysis Inc.[22] While the quantity earned

[20] See, for instance, https://blockchain.info/charts/n-transactions.
[21] The co-founder of bitcoin payment processor BitPay estimated the number of worldwide businesses at 26,000 in a separate interview given contemporaneously. See http://money.cnn.com/2014/03/17/smallbusiness/bitcoin-bitpay
[22] See https://www.bloomberg.com/news/articles/2018-08-01/bitcoin-s-use-in-commerce-keeps-falling-even-as-volatility-eases

by dealer facilities like BitPay, Coinify, and GoCoin mildly improved to USD 69 million in June, Chainalysis discovered that was still far away from the USD 270 million provided a year earlier.

Figure 1-6. Amount of Bitcoin Received by Top Merchant Processors, in millions of USD



(Source: https://bitinfocharts.com/bitcoin/)

Bitcoin supporters have long proposed that the cryptocurrency would one day substitute fiat money as a means of doing business, but bitcoin lost what little attraction it had as a way to purchase products, after an increase in usage last autumn. "It's not actually usable," said Nicholas Weaver, a senior researcher at the International Computer Science Institute. "Often, the net cost of a Bitcoin transaction is much more than a credit card transaction", he said. And Bitcoin-based purchases or transfers can't be undone. This becomes a serious problem if a dealer or customer faces cheating.

The decrease in transaction usage coincided with the rise in speculative investment, which in December 2017 pushed the value of the most widespread cryptocurrency to a record high of nearly $20,000. While the exchange rate of Bitcoin has been somewhat steady for several months after falling more than 70 %, customers still seem unwilling to use electronic currency for payments in 2019. When exchange rate volatility is as high as it was last year, one might gain or lose 20% of his/her wealth in one day. Moreover, increased transaction charges turned it impractical to pay for simple products such as coffee with Bitcoin (see Figure 1-7).

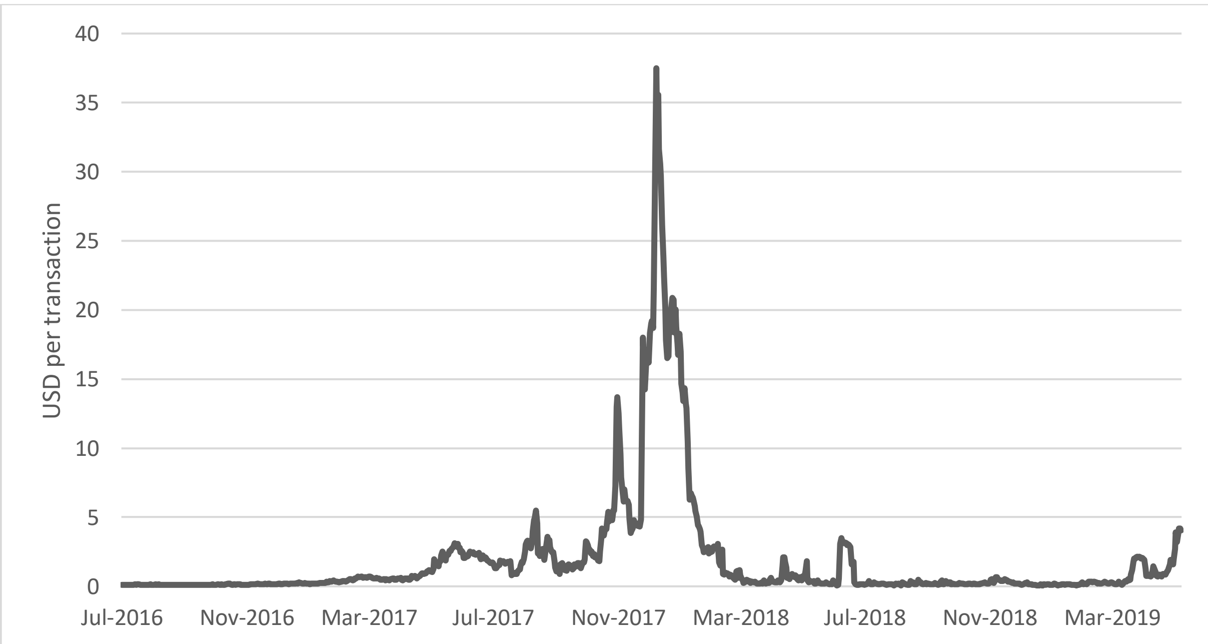
Payment processing service Stripe Inc. ceased supporting Bitcoin in February 2018 as demand decreased and market fluctuations worsened.[23] Several businesses like travel facilities supplier Expedia also discontinued recognizing the cryptocurrency. This is a disturbing indication

[23] See https://stripe.com/blog/ending-bitcoin-support

for some cryptocurrency holders who retain the conviction that bitcoin must be used around the world versus being merely a temporary tool with a long-term significance. Over the previous one-two years, Bitcoin has developed to become more suitable as an investment than as a medium of exchange, as block size boundaries have been achieved. Given the general progress the Bitcoin community has accomplished, the choices taken along the path are difficult to quibble with. However, this has made the cryptocurrency less usable for payments. The reason is that transaction verification times have increased significantly; it, therefore, has resulted in increased transaction failure frequency denominated in fiat money. (Bitcoin price volatility implies that the quantity of money is incorrect by the moment the transaction is verified.) In addition, charges have increased significantly. A charge of tens of U.S. dollars is frequent for a Bitcoin transaction, making Bitcoin transfers as costly as bank transfers or even more.

The way Bitcoin is used is also evolving. Because transaction charges can sometimes be high and volatile - they reached USD 54 in December 2017 but are down to less than USD 1 as of May 2019 - not many individuals use cryptocurrency for tiny purchases, such as purchasing a cup of coffee. They spend the digital currency more to compensate suppliers like freelancers abroad: Bitcoin can sometimes be quicker and cheaper for such occasions than using traditional economic facilities.

Figure 1-7. Historic daily average Bitcoin transaction fees (USD per transaction)



(Source: bitcoinfees.info)

"We've seen a huge rise in crypto businesses paying their suppliers in Bitcoin over the past six months, including legal firms, hosting firms, accounting firms, landlords and software suppliers," says Sonny Singh, chief business officer of BitPay processor. His firm has seen a fivefold rise in crypto businesses paying their bills since last year, he said. Bitcoin faithful consumers are still buying bigger-ticket products like furnishings and the occasional sports auto using bitcoin. “Crypto-based revenues at Overstock.com Inc. have doubled in the first quarter of 2018 compared to a year earlier”, the firm said. According to the website, “top products purchased with bitcoin include living-room furniture, bedroom furniture and laptops, according to the site”.

But again, many individuals only speculate with cryptocurrencies or sell tiny quantities to turn it into a fiat money and use it to pay for products and facilities. Graham Tonkin, chief development officer at Mosaic, a crypto-financial research firm, said he sometimes exchanges his Bitcoin and Ethereum to settle credit card charges. "I suppose that many individuals are like me, they are not going to do their daily operations in cryptocurrencies. I don't think it fits the characteristics of money very well." said Tonkin.

One barrier to Bitcoin becoming a commonly utilized medium of exchange is the challenge of obtaining new Bitcoins. Unless a customer succeeds as a bitcoin miner (an industry now overtaken by super computers demanding significant investment in resources), he or she must purchase bitcoins from internet exchanges or merchants and then find a method to safely store them. Typically, these transactions cannot be done using a credit card or PayPal, but somehow the customer must do a bank transfer or connect his bank account to the digital exchange[24]. Often, cryptocurrency markets have poor liquidity, substantial bid-ask spreads, as well as a certain level of custody and execution risk.

Finally, one cannot avoid the necessity of owning bitcoins before obtaining products or services from a seller. In most retail marketplaces, the possibility to buy goods not having money in your hand happens routinely, as people often use merchant-funded consumer loans or a third-party credit card supplier. For bitcoins, these alternatives are not accessible, although there are several startups, for example Revolut, which provides a possibility to exchange fiat money for cryptocurrency in the app and pay with cryptocurrency using its credit card.[25] But these kinds of services are not very popular and widespread yet.

1.3.2 Medium of account

People and companies must regard a currency as a numeraire when considering the prices of products and services in order for it to work as a unit of account. For example, a bottle of water in one shop costing USD 1.00 is quickly learned to be twice as expensive as a bottle of water in another shop down the road selling for USD 0.50. To become a valuable account unit, Bitcoin encounters a variety of barriers. One issue stems from its huge market volatility, a subject

---

[24] See http://howtobuybitcoins.info/us.html
[25] https://www.revolut.com

described below in more detail. Since a bitcoin's exchange rate varies widely on a daily basis relative to other currencies, businesses who accept the cryptocurrency must readjust prices very frequently, a procedure that would be expensive for the vendor and perplexing for the consumer. This problem would, in theory, disappear in an environment which used bitcoin as its main currency, but in today's world there is no such location.

A connected issue arises from the variety of "current exchange rates" that can be obtained at any specified moment for bitcoin. For example, while consulting a commonly used website coinmarketcap.com, which publishes the exchange rates of bitcoin on exchanges around the globe, as of May 2019 the top five exchanges by volume traded quoted U.S. dollar rates for one bitcoin of USD 7860, USD 7962, USD 8377.52, USD 7953, and USD 7959.6 for the most recent trades. This gap in market prices, sometimes reaching significant percentage of bitcoin price, is a simple violation of the law of one price, and it would be impossible for these circumstances to continue in an efficient market because of the ease of arbitrage. For any third-party seller who wants to create a legitimate benchmark for setting consumer prices, an inaccurate estimate of the value of one bitcoin is a problem. As a consequence, many web pages relied on not very accurate price aggregates, such as the average bitcoin exchange rate in several markets over the past 24 hours. But the problem with these aggregates is that they do not indicate the actual cost of acquiring or selling a bitcoin at current time.

Perhaps the most significant barrier for Bitcoin to become a commonly used unit of account and one that Bitcoin supporters often disregard or trivialize, happens due to high price of one Bitcoin relative to most normal goods and facilities. This would force vendors to cite bitcoin prices for most products and services in four or more decimal digits. While such arithmetic is simple, these decimal digits are probable to be perplexing for customers. A reference to one online food vendor for example, provides the following prices of a salsa bottle for 0.001694 BTC, a box of fruit for 0.000738 BTC, and a pack of tea for 0.004361 BTC. Alternatively, in scientific notation such rates could be displayed as 1.694x 10-3 BTC, 7.38 x 10-4 BTC, and 4.361 x 10-3 BTC, respectively. It is difficult to find any other currency in the world for which such numbers quote real prices of goods, and indeed, many commonly used billing software platforms can only fit two decimal digits in the cost of a product.

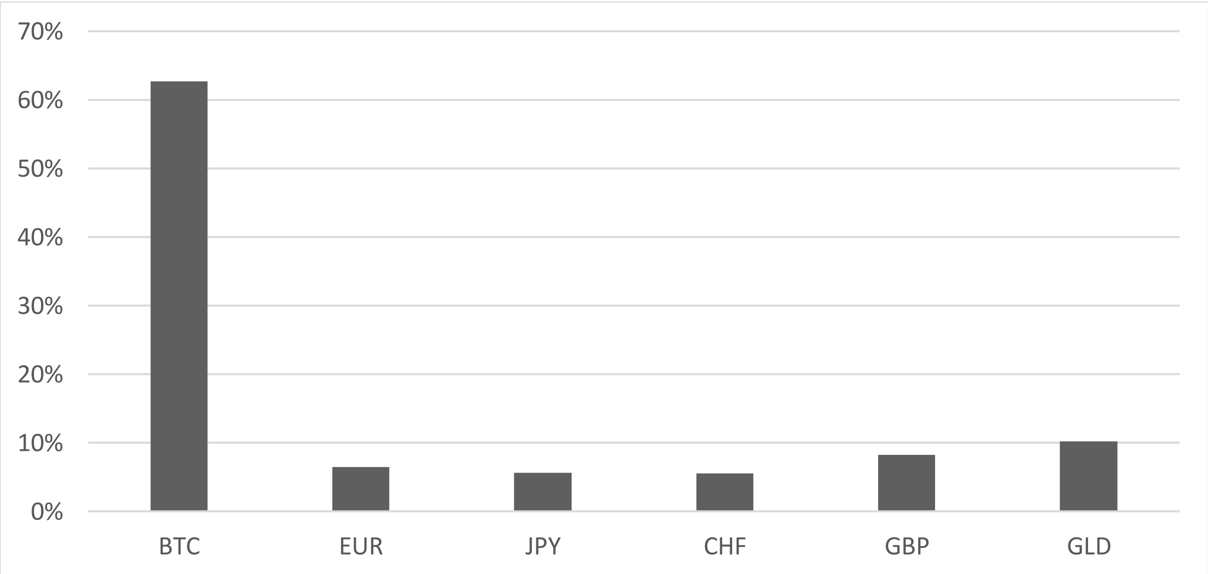
The trend generally happens in the reverse direction because of elevated inflation to the point that some countries have retail prices that deviate from ordinary integers. Italy could be a good example before the introduction of the euro; for example, one could have paid 5,000 liras for an ice cream bought from a street vendor in Rome in the 1990s, and people generally adapted to these figures ignoring zeroes at the end. Several studies in marketing have recognized customer pricing heuristics that include multiple points of reference for integer, left-digit, and right-digit.

Many products, for instance, have rates that end in number 9.[26] One problem in these studies is the consumer's computational complexity in comparing products prices. In the example above, tea is more than 5 times as costly as fruit, but even university graduate customers would often fail to achieve this result easily because of the number of digits and the existence of leading zeros. Supporters of Bitcoin tend to disregard the inability of the cryptocurrency to produce values that match normal customer reference points.

1.3.3 Store of value

If currency operates as a store of value, one is able to buy the currency at certain time and exchange it at some time in the future for products or facilities. And when the currency gets spent, one expects to gain the very same economic value as it was when he first purchased it. Throughout most of history, handling money as a store of value implied saving it from robbery, either by hiding it physically or by placing it in a bank (which then presumed the safety issue). Approaches to hide cryptocurrencies under mattresses or elsewhere cannot happen as there is no physical representation of the cryptocurrency. Alternatively, bitcoins should be kept in online accounts recognized as digital wallets. Safety is also a significant challenge for the bitcoin economy. Several electronic wallet enterprises have agreed to provide a coarse type of deposit insurance with third-party insurance carriers. Although this approach may operate in theory, it requires the client to carry the cost of assessing both the wallet service provider and the insurance company's security (economic and otherwise).

Figure 1-8. Annualized volatility of bitcoin compared to major currencies and gold

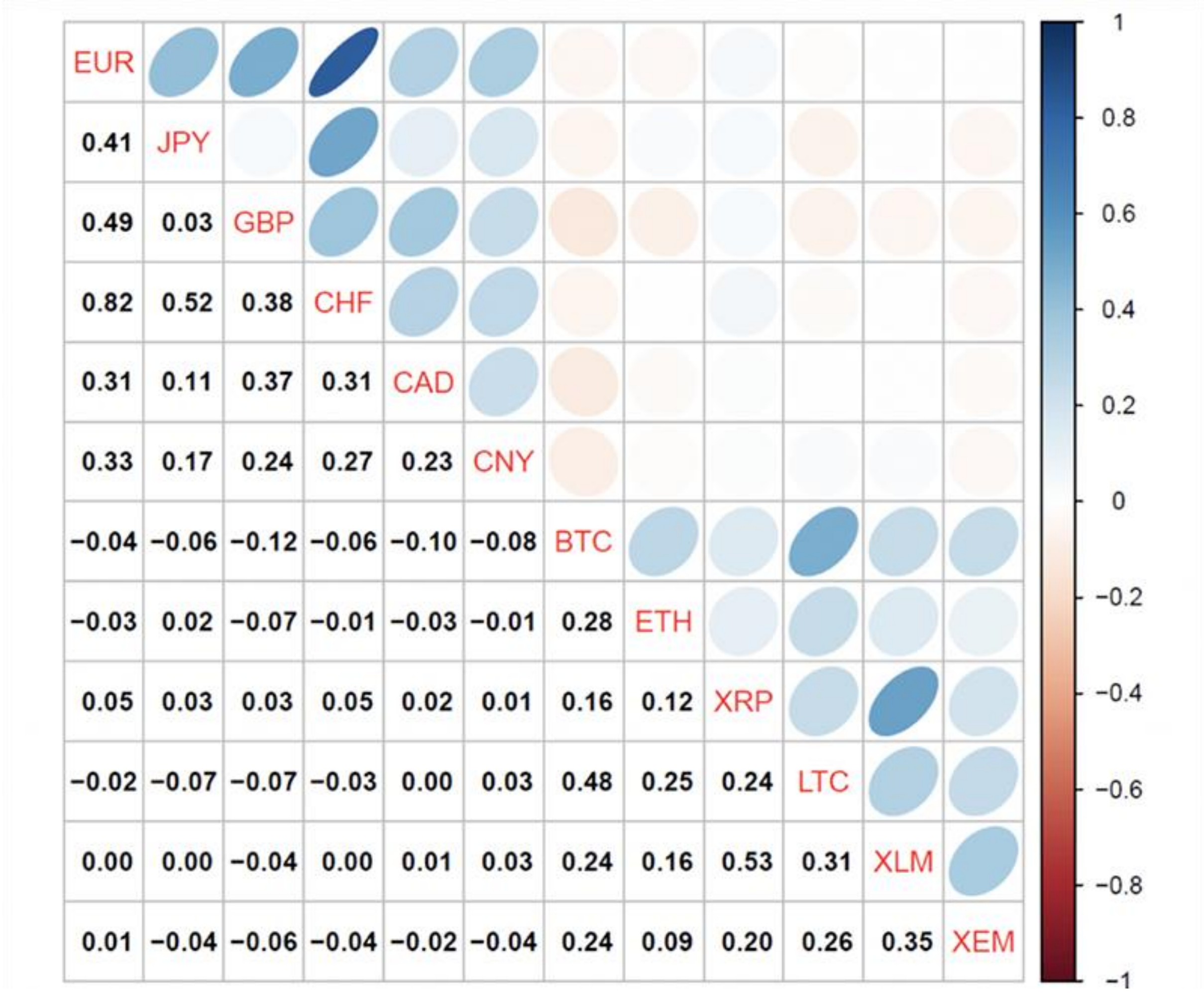


(Source: computed using data from https://www.investing.com)

[26] See Thomas and Morwitz (2009) for an analysis of this study.

If a customer discovers a decent way to keep and protect his bitcoins, he encounters the additional issue of handling the danger caused by the volatility of bitcoin. The Figure 1-8 above demonstrates the annualized volatility of the spot exchange rate of bitcoin to USD, calculated using the daily returns starting from January 2018 until May 2019. The chart provides a comparison of the volatility of the main fiat currencies in the world, Euro, Yen, British Pound, and Swiss Franc exchange rates as well as the gold price in USD and Bitcoin. The annualized volatility of Bitcoin for that period was above 60%, an order of magnitude much greater than the volatility of the other currencies, which does not go higher than 10%. Gold, which is considered a classical value-storing instrument in the world, had a volatility of returns equal to 10.17% during that period (relying on exchange rate in USD). To compare, the top 15 stocks by volume traded have volatilities ranging from 20% to 30%, and if even if we look at the riskiest stocks, they rarely display volatility as high as 50%. From the diagram above it can be concluded that keeping bitcoins even for a brief span of time is somewhat dangerous, incompatible with a currency that acts as a store of value and it significantly discredits a currency's capacity to work as an account unit.

Figure 1-9. Visualization of correlation matrix



(Source: computed using data from investing.com)

Figure 1-9 above shows the simple correlations of the percentage changes in daily exchange rates to USD of Bitcoin most widespread and stable fiat currencies, mentioned above. In addition, it shows correlation between each currency and the main altcoins price, nominated in USD. Friday-to-Monday returns for cryptocurrencies and fiat currencies exchange rates were used. As can be seen in the table most of the fiat currencies exhibit somewhat significant correlation with each other. There is a very strong positive correlation between main European currencies, with the Euro getting 0.82 correlation with the Swiss Franc and 0.49 correlation with the British Pound. The exchange rate of the Japenese Yen, Chinese Yuan and Canadian Dollar exhibit lower correlations, but still quite linked to American Dollar. By comparison, the daily returns of bitcoin-dollar show nearly zero linear relationship with the exchange rates of the other fiat currencies.

Bitcoin’s complete separation from other prominent international currencies and from gold seems telling. Macroeconomic events that cause similar impacts on the value of various currencies do not seem to affect bitcoin either positively or negatively. These measures are in line with those of previous empirical research (e.g., Yermack, 2015; Bouri et al., 2017a; Bouri et al., 2017b). The correlations between forex currencies and cryptocurrencies are practically zero; thus, cryptocurrencies could be viewed as a diversifier for foreign exchange investors (neglecting all other factors apart from the correlations).

1.3.4 Obstacles faced by bitcoin

For bitcoin to become established as a bona fide currency, its daily value will need to become more stable so that it can reliably serve as a store of value and as a unit of account in commercial markets. The excessive volatility shown in Figure 1-8 is more consistent with the behavior of a speculative investment than a currency. As described above, bitcoin also faces difficulties due to its decimal pricing of common household goods, the scarcity of merchants who accept it, and the cumbersome process of procuring bitcoins from a vendor/exchange, among other issues. The relatively high level of computer knowledge required for using bitcoins represents a further barrier to bitcoin's wide adoption.

Moreover, bitcoin's legitimacy as a currency should also hinge on its integration into the web of international payments and risk management transactions. Even though it is not issued by a sovereign state, bitcoin imparts risk to any business that accepts it for transactions, just like all other currencies. Major companies that deal in more than one currency, such as multinationals, attempt to hedge themselves against risks related to changes in those currencies' values. Several researches (Chan, 2018) suggest that no effective way exists to hedge bitcoin against the value of other currencies, although there exists a futures contract on bitcoin that has recently started trading on CBOE. Bitcoin transactions also are risky due to the absence of basic consumer protection, such as the provision of refunds that result from disputes between merchants and customers. While local laws may provide ground rules for resolving such disputes, because a government has no legal way to foreclose and take possession of bitcoins, it ultimately has little ability to step in and enforce its laws. Similar problems arise in attempting to secure consumer credit denomiated in bitcoin or to pledge bitcoins as a collateral for a consumer loan. Again due to its lack of affiliation to any sovereign, bitcoin is ill-suited for use in credit markets because no government can foreclose and seize it in the event of a default.

Bitcoin appears to suffer by being disconnected from the banking and payment systems of the U.S. and other countries. Most currencies are held and transferred through bank accounts, which in turn are protected by layers of regulation, deposit insurance, and international treaties. Without access to this infrastructure, Bitcoin is vulnerable to fraud and theft by skilled computer hackers. However, adherents of Bitcoin argue that bitcoin bypasses the well-known flaws in standard financial security systems, which have spawned epidemics of identity theft and related problems for ordinary customers of mainstream businesses.

Finally, bitcoin faces a long-term structural economic problem related to the absolute limit of 21 million units that can ever be issued, with no expansion possible of the bitcoin supply after the year 2140. If bitcoin becomes wildly successful and displaces sovereign fiat currencies, it would exert a deflationary force on the economy since the money supply would not increase with economic growth. This situation would require most workers to accept pay cuts every year, for instance, likely leading to political protests against the currency similar to those experienced in the U.S. during the Populist movement at the end of the 19th century. One can imagine a revival of William Jennings Bryan's 1896 "cross of gold" speech in the next century, updated with futuristic rhetoric about the economic tyranny of an uber-currency with an inflexible supply.

Chapter 2. Arbitrage in cryptocurrency markets

2.1 Literature review

Cryptocurrency, as a young asset form, provides a lot of opportunities for research, in particular, from econometric point of view. Adoption of this modern technology has increased quickly and activity in trading cryptocurrency has resulted in more than 200[27] (as of May 2019) extremely fragmented, mostly uncontrolled, cryptocurrency exchanges that behave more like brokers and dealers rather than exchanges in traditional markets (Hansen, 2018). Also, there is significant amount of "off chain trading". This could be either interexchange trading or dark pool trading that could result in price jumps on a particular cryptocurrency exchange (Sharma, 2018). Various forms of digital currency are bought and sold at different prices on different exchanges. These simultaneous sources of data produce dynamic interconnections of high dimensions.

Ong et al. (2015) utilize social network information and discover four main factors linked to a cryptocurrency's market cap: 1) combined GitHub pull requests, 2) number of mergers, 3) number of active accounts, and 4) number of comments. The cryptocurrencies with highest market cap have the highest activity, unsurprisingly. This type of database is new to digital assets. Similar statistics, in the stock markets for instance, can be obtained from third-party data vendors or gathered from analytical reports and researches, and maybe news streams and management conference calls. The more various data sources become available for cryptocurrencies the more opportunities in the market.

Studies of trading patterns, herding impacts, and decision-making has begun on analyzing and forecasting sentiment. In conjunction with other machine learning methods, natural language processing methods enable professionals to construct measures of sentiment. Cretarola and Figa-Talamanca (2017) are proposing a confidence-based approach for pricing cryptocurrencies and derivatives on them, where price movements of different cryptocurrencies are partially driven by the confidence in the technology underlying. Aste (2018) investigates the present digital currencies dependencies and causal relationships. The research involves analysis of both price-volume data and social network sentiment for nearly 2000 of cryptocurrencies traded in 2018. His findings reveal a complicated interrelationship framework in which prices and news sentiment both instantly and with lead-lag relationships affect each other across various currencies.

Nasekin and Chen (2019) used a cryptocurrency-specific terminology suggested in Chen et al. (2018b) and machine learning techniques to study market sentiment on digital currencies. By

---

[27] See coinmarketcap.com/exchanges

studying word embedding, accounting for situation-specific data and word resemblance, they utilize natural language processing techniques for the formation of sentence-level grouping and market sentiment index. They claim that the developed market sentiment indices are value-relevant for cryptocurrency market indices in terms of their return and volatility forecasting. Pagnotto and Buraschi (2018) also discuss cryptocurrencies valuation and characterize bitcoin demand by the current hash-rate and demonstrate that a fixed-point issue solves the equilibrium price. They discover that "price/hash-rate spirals" counteract shocks in demand and supply. Schilling and Uhlig (2018) are analyzing the USD-Bitcoin coexistence and rivalry. They evaluate the price development of Bitcoin and the relationship between the price of Bitcoin and the monetary policy that addresses the USD and derive a fundamental price equation that, in its basic form, shows that price process of Bitcoin forms a martingale.

Chaim and Laurini (2019) evaluate bitcoin's daily yields from January 2015 to March 2018 to empirically explore the hypothesis of speculative price bubble. Bitcoin yields have features that one might expect from a speculative bubble: it is highly volatile, negatively skewed and exhibit high kurtosis. (Camerer, 1989). They claim that BTC-USD prices being a speculative bubble is possible, but the evidence is incomplete. On the other hand, Henry and Irrera (2017) claim that there is bubble-like pattern in digital currencies. Recent Hafner (2018) study expands traditional bubble testing to the situation of time changing volatility. Dong et al. (2018) explore the positive and negative results of a cryptocurrency model as costly and risky bubbles in an infinite-horizon manufacturing economy with incomplete markets with the following framework for bitcoin: 1) huge price fluctuations, 2) returns dynamics is considerably vulnerable to market sentiment and political issues, and 3) the market displays various cyclical characteristics. Though, their quantitative findings depend strongly on the severity of the distortion of the market, i.e. the intervention of a government authority in the specified market, which in turn dictates the size of the speculative bitcoin bubble.

Bitcoin utilizes an especially electricity-intensive technique that increases environmental issues, particularly with the prevalence of coal-fired power plants based on bitcoin mining in China (Hileman and Rauchs, 2017). Cong et al. (2018a) demonstrate that mining pools considerably intensify the consumption of electricity for proof-of-work blockchains. As of April 2018, total amount of energy alone for mining bitcoin surpassed 60 TWh, approximately the annual energy consumed by Switzerland (Lee, 2018). Mishra et al. (2018) explore how the bitcoin mining procedure affects miners' computational power requirements and show that the mining algorithm as well as the quantity of transactions boost computing resource requirements, which in turn increases energy consumption. They eventually contend about resource demands from both computer hardware and energy consumption needs that the future development of the bitcoin protocol and the use of bitcoin as a currency might be dubious. Blockchain technology provides several innovative research possibilities related to the environment (Hayes, 2017; Pop et al., 2018).

There are hundreds of exchanges of cryptocurrencies worldwide. Binance, Kraken, Bitfinex, Coinbase, itBit, Gemini, Bitflyer, Bitstamp, Bittrex and Poloniex are probably the most famous ones. All of these crypto exchanges have their own particular characteristics. Kraken claims to be the biggest bitcoin to EUR exchange by trading volume and market liquidity and it has a partnership with the German BaFin-regulated Fidor bank, developing first cryptocurrency bank. In comparison, Shapeshift is an exchange that enables trading to take place without having to sign up for an account. Gemini, which is a fully regulated and licensed United States cryptocurrency exchange, meets its capital requirements by putting all USD accounts in an FDIC-insured bank. Coinmarketcap lists 245 exchanges as of May 2019, but its reported trading volume numbers are questionable. A latest submission to the SEC (2019) claims that 95% of bitcoin's trading volumes are fake. The study lists 10 exchanges with non-manipulated trading volumes (out of 81 present in the report), which are Binance, itBit, Bitfinex, Kraken, Coinbase, Bitflyer, Bitstamp, Gemini, Bittrex and Poloniex.

Because of the big amount of exchanges and growing amount of cryptocurrencies, price differences among exchanges naturally occur. The low regulatory standard and emotion-driven prices create price differences much bigger than in other financial markets, such as stocks, bonds, FX, etc. Though, some of the price inconsistencies may not be actual if one operates outside the ten exchanges with reliable trading volume. By using a theoretical model and an empirical approach, Bistarelli et al. (2019) demonstrate that arbitrage opportunities are still feasible by trading on various exchanges (Cretarola et al., 2017). Their methodology complements other theoretical research on bitcoin arbitrage such as Barker (2017) or Pieters and Vivanco (2015), where authors are studying triangular arbitrage with bitcoin, i.e. purchasing bitcoin in USD and selling it in EUR.

Makarov and Schoar (2018) observe big recurring opportunities for arbitrage trading across exchanges in cryptocurrencies prices relative to fiat currencies. These opportunities are often open for several days/weeks, and interestingly even in the presence of significant volumes traded and liquidity prices exhibit variation across exchanges. Krueckeberg and Scholz (2018) are conducting further studies on arbitrage trading in the Bitcoin markets. Bistarelli et al. (2018) demonstrate that arbitrage trading approaches on cryptocurrencies are profitable because the short-term rates of the exchanges differ. In fact, exchange fragmentation is perfect for high-frequency trading bots. Bloomberg (2017) claims that Chinese high-frequency traders have used trading algorithms which find mispricings and arbitrage opportunities across China's many exchanges. However, China prohibited all exchanges for cryptocurrency later in 2017.

Hautsch et al. (2018) observed that consensus protocols confront traders with unpredictable waiting times until the cryptocurrency transfer is completed. This method of settlement exposes arbitrage traders to price risk and limits arbitrage opportunities. Under general assumptions, they derive theoretical arbitrage limits and demonstrate that they rise with predicted latency, latency variability, spot volatility, and risk aversion by using high-frequency bitcoin price data. They state that distributed systems settlement induces non-trivial frictions influencing market efficiency and price formation.

2.2 Liquidity estimation

The main data for this research is minute timeframe prices data obtained from different sources, among which Enigma database,[28] using its specialized Python library, called Catalyst, and data obtained directly from exchanges using APIs. Catalyst is an algorithmic trading library for crypto assets. It has several features among which access to historical data of daily and minute resolution, possibility to backtest trading strategies, and provision of analytical tools regarding a particular strategy's performance. Moreover, Catalyst supports live trading of cryptocurrencies on 4 exchanges, Bittrex, Binance, Bitfinex and Poloniex. Catalyst covers more than 130 exchanges, which, besides bitcoin provide trading data on other coins, such as Ethereum, Ripple, Bitcoin Cash, Litecoin, etc. Catalyst collects the data by querying APIs provided by exchanges. Generally, their data is viewed as reliable in this paper, but where available, confirmations of the quality of the data by comparing them with the data reported by the exchanges to bitcoincharts.com are provided. Also, data from an independent resource, coinmarketcap.com is used. It aggregates information on trading volumes by exchange and by coin.

The variables contained in the data are essentially open, high, low, close prices of the cryptocurrencies for a given timeframe and volume traded in this period on the given exchange. For shorter time periods (from February 2019 to May 2019) more frequent data was used: limit orderbook snapshots and its updates parsed directly from exchanges. Usually frequency of updates of orderbooks by exchanges lies somewhere between several times in one second or once in a second for some particular exchanges.

When it comes to choosing cryptocurrency exchanges, unfortunately, anyone, who would like to do a research or just trade cryptocurrencies, faces several problems. Wide audience of speculators and projects seem to care only about one metric regarding exchanges which is trading volume. Therefore, exchanges tend to overstate their reported trading volumes. Most of the cryptocurrency exchanges execute trades in a centralized database and can therefore trade the same coins back and forth between two bot accounts, not providing any real liquidity into the market.

[28] https://enigma.co/catalyst

Thus, people or companies who want to trade cryptocurrencies receive biased picture of the real liquidity on exchanges, if they sort them just by volume traded (as it is done on the Coinmarketcap.com), and therefore, make wrong choices. Users join an exchange and think that they can buy or sell rather big amounts of cryptocurrency without affecting prices, which is apparently not the case, because there might be only a few bots trading on the platform. Of course, volume traded is not a bad measure overall, but it should be used together with the other metrics.

2.2.1 Liquidity

What people care about the most is probably if the exchange has enough liquidity in its order book rather than if it has high volume traded. Because only in this case they can execute their orders without moving the price by, say, 10%. Thus, a clearer illustration of the liquidity of an exchange/trading pair should be given by contrasting stocks and trading pairs by orderbook depth rather than by trading volume. A recursive strategy could be selected to best assess the trading liquidity supplied by the orderbook. For example, if you sell an equivalent of $10k using a market order on a specific trading pair that you are interested in, you could evaluate how much the price of this asset would change.

Further in this section it is done for top 10 exchanges by volume, as reported by Coinmarketcap, using an Ethereum to USD and Bitcoin to USD pairs. The data from the orderbooks for abovementioned trading pair has been collected on the exchanges, and the percentage change in price in case of selling different amounts of the asset has been estimated. Unfortunately, many of them in their customer interface provided only a tiny overview of their orderbook by displaying only the best bids and asks. However, it becomes clear that the orderbook liquidity does not necessarily follow trading volume.

For instance, OEX and ZBG exchanges, listed on places 1 and 4 of CoinMarketCap's list of Ethereum's top markets by trading volume on 14-05-2019, did not provide sufficient liquidity to supply even 50 ETH at the present exchange cost. This is an equivalent of about USD 12 000 at the current market price. If you sold 50 ETH using market order on ZBG, you order would have been matched with buying orders 10 percent below the initial price. If you placed an order on OEX, selling 50 ETH, the market exchange rate would have fallen by 33%. BitForex and Bibox offered higher liquidity for low sales amounts, so that it would be possible to place a selling order of 50 ETH and get actual price close to the global market price. But if you tried the same with 500 ETH, buying orders would have come 80-99% below the real exchange quote. Huobi, Bitfinex, Okex and Binance offered the best liquidity among the exchanges that provide trading in Ethereum to US Dollar pair. If you placed an order one of these exchanges, you could have sold more than 1000 ETH without dropping market price even 1% lower than the initial one.

This strategy is a little difficult to implement, so another attempt to assess orderbook liquidity was made. Its essence is evaluating what quantity of a particular coin can be purchased/sold for the current market price via simple market order. However, cryptocurrency markets are generally quite volatile and not very liquid, that's why let's assume that actual market price of a token is the price it is traded at +-1%. Thus, the objective becomes to assess the quantity of coins that can be purchased/sold without reaching sell/buy orders 1% below/above the actual market price. It should be stated that these figures change several times a second. Nevertheless, this approach still should provide valuable insights into which exchanges' orderbooks offer sufficient liquidity and which ones do not.

Further, orderbooks of 75 digital exchanges, which had a 24-hour trading volume higher than USD 35 million, were analyzed. The exact trading pairs taken into consideration are ETH/USD, ETH/USDT and ETH/BTC. The analysis was done on 15-05-2019. At that time, one could have sold 1000 ETH by using simple market order on 19 of these 75 exchanges without touching buy orders 1% below the current market price. Table 2-1 below summarizes the analysis performed and is structured as follows: those 19 exchanges are sorted by the amount of ETH that could have been sold on the exchange using purely market orders without moving current market price more than 1% down. The reported traded volume is also added to the table to illustrate the lack of a relationship between real orderbook liquidity and the position of the exchange in the Coinmarketcap top chart. It can be noticed that the exchanges that have been operating for a long time are also the ones that provide the best liquidity to the users. Bitfinex, Bitstamp, HitBTC, Kraken, Coinbase Pro are well-known exchanges which exist for relatively long period of time (for cryptocurrency industry), and are included in the top 5 group, while newer exchanges, that claim to have giant trading volumes, fail to make that list.

Table 2-1. Cryptocurrency exchanges limit orderbook liquidity assessment (Ethereum pairs)

Exchange	Pair traded	Amount of ETH can be sold, without affecting market price more than 1%	Amount of ETH converted to USD at current exchange rate, thousand USD	24h volume traded reported by exchange, million USD
Bitfinex	ETH/Tether	7787	1946.8	83.5
Bitstamp	ETH/USD	6180	1545.0	7.5
HitBTC	ETH/BTC	4343	1085.8	9.7
Kraken	ETH/EUR	3917	979.3	24.7
Coinbase Pro	ETH/USD	2930	732.5	33.3
Poloniex	ETH/BTC	2764	691.0	5.1
Gemini	ETH/USD	2598	649.5	6.9
Huobi	ETH/Tether	2114	528.5	133.0
Bitrue	ETH/Tether	1955	488.8	1.1
Bithumb	ETH/KRW	1760	440.0	83.5
Binance	ETH/Tether	1650	412.5	133.5
IDAX	ETH/Tether	1644	411.0	45.0
Bittrex/UPBit	ETH/BTC	1426	356.5	3.2
Coinone	ETH/KRW	1407	351.8	9.8
Okex	ETH/Tether	1078	269.5	146.9
Coinhub	ETH/BTC	1069	267.3	7.1
Coineal	ETH/Tether	1066	266.5	96.8
Topbtc	ETH/CNY	1030	257.5	26.5

(source: Calculations based on limit order book data provided by exchanges)

The same analysis was undertaken for Bitcoin to fiat and Bitcoin to stablecoins trading pairs. At the time of the analysis one could have placed 100 BTC order on 15 out of 104 exchanges, without moving the exchange rate higher than 1%. Their ranking is presented in the table below. It probably wouldn't be a surprise to anyone that overall liquidity in Bitcoin instruments is higher than in Ethereum instruments, besides this fact, the analysis on Bitcoin provides approximately the same picture. There are some rather "old", well-known exchanges, which provide relatively good liquidity, and there are a lot of new exchanges which fabricate their volume traded reports.

It is interesting also to estimate how the assessed limit order book depth correspond to the trading volume recorded on these 100+ exchanges. The results indicate that there are three groups of exchanges: some don't seem to inflate their trading volumes, some of them inflate volumes at some extent, but they also have rather high actual trading activity, and some exchanges have mostly artificial volumes of trading.

According to the data gathered, the exchanges with the greatest liquidity in Bitcoin trading pairs produced mostly a double-digit million USD trading volume. For instance, $41 million for Bitfinex, $20 million for Kraken, $28 million for Bitstamp and $61 million for Bithumb. An interesting fact is that the average daily volume of trading in the assessed trading pairs across all 104 exchanges in the long list was $44 million. Therefore, an average exchange in the long list had more daily trading volume reported than Coinbase Pro or Bitfinex.

Correlation analysis was conducted between the trading volume and order book depth, to further evaluate the link between them. A positive correlation between these two variables should be expected as at first, they seem to be interconnected. For the cryptocurrency markets this is not the case though. The correlation coefficient calculated on the gathered data is equal 0.02. It demonstrates that both variables are totally unrelated and there is no even a weak linear relationship between them.

Assume that the trading volume reported by Coinbase Pro, Kraken and Gemini is real. All of these exchanges have filled limit order books, but don't report that high trading volumes. For instance, in the investigated data set, the limit order book depth of Gemini, Huobi and Liquid was within the same range (about 200 BTC could be bought/sold without changing the price by more than 1 %), but the reported trading volume on Huobi and Liquid (approximately 100 million USD) was 10x higher than on Gemini (app. 10 million USD). In addition, the reported daily trading volume on Binance, for instance, was 2.5x of that on Coinbase Pro, although Binance had only 22% of the limit orderbook liquidity of Coinbase Pro.

Of course, these results do not imply that all the trading volume on Binance, Huobi and Liquid for example, is falsified. Perhaps, trading volume on these exchanges is partially inflated with artificial volume, but it's not necessary. The exchanges are interested to overstate their actual trading volumes with artificial to attract investors. It might be that their entire trading volume is genuine and created by investors or traders who simply prefer to use these markets to the others, because of some good features that these exchanges provide, for example favorable trading platform, speed or simply low fees. It is harder to prove that the reported trading volume on the exchange is falsificated, if it provides good liquidity.

But obviously, volume reported by some cryptocurrency exchanges, which claim to achieve several million dollars in daily trading volume, but cannot even consume an order of few BTCs without collapsing the exchange rate, is most likely false. Good illustration to that would be that 16 of the 104 exchanges examined reported a daily volume of trading in a Bitcoin to Fiat or Bitcoin to Stablecoin trading pairs of more than USD 1 million, but you couldn't even sell one Bitcoin without a drop in price of more than 1%. Obviously, these exchanges are simply trying to trick clients and should be removed from rankings.

Table 2-2. Cryptocurrency exchanges limit orderbook liquidity assessment (Bitcoin pairs)

Exchange	Pair traded	Amount of BTC can be sold, without affecting market price more than 1%	Amount of BTC converted to USD at current rate, million USD	24h volume traded reported by exchange, million USD
Bitfinex	BTC/Tether	727	5.5	40.8
Coinbase Pro	BTC/USD	551	4.1	34.1
HitBTC	BTC/Tether	486	3.6	53.1
Kraken	BTC/USD	430	3.2	20.3
Bitstamp	BTC/USD	343	2.6	27.9
Gemini	BTC/USD	227	1.7	10.8
Huobi	BTC/Tether	219	1.6	110.6
Liquid	BTC/JPY	209	1.6	105.0
Bitflyer	BTC/IDR	162	1.2	13.2
Binance	BTC/Tether	121	0.9	91.4
Bithumb	BTC/KRW	119	0.9	60.9
Coincheck	BTC/JPY	114	0.9	6.6
Indodax	BTC/UIDR	111	0.8	1.3
Idax	BTC/Tether	110	0.8	184.4
Coinall	BTC/Tether	109	0.8	55.4

(source: Calculations based on data provided by particular exchanges; volume traded data is taken from coinpaprika.com)

However, there are several limitations to this “true liquidity estimation” approach. Probably the most important of them is that dataset used it limited. The liquidity assessment is based on a limit orderbook snapshot in a single day. But anyway, it doesn’t make the results insignificant, because if an exchange couldn’t provide liquidity in the snapshot on a random day, then it is very likely that it won’t be able to provide liquidity in general. Sadly, there is no centralized database regarding cryptocurrency exchanges’ orderbooks and it makes the collection of the data quite hard. Also, not all exchanges provide access to their orderbooks, and it automatically excludes them from this type of analysis.

Another limitation is that the volume of Ethereum or Bitcoin that can be sold, stated in the tables above, comes from the orderbooks reported by the exchanges themselves. These entries in an orderbook may also be falsified by exchanges. They could make orders vanish at the same moment that a trader puts an order and bring up these orders again after the trade was executed. It would definitely harm the credibility of these exchanges, thus, it's not obvious for how long such exchanges are able to maintain their business with fake orderbooks, because any more or less experienced trader will notice it very fast.

Of course cryptocurrency exchanges from the top of the Table 2-1 and Table 2-2 do not necessarily constitute the finest solutions for trading cryptocurrencies, simply because they provide liquidity. For previous years, a lot of debate has surrounded Tether and Bitfinex; HitBTC has lately been in the news headlines for not enabling their customers to withdraw their money. But these tables and the underlying orderbook depth measure should be seen as an improvement towards ranking of exchanges simply by volume reported, since it is often interpreted as liquidity, but in the present condition of the cryptocurrency markets, the volume traded and real liquidity reported on an exchange are completely unrelated.

This quick analysis should be treated as a general overview of which exchanges should be avoided when trading, but for studying arbitrage opportunities it may not be sufficient and more detailed research is required. The obtained list of potentially “good” exchanges will be further analyzed in the next part.

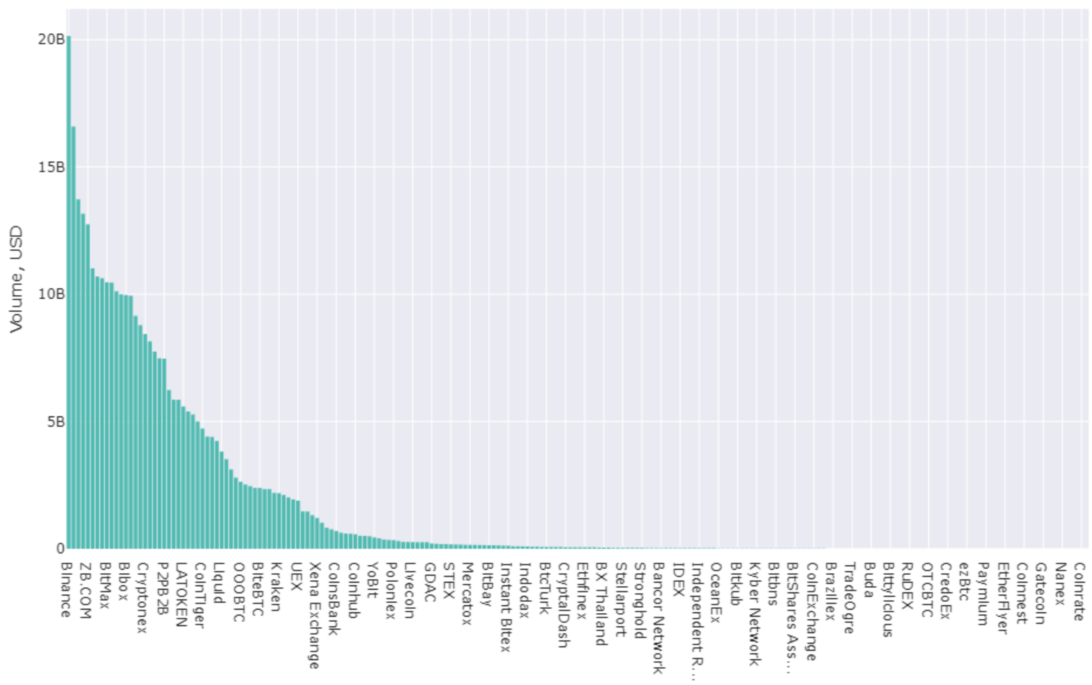
2.2.2 Exchanges

The motivation for this part of research was to test specific exchanges from the long-list, obtained in the previous part, which were several times mentioned to be suspicious[29] in terms of falsifying their volume traded, botting etc; and maybe add to the list some more exchanges with good test performance if any. At the time of study, Coinmarketcap reported 245 exchanges, although the real number was a bit higher. Of 245, there were 27 exchanges with 0 adjusted trading volume[30] and $93.5 billion (in sum) reported volume (21 percent of complete traded quantity recorded). Almost all these exchanges either provide future contracts trading or have zero trading fees (BitMEX, Bithumb, etc). This study includes 14 digital exchanges from Coinmarketcap rating of top exchanged ranked by adjusted volume traded for the last 30 days, among them there are some exchanges chosen in the previous section, for which it was possible to get high frequency data. In May 2019, these exchanges corresponded to 34% of the total adjusted volume traded. Within each of these exchanges the data for the most liquid cryptocurrencies is analyzed.

[29] There are numerous investigations on this topic. See, for instance https://www.bti.live/december-2018-rankings or https://blog.cer.live/investigations/

[30] Adjusted volume is a metric reported by Coinmarketcap defined as volume from spot markets excluding markets with no fees and transaction mining

Figure 2-1. Trading volume in May 2019 of all exchanges with non-zero adjusted volumes



(Source: coinmarketcap.com)

Table 2-3. Summary of the data used

	Volume	% of total volume
Total 30-day volume (%)	$441,557,469,457	100%
Excluded (%)	$93,511,692,648	21%
Adjusted (%)	$348,045,776,809	79%
Researched (%)	$118,802,880,621	27% (34% of adj.)

(Source: coinmarketcap.com)

2.2.3 Methodology

It is hard to overestimate the importance of liquidity and trading volume indicators a trader or investor. Non falsified, true trading volume is an insight into the intrinsic value of the asset. It can be used as a leading indicator of future market price of an asset by traders. In theory, market price of a bitcoin or any other cryptocurrency (or stock, commodity, etc.) is a result of interaction of supply and demand. Demand, in turn, can be driven by different factors. One of them is utility, or in other words how many people will use it in future. Another factor affecting demand is investment value of an asset, or how many people will buy it, hoping for future cash flows or market price increase. In short, the demand should reflect people’s interest in the underlying technology, its future prospects and its mass adoption among investors. The higher the interest, the higher the volume. Further, the trading volume is a function of orderbook liquidity among other factors. To formulate it mathematically:

$P = f(V, p_1, \dots, p_n)$, where P is market price, V is the volume, and p_i are some other parameters;

$V = f(L, k_1, \dots, k_n)$, where L is orderbook liquidity, k_i are some other parameters.

Things considered, understanding of an asset's true liquidity and volume traded is necessary to make a balanced investment decision. Although traditional economic markets have a set of techniques and regulations to avoid distortion of market data, cryptocurrency transactions remain susceptible to fraudulent activities that mislead shareholders and traders. In controlled financial markets, there are several fraudulent trading activities that are forbidden. Wash trading and disruptive activities like flipping, spoofing, quote stuffing are the most prevalent examples. This part of the research is an attempt to detect exchanges on wash trading as the cryptocurrencies world's most prevalent fraud.

The three key techniques — from the simplest to execute (and identify) to the hardest:

1. in-spread trades without limit orders;
2. in-spread trades with short-lived limit orders;
3. trades near bid and ask prices using short-lived limit orders.

The first of these practices is the most used one, especially among fresh new exchanges, and its mostly because it is easy to implement and almost riskless. An exchange that implements this method claims a trade while the order book is completely unchanged. This practice bears no risk because, even for fractions of seconds, there are no limit orders in the order book, so other traders cannot make counter orders to them. Only the exchange itself can sustain this kind of trading activity.

In-spread trades with short-lived limit orders. Anyone can create a limit order with price between bid and ask prices. Such activities make the spread on the instrument smaller and add liquidity to the market. This manipulative technique, on the other hand, implies that only orders with very short lifetime in the order book (milliseconds), so that a human is not only unable to meet them but also to just notice them in the order book. Not all of the short-lived limit orders are malicious, practically speaking. There are algorithmic traders and arbitrageurs who track all oncoming orders and rapidly meet those which bring them the chance to trade. But there is no justification for the dominance of transactions induced by such orders.

Trades near bid-ask caused by short-lived limit orders. Advanced wash trading algorithms are adjusted to generate sell orders close to the best bid offer and vice versa buy orders near the best ask to make the trading flow look more natural. But there is no intention of wash traders to genuinely trade with anyone. They set limit orders for fractions of a second and then immediately cover them with counter orders in order to minimize this risk. Naturally, a limit order of 1 BTC hit after, say, 50 milliseconds by a market order of 1 BTC can readily be categorized as a wash trade. The fakers are therefore trying to add distortion to their operations:

- divide limit order into several orders of different size;
- put aggressive orders of bigger or less size than the size of the limit;
- introduce noise to life of duration of limit orders;
- create aggressive orders with different lifetime (can be the case if the size of an aggressive order is bigger than the size of a resting one, then the not-filled part of the aggressive order becomes new resting order).

The algorithm calculates 3 metrics of liquidity for the most liquid trading instruments on every analyzed exchange:

- Handy liquidity defined as the cumulative volume in an order book at prices further from the mid-market (average of bid and ask) price by 0.5% or less. The higher the liquidity of an asset, the higher should be the handy liquidity.
- Bid ask spread, defined as:

$$\frac{(best\ ask - best\ bid)}{\frac{best\ ask\ +\ best\ bid}{2}} * 100$$

The higher the liquidity, the narrower the bid ask spread.

- Bid ask spread by 10 BTC, calculated as:

$$\frac{(ask - \ bid)}{\frac{(ask + bid)}{2}} * 100,$$

where ask and bid are weighted by aggregated volumes of 10 BTC average prices. This metric should be lower for liquid assets.

The algorithm examines market data collected using APIs of every exchange in the list. Data used is order books and trades, some of the exchanges also provide the updates of order books. The rules by which the system believes that transaction is suspicious (wash trade):

1. If a trade takes place at a price lower than the best bid or higher than the best ask price, it is classified as “out-of-spread trade”;
2. If a trade takes place at a price lower than the best ask or higher than the best bid price, it is classified as “in-spread trade”;
3. If trade happens at bid or ask price, it is referred to as normal trade.

Sum of the out-of-spread and in-spread trading volumes combined make up total artificial trading volume. The overall statistics are calculated as follows:

1. Share of total artificial volume and of total artificial trades (quantity) per trading pair per each exchange examined is calculated;
2. Arithmetic mean of the shares of total artificial trading volume and total artificial quantity of trades across all selected trading pairs of all exchanges examined is calculated.

The algorithm can identify fake trades when there is no wash trading performed (Type I error) due to probable problems with data gathering and time synchronization. Coinbase Pro was regarded as a benchmark for 100% organic trading flow, although the algorithm discovered minor anomalies.

Table 2-4. Detection of in-spread and out-of-spread trading activities for Coinbase Pro exchange

Coinbase Pro	BTC/USD
Fake In-spread Trades, %	0.00
Fake In-spread Volume, %	0.00
Fake Out-of-spread Trades, %	1.26
Fake Out-of-spread Volume, %	1.19
Fake Total Trades, %	1.26
Fake Total Volume, %	1.19

(Source: computed using data provided by exchanges)

Several chosen trading pairs have been examined in this research, as stated above, all of them are labeled in the Table 2-5 below. The analysis is conducted on randomly chosen samples of data (in February, March and May 2019).

Table 2-5. Trading pairs and exchanges analyzed

	BTCUSDT	EOSUSDT	ETHUSDT	LTCUSDT	XRPUSDT
Lbank	●		●	●	
bitmart	●	●	●		●
BW	●	●	●	●	●
BitZ	●	●	●	●	●
CoinTiger	●	●	●	●	
HitBTC	●	●	●	●	●
ZB	●	●	●	●	●
IDAX	●	●	●	●	●
Huobi	●	●	●	●	●
Coinbene	●	●	●	●	●
OKEX	●	●	●	●	●

The study is focused on raw data provided by exchanges: transactions and order books. There are three types of market data protocols:

1. Full order log, which provides the most details among other types of protocols, including the history of every placed order—time it was placed, execution, cancellation;

2. Level 2 updates provides all the updates (snapshots) of the order books up to N (which depends on particular exchange) price levels; in other words, it gives changes in aggregate sizes at a particular price level;
3. Level 2 snapshots provide an order book state at the time of a snapshot; they may be configured by an exchange (e.g., Bittrex sends marked data updates every second using WebSocket) or limited by the protocol and rate constraints (e.g., Yobit constraints the quantity of requests for an order book snapshot to 100 per minute).

Not all of the exchanges provide timestamps of their orderbooks. If it is the case, local server time (time when either an orderbook or trade was received by the server) is used for both trades and orderbooks. The summary of market data protocols is provided in a table below.

Table 2-6. Market data protocols summary

Exchange	Protocol	Type	Time Book	Time Trade	Used Time
Bitmart	WS	Updates	n/a	1 ms	Local
Bit-Z	Unofficial WS	Updates	1 ms	1 sec	Exchange
BW	Unofficial WS	Updates	n/a	1 sec	Local
Coinbene	Unofficial WS	Updates	1 ms	1 sec	Exchange
Cointiger	Unofficial WS	Snapshots	1 ms	1 ms	Exchange
HitBTC	WS	Updates	100 ms	1 ms	Exchange
HuobiGlobal	WS	Snapshots	1 ms	1 ms	Exchange
IDAX	Unofficial WS	Updates	1 ms	1 ms	Exchange
Lbank	WS	Snapshots	1 µs	1 µs	Exchange
Livecoin	WS	Updates	n/a	1 ms	Local
OKEx	WS	Updates	1 ms	1 sec	Exchange
ZB	WS	Snapshots	1 sec	1 sec	Exchange
Coinbase	WS	Updates	1 ms	1 ms	Exchange

(Source: CryptoIntegrity)

The more extensive data is used, the more comprehensive and accurate research can be done. Milliseconds timestamps of both orderbooks and trades (as recorded by an exchange) are needed to match order books and trades accurately. Bad information quality makes manipulations more difficult to identify. Hence, as a rule of thumb, the credibility of transactions with good API is greater.

2.2.4 Issues with the data

Huobi. While in the documentation of the Huobi's API it is written that the user will obtain updates in the limit order book after any change, it was noticed that actual market information updates are consolidated and transferred with frequency of 1 second. Firstly, it means the orders created and filled/canceled between two updates (within a second) with a lifetime of less than 1 second are absent from their updates. Moreover, the actual time of placing, filling or canceling an order is unknown. The actual pattern can be observed if we look at histogram of frequencies of updates.

HitBTC. While the API documentation says that each update will be sent, the same bias as with Huobi was observed. Market data requests from WebSocket API seems to have an update frequency of approximately 1 second. The consequences are also comparable to Huobi's. In addition, there was a problem with missing updates of order cancellation, which might have resulted to inaccurate bid ask spread and subsequently inaccurate statistics.

Figure 2-2. Distribution of Huobi market data updates frequency (in microseconds) for BTC-USDT pair.

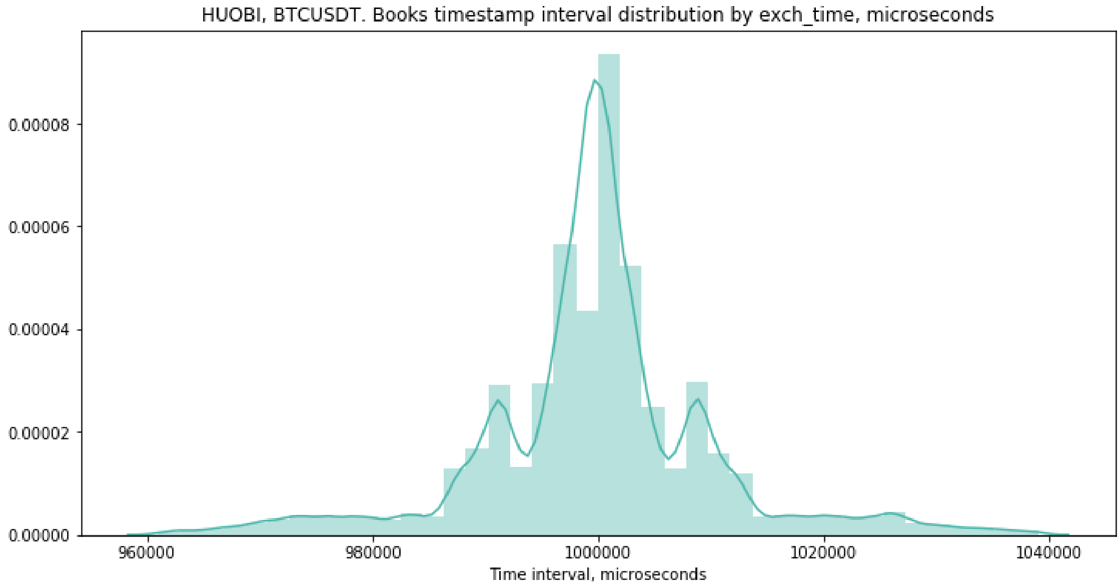


Figure 2-3. Distribution of HitBTC market data updates frequency (in microseconds) for BTC-USD pair.

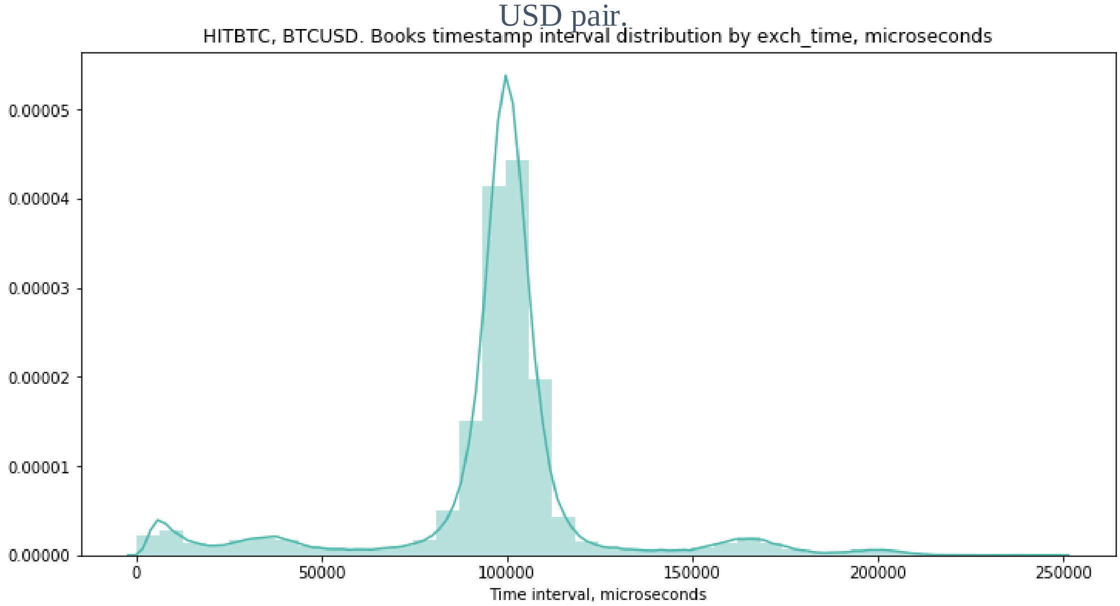
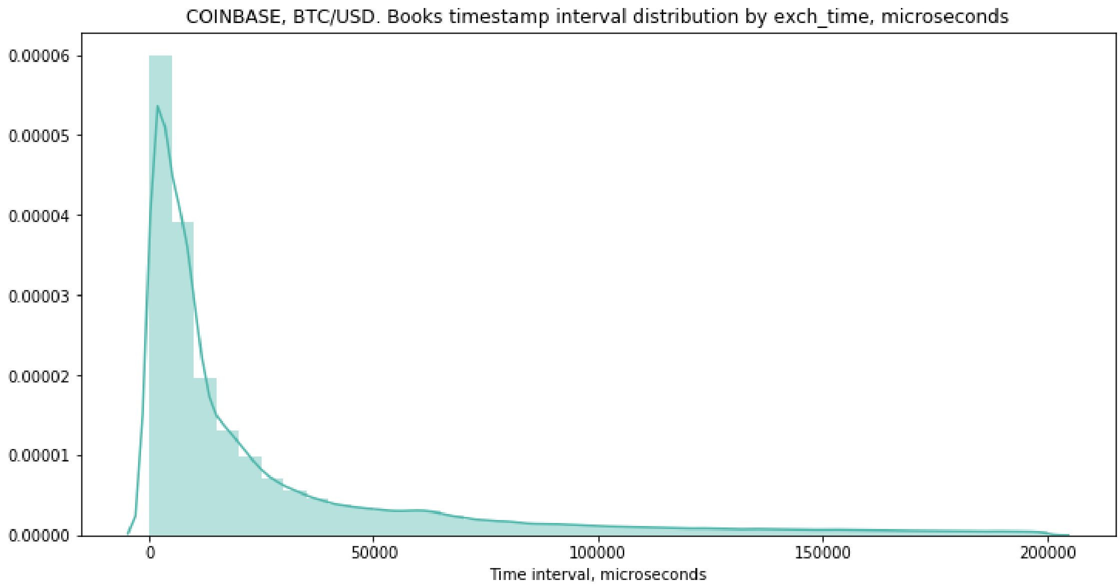


Figure 2-4. Distribution of frequency of updates for the benchmark exchange – Coinbase



(Source: computed using data provided by exchanges APIs)

From the distribution plots above it is clear that the data provided by APIs of HitBTC and Huobi cannot be trusted, because they provide not updates of order book, but rather orderbook snapshots with 1 second interval.

2.2.5 Results

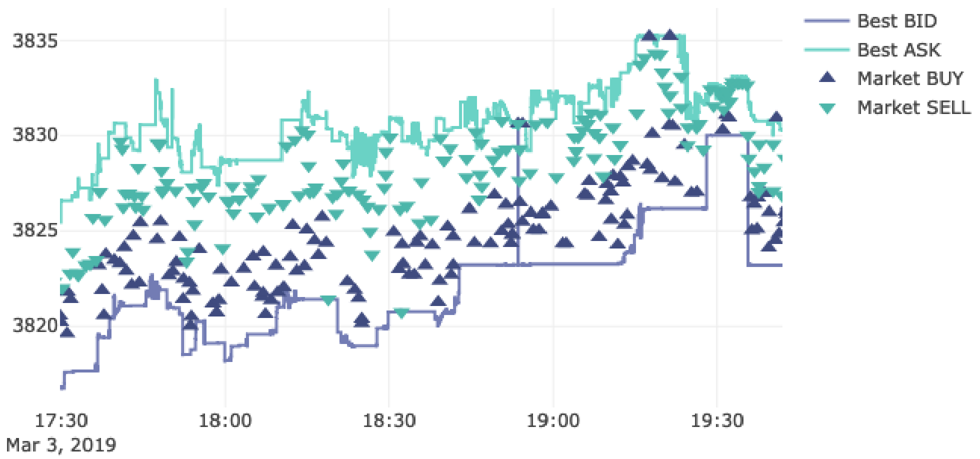
The graphs below demonstrate some cases of suspicious trading behavior with a brief remark on the chosen markets. The summary of all kinds of instruments which produce suspicious (wash) trading are presented in Table 2-7 below.

Table 2-7. Identified mechanisms of suspicious (wash) trading activity

	In-spread trades without limit orders	In-spread trades with short-lived limit orders	Trades near bid-ask caused by short-lived limit orders
OKEx	●		●
Bit-Z	●		
ZB.COM	●		
CoinBene		●	
LBank	●		
Huobi Global	n/a	n/a	n/a
BW	●		
HitBTC	n/a	n/a	n/a
IDAX	●	●	
BitMart	●	●	
CoinTiger	●		
Complexity			
	lower		*higher*

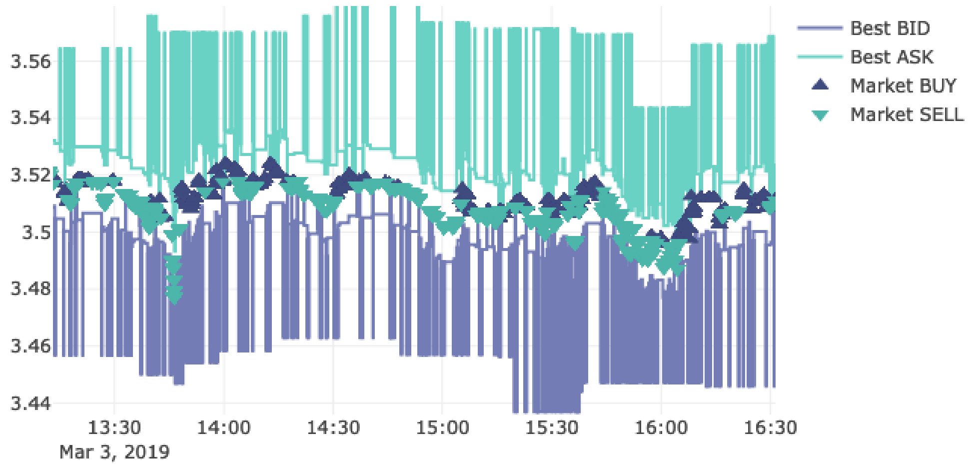
Some results are presented below, but most of the figures are skipped as they take a lot of space, and they can be found in the appendix (see Figure 1).

Figure 2-5. Identification of wash trading on LBank exchange (BTC_USDT pair)



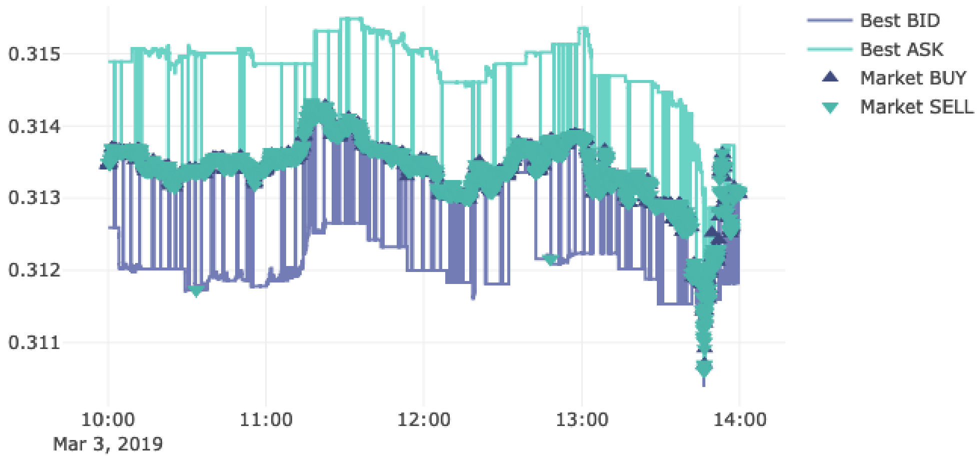
LBank orderbook data shows a lot of in spread-trades at random prices within the spread.

Figure 2-6. Identification of wash trading on Bitmart (EOS_USDT pair)



BitMart shows two patterns of wash trading, among which trades in-spread with limit orders and without limit orders.

Figure 2-7. Identification of wash trading on IDAX (XRP_USDT pair)



IDAX orderbook shows big spreads which is not a good sign in terms of liquidity, and the simplest wash trading activity - mid-market in-spread trades with short lifetime limit orders.

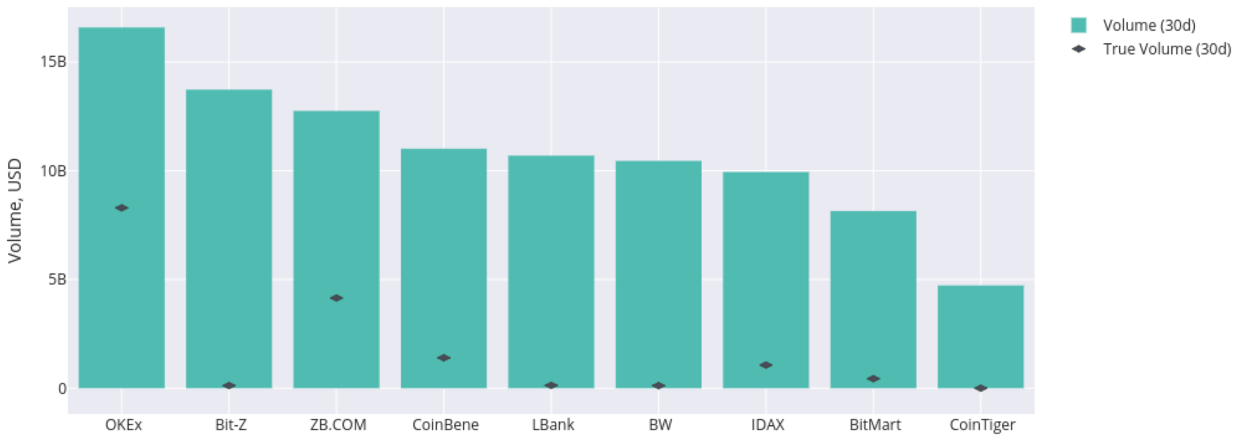
Table 2-8. Summary of liquidity by exchange. Green means high liquidity, red – low. Coinbase is regarded as a benchmark.

BTC/USDT	Handy liquidity mean, BTC	Spread 10 BTC, mean, %	Bid-ask spread, mean, %
Lbank	6.13	0.7	0.21
BitMart	24.96	0.25	0.08
BW	5.44	0.24	0.12
Bit-Z	5.94	0.27	0.11
CoinTiger	0.83	0.82	0.68
HitBTC*	4.65	0.02	0.09
LiveCoin	2.23	0.84	0.35
ZB	1.24	0.1	0.05
IDAX	9.24	0.19	0.1
Huobi*	5.86	0.02	0.01
Coinbene	3.72	2.18	0.21
OKEx	11.32	0.04	0.01
Coinbase	111.34	0.03	0.003

(note: HitBTC and Huobi are marked with a star, because there are issues with the data from these exchanges, thus, results regarding them are not truly reliable)

Interestingly, during the period studied BitMart had the greatest handy liquidity. All three measures had excellent estimated scores for OKEx. However, further assessment may be necessary as the liquidity presented may not be equivalent to the liquidity that is genuinely available. It is alleged that some fraudulent cryptocurrency exchanges display limit orders of special kind that other traders cannot meet. This is another limitation of this research.

Figure 2-8. Reported and true trading volume by exchange



Finally, Figure 2-8 represents the scale of trading volume manipulations by some cryptocurrency exchanges. This analysis has reduced the number of exchanges in the short list to 9, which are the most liquid and have significant trading volumes. In the next section, based on the data obtained from these exchanges, an analysis of arbitrage opportunities will be conducted.

2.3 Arbitrage opportunities

Arbitrage is "the simultaneous purchase and sale of the same, or essentially similar, security in two different markets for advantageously different prices".[31] The word should be read in a very general sense because both legs of the arbitrage can be a number of securities or derivatives. The definition essentially underlines two structural features of these transactions. First, the operations on the different instruments must be simultaneous. This characteristic is a direct consequence of the fact that arbitrage opportunities must be risk-free and are consequently short-lived. Secondly, the securities should be the same or essentially similar (believed to provide the same payoffs). This similarity in essence is an insurance contract, and by construction, it eliminates all and every risk beyond immediate execution. This idea is commonly described as the *law of one price,* and it is used extensively in the context of derivative pricing: If two portfolios have the same payoff at maturity, their price should be equal today, and if they are not, they can be expected to converge back toward equality.

The efficient market hypothesis states that "prices of securities fully reflect available information".[32] It basically comes in three flavors depending on the strength of the underlying assumptions about information dissemination. The weak form asserts that stock prices already reflect all information contained in the history of past prices. The semi-strong form asserts that stock prices already reflect all publicly available information. The strong form, which has been made illegal in some countries by laws against insider trading, states that stock prices reflect all relevant information including insider information.

The most important consequence of an efficient market is that most securities are, by definition, fairly priced by the market. There may be issues about liquidity to weaken this general statement, but certainly there cannot be any gross mispricing in an efficient market. In other words, arbitrage opportunities cannot exist. If they do, they may be the result of an optical illusion - typically non-simultaneous prices - or their magnitude is such that they do not create profitable opportunities when transaction costs are included - for example, the put-call parity on listed European options.

For the most part all these arguments hold in the traditional financial markets. However, there are still numerous opportunities that support profitable dedicated traders. For example, the

---

[31] [1] Zivi Bodie, Alex Kane, and Alan J. Marcus, *Investments,* 4th ed. (New York: Irwin/ McGraw-Hill, 1999), p. 307. The common term *arbitrage* is in fact used in many more situations. See, for example, Jochen E. M. Wilhelm, *Arbitrage Theory* (New York: Springer-Verlag, 1985), which makes a distinction between *arbitration* ("the search for the lowest cost in achieving a certain intended financial position"), *spreads* (simultaneous sale and purchase, i.e., *our* arbitrage), and *free lunches* ("turning from one combination of assets to another one which is equivalent but has a lower market price"). In this terminology, we are primarily concerned with spreads.

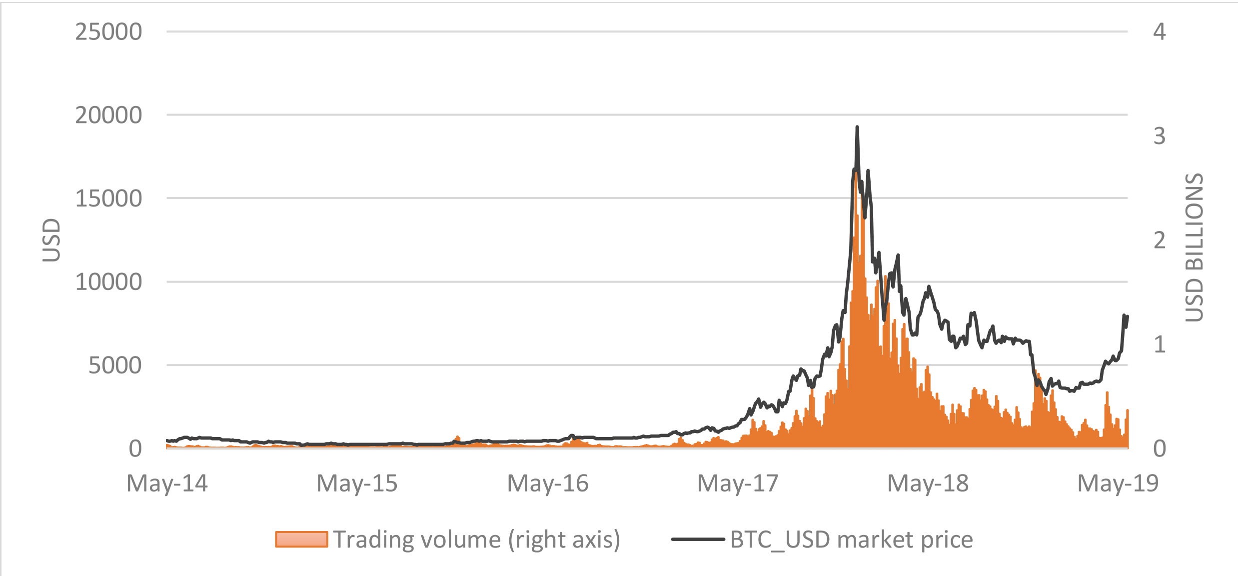
[32] Bodie, Kane, and Marcus, Investments, p. 933.

three main types of arbitrage - index arbitrage, risk arbitrage and pair trading - are actively and profitably traded in the world's largest markets. This apparent paradox has been studied extensively in the academic sphere.[33] However, for cryptocurrencies even the simplest, between-exchange arbitrage opportunities are present and quite large, as it will be shown further.

2.3.1 Prices and returns

The figure below shows bitcoin price and volume dynamics. Price data is volume-weighted weekly average price from Coinbase Pro and Bitstamp, which are the largest US based cryptocurrency exchanges, from May 2014 to May 2019. Data, gathered from Coinmarketcap and Bitinfocharts, shows the sharp increase in bitcoin-dollar exchange rate from January 2017 to January 2018, which in the common media has attracted a ton of publicity. At the beginning of 2017, it grew from less than \$1000 to nearly \$20,000, with a particularly fast rise after November 2017. By the beginning of February 2018, the price dropped back to just under \$10,000, and continued falling after, but not that fast, reaching \$3000 by beginning of year 2019. Currently bitcoin is priced at around \$ 8000. Thus, from the beginning of 2016 until now, bitcoin return was impressive 1700%.

Figure 2-9. Dynamics of trading volume (right axis) and price of Bitcoin since 2014



(Source: bitinfocharts.com and coinmarketcap.com)

Table 2-9 below demonstrates the higher moments of bitcoin returns - standard deviation, kurtosis and skewness, as well as autocorrelation of 1-3 lags, and cross-correlation between exchanges, from May 2014 to May 2019 at the 5-minute, hourly and daily frequencies. These stats

[33] Craig W. Holden, "Index Arbitrage as Cross-Sectional Market Making," Journal of Futures Markets, 15, 1995, pp. 423-455, for a general discussion of market efficiency and index arbitrage.

are computed by averaging the moments across all feasible exchanges. The estimates of standard deviation reported in the first column shows rather high volatility of returns. If we compute annualized deviation even from daily standard deviation it is still higher than 100%. By contrast, for Nasdaq shares this metric is equal to 0.17 from 1999 to 2018. Though, kurtosis, if we look at daily frequencies, is close to that of the Gaussian distribution. Positive skewness of bitcoin returns is probably not very surprising, due to significant increase in bitcoin price during the period analyzed. Low autocorrelations of the first 3 lags suggest that there is little predictability of bitcoin price.

Table 2-9. Descriptive statistics of Bitcoin returns (from May 2014 to May 2019)

Frequency	Standard deviation	Kurtosis	Skewness	ρ_1	ρ_2	ρ_3	Cross correlation
5 minute	1.45	147.95	-0.33	0.02	0.08	-0.01	0.67
Hour	1.34	7.31	-0.06	0.04	-0.05	-0.04	0.85
Daily	1.07	3.45	0.15	-0.07	-0.06	0.04	0.93

(Source: computed using data from bitinfocharts.com)

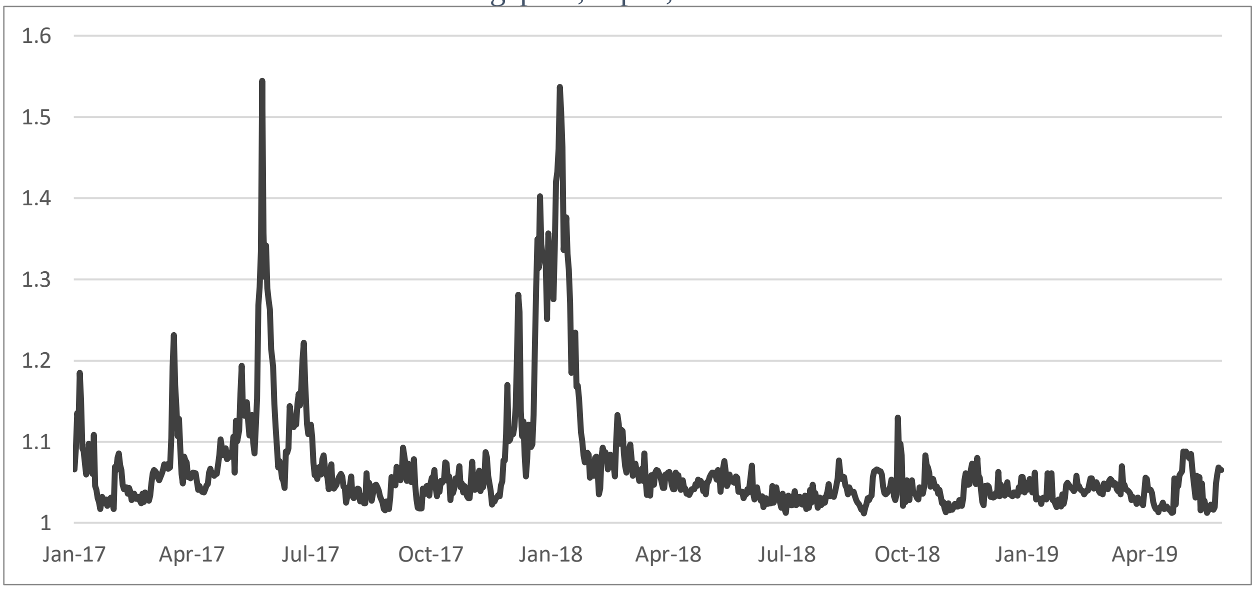
The last column shows the cross correlation of bitcoin returns across several exchanges. It was computed as average between all pairwise correlations. It is noticeable that correlation computed using high frequency returns is not very high and it gradually increases with decrease in frequency. These findings are comparable to those found in developed financial markets. But there is a significant difference between these cases. On the stock market, for instance, low correlations usually happen at millisecond timeframe (Budish, Cramton, Shim 2015), while for bitcoin market it is present even for minute data. These findings suggest that there might exist arbitrage opportunities between exchanges. In the following subsections it will be researched in more details.

2.3.2 Arbitrage index

Decentralization of cryptocurrency trading creates an interesting environment for studying exchange-wide arbitrage. The cross-correlations computed in the previous chapter indicated that cryptocurrency markets may not be completely effective. To quantify the price differences across exchanges at every time period, arbitrage index is computed. Its essence is that it measures the maximum price difference across given exchanges. To begin, the index is computed using minute timeframe data. It is calculated using volume-weighted price for every exchange in the sample. The index is computed by dividing max price across all short-listed exchanges for a given minute by min price. To clean the data from shocks, any movements of more than 10% between two neighboring transactions are excluded. At last, to decrease the effect of intra-day price fluctuation, we compute daily average of minute-level indexes.

Resulting index must be equivalent to 1 all the time if the markets are efficient and there are no arbitrage opportunities. Firstly, arbitrage index is computed using the exchanges from the previous parts, the sample is quite representative, there are exchanges from all over the world: United States, Europe, Hong Kong, Singapore, Korea, Japan, Australia. There are no exchanges from China, because many exchanges from there moved their headquarters to other countries due to strict regulations and ban on cryptocurrency exchanges in 2017 by Chinese government.[34]

Figure 2-10. Arbitrage index. Regions included are United States, Europe, Korea, Hong Kong, Singapore, Japan, Australia.



(Source: computed using data gathered from Catalyst library and particular exchanges API)

Figure 2-10 above indicates that the arbitrage index value is almost always above 1 for the period since January 2017 until May 2019. Average value of the index across this period is 1.07, median 1.05 and maximum value is 1.54. This implies that the distinction between the prices across various markets was 7% on an average day. Figure above also shows that important variation occurs during the analyzed period. It is noteworthy that there are several months in this period in which the arbitrage index remains constantly remains higher than 1.1, for instance December 2017 to February 2018.

It is not included in the graph, but the value for the index for the period since year 2014 was also computed. Interestingly, the arbitrage opportunities are much smaller during previous years, even if we take into account Chinese exchanges. Average of the index in 2016, for

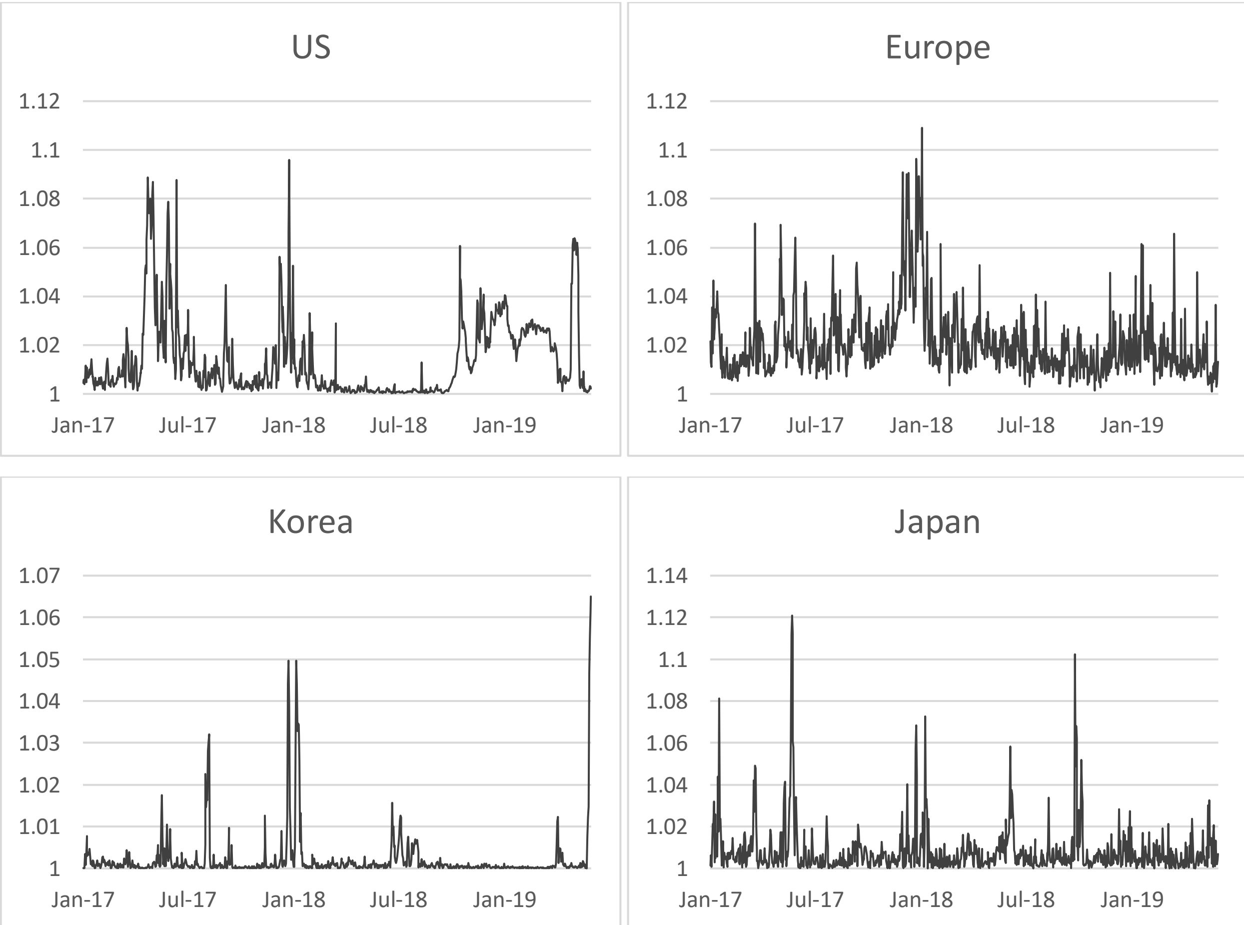
[34] See https://www.reuters.com/article/us-china-bitcoin/chinas-okcoin-huobi-exchanges-to-stop-bitcoin-withdrawals-idUSKBN15P0HE

example, was 1.03 as opposed to 1.07 in 2017 and the highest peaks happen in June and December. These findings indicate that the excessive possibilities for between-exchanges arbitrage opened up when volatility (as well as prices) raised significantly in 2017.

2.3.3 Arbitrage index within geographical regions

To examine further what are the key drivers of the arbitrage opportunities, the arbitrage index is computed across geographical regions (instead of within regions). Firstly, the index is computed for every geographical region in which more than 1 cryptocurrency exchange is present in the data, these would be Europe, US, Japan and Korea. For every region stated above, there is data for about 3-4 major exchanges, except for Korea, which has only 2, with proved liquidity, and time earned reputation.

Panel 1. Arbitrage index within geographical regions.



(US exchanges: Gemini, Kraken, Bitstamp, Coinbase Pro, currency used - USD. Japanese exchanges: Zaif, Bitflyer and Quoine, currency - JPY. European exchanges: Kraken, Bitstamp and Coinbase: currency - EUR. Korean exchanges: Bithumb and Korbit. KRW)

In the Panel 1 above, the graph on upper-left represents the arbitrage index computed on data from 4 major US cryptocurrency exchanges from the beginning of 2017 until May 2019. The approach was the same as for the previous case with all exchanges in one sample. The graph still has several spikes almost reaching 1.1, but on average variation of price is much smaller. Average value of the index is only 1.014. Spikes are about the same time as in the previous case: May, June and December 2017, but range is about 1.06-1.08. Although the spread is much smaller, compared to the total index, it is still huge if we compare it with stock market spreads, for example.

Table 2-10. Descriptive Statistics for Arbitrage Index in different regions

Arbitrage Index	**Average**	**Median**	**Maximum**
United States	1.014	1.007	1.096
Europe	1.021	1.017	1.109
Korea	1.002	1.000	1.065
Japan	1.008	1.005	1.121
Overall	1.069	1.048	1.545

(Source: computed using data gathered from Catalyst library and particular exchanges API)

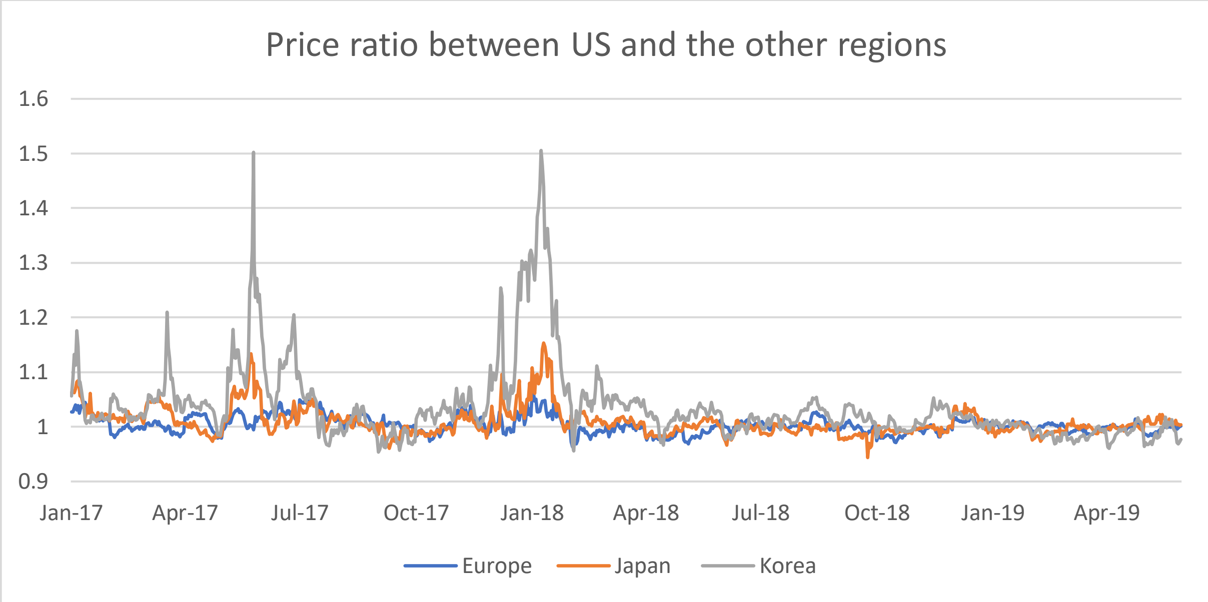
Then the same approach was repeated for the European cryptocurrency exchanges arbitrage index for the same period. Upper right graph shows that price variation within European exchanges is much lower than the arbitrage index as a whole. The average index is 1.021, median is 1.016. Compared to US the index for Europe has more spikes, and on average it is higher than for US. And again, a comparable image appears in Japan, where the arbitrage index on average is lower than for both US and Europe (1.008). In January and May 2017, and September 2018 there are brief intervals when the index moves to about 1.08. The peak in winter 2017 is similar to the trends in the other areas, but January 2017's rise is peculiar to Japan only.

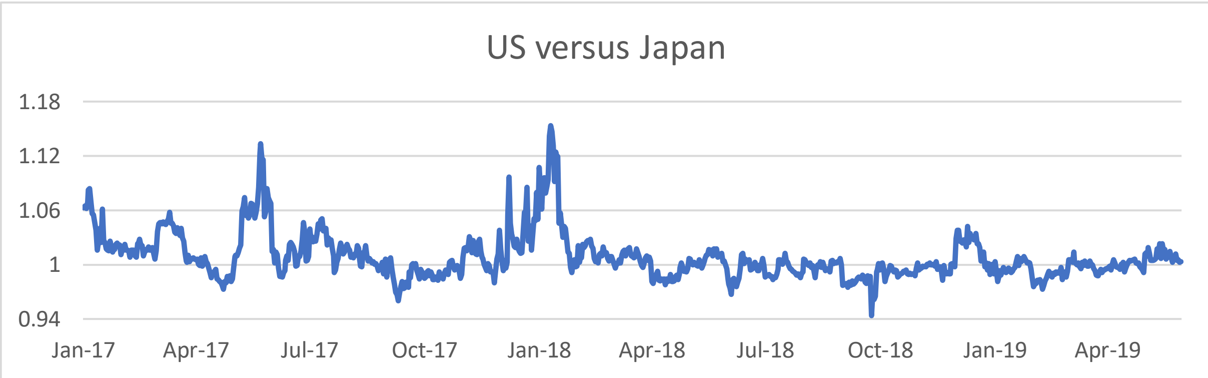
And lastly, we're looking at Korea, for which there is data for only two exchanges. Here the picture is different from other regions, we analyzed. The arbitrage index has also common peaks in September and in December 2017, but the range is much smaller than for other regions. At most it reached 1.05. Also, there is noticeable recent spike in May 2019, which is unique to Korean exchanges only. The index computed on Korean exchanges has lower average, mean and max values than for any other regions. Although, it might be due to lack of exchanges.

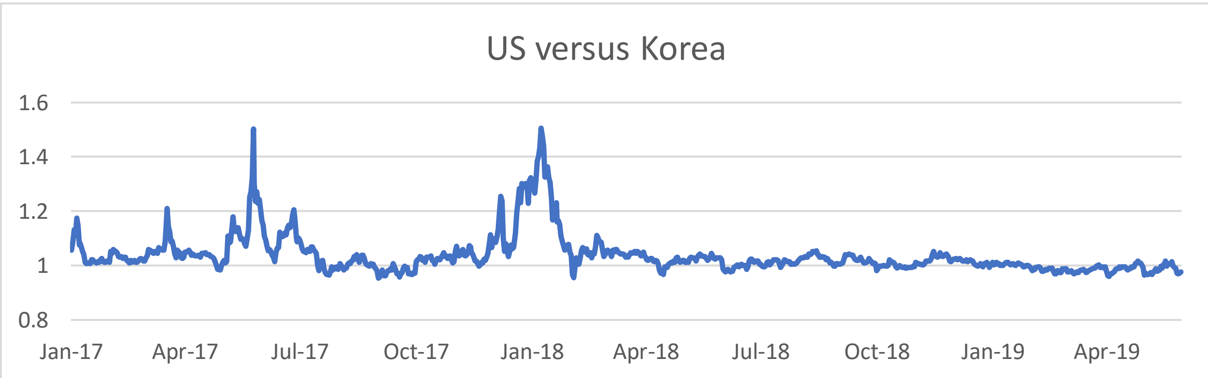
Overall, the findings indicate that the possibilities for arbitrage trades within geographical areas are much lower than between them. This implies that the most noticeable price variation is likely driven by country-wide price variation. The findings therefore indicate that crypto exchanges within specified areas appear to be much greater embedded than across areas.

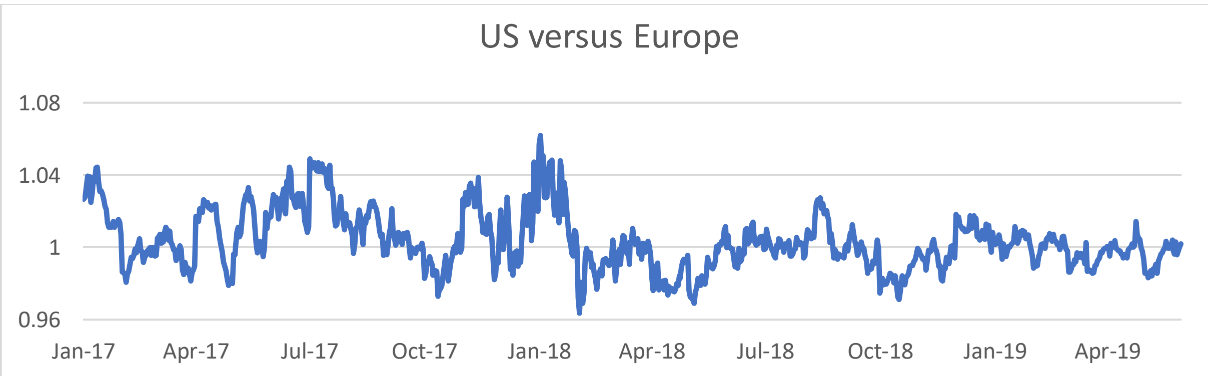
In order to verify the assumption that a substantial portion of the arbitrage spread is powered by price variation across geographic areas, the ratio of prices between United States exchanges and every other geographic region exchanges at minute timeframe is plotted in Panel 2 below. Ratios are calculated on the volume-weighted prices using minute timeframe and then averaged at the daily level. The graphs represent that prices on Korean crypto exchanges were two times 50%+ higher than on US exchanges and had the largest spread overall. The fact that Korean crypto exchanges had such a big premium during 2017 even has the special name - "Kimchi premium". At the same time as when Korea had peak spread, Japan also had significant bitcoin price deviations from the US. But Japan's bitcoin price premium to the US had a peak of about 1.15. In January 2018 Europe, on the other hand, experienced the largest price distinction to the United States of around 6% above US rates. Compared to other areas, the price gap between the United States and Europe is low, which is not strange given that the same exchanges operate in both US and Europe. The findings indicate that a large percentage of the huge arbitrage spreads calculated for bitcoin are dictated by region-wide price variations. And in many cases, these variations persist over rather long time periods.

Panel 2. Price ratios across regions. These graphs plot the average price ratio between the price of bitcoin to USD across several regions from the beginning of 2017 until May 2019.







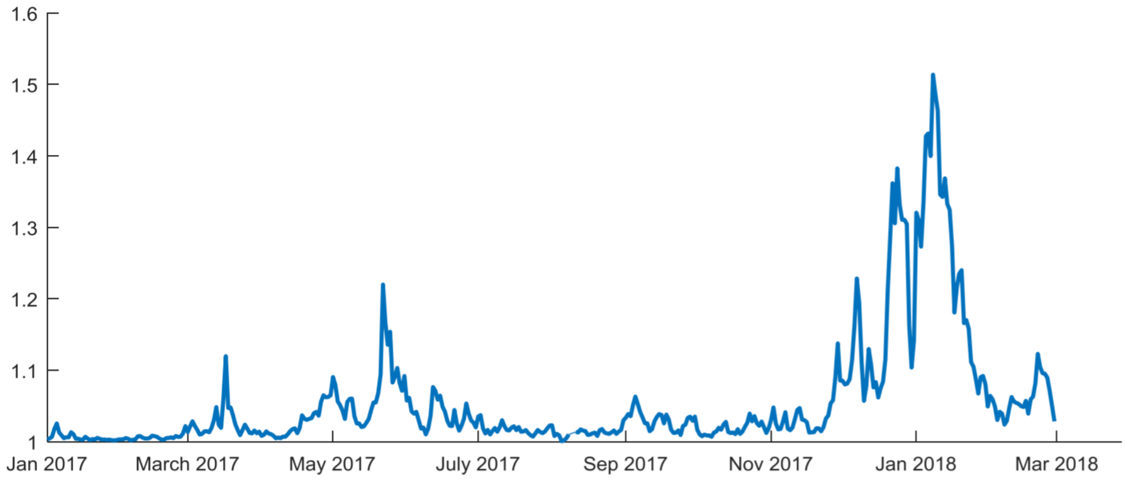


2.3.4 Arbitrage in other cryptocurrency markets

In the previous chapter it was indicated that there are decent arbitrage opportunities in bitcoin to fiat pairs across different geographical areas, especially between the United States, Japan, and Korea. In this subsection, it will be examined if there is arbitrage in the other cryptocurrency markets or it is unique to bitcoin markets only.

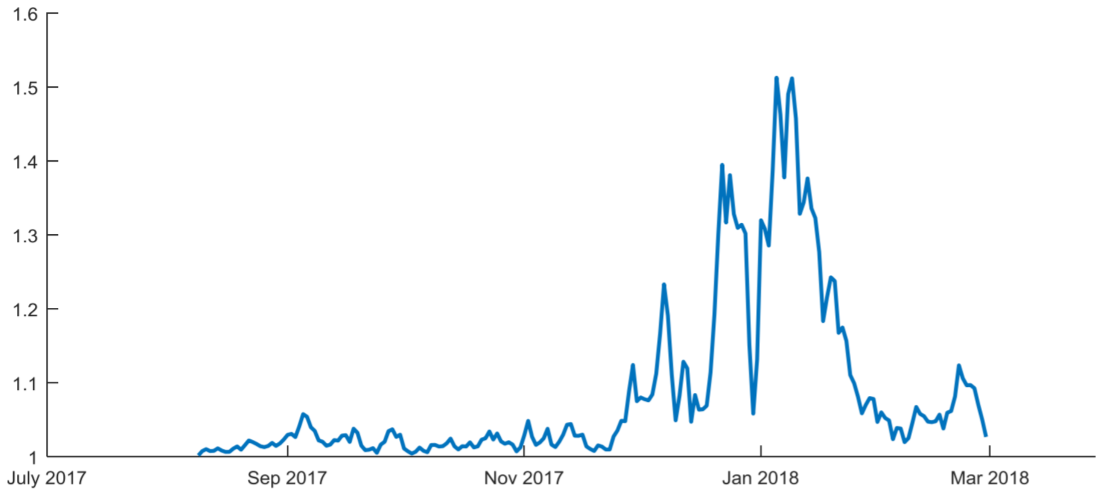
The figure below shows the dynamics of the arbitrage index for Ethereum and Litecoin to fiat currencies pairs. Although, these coins are together with bitcoin top 3 by volume traded, there is far less volume for them, and it is much harder to find good data without many errors. Because of the data problems it was decided to limit the time horizon for the most liquid and volatile period which is July 2017 - March 2018. It is noticeable that there is large variation in the index during this time period, which looks very similar to that of the bitcoin. In this time horizon there are periods of comparatively small values of the index and there are extended peaks at about the same time as on the bitcoin index graph. Both Ethereum and Litecoin similarly to Bitcoin reach 1.5 value of the index in December 2017 - February 2018.

Figure 2-11. Ethereum arbitrage index



(Source: Catalyst library)

Figure 2-12. Litecoin arbitrage index



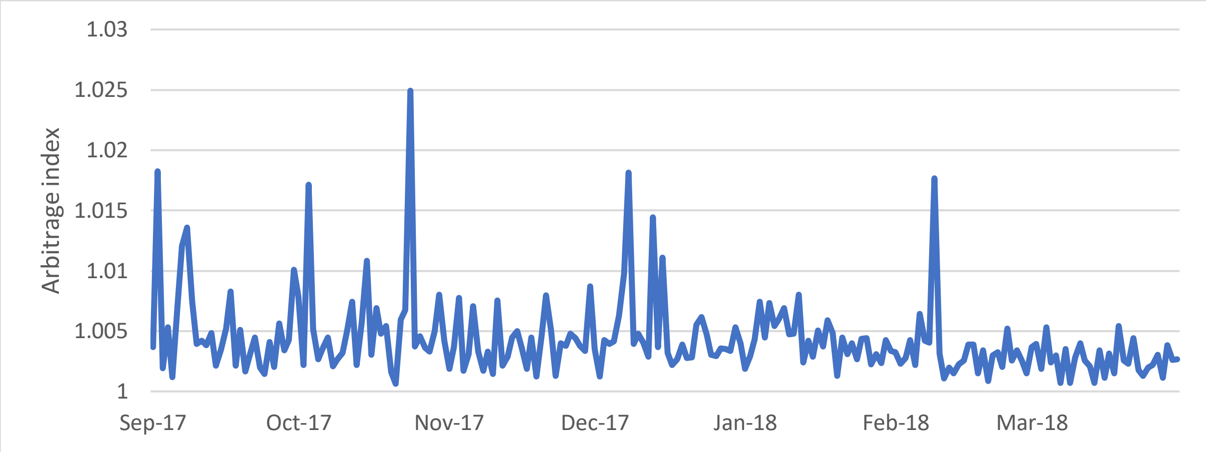
(Source: Catalyst library)

Similar to the bitcoin, price variability in Ethereum and Litecoin is due to differences of price across geographic areas. The graphs representing differences of the ETH and LTC prices across regions are not included in the paper, as they look almost the same as of bitcoin. In addition, even without statistical analysis one can observe similarities and most likely strong correlation between them by just looking at the graphs. All of the indexes rise at approximately the same moment and even reach comparable values. This is not a random coincidence, but due to the arbitrageurs' activity between crypto-to-crypto pairs, as it will be shown in the next subsection.

2.3.5 Arbitrage between cryptocurrency-to-cryptocurrency pairs

To assess whether arbitrage opportunities similar to crypto-to-fiat pairs, exist among crypto-to-crypto pairs, we examine the ETH to BTC pair, which is second cryptocurrency by volume traded. Similar to the previous part, time horizon is changed to the most liquid and volatile one (from September 2017 to March 2018). When price of the bitcoin increased in that period, trading in other cryptocurrencies also became more liquid. If limitations in the flow of funds drive the arbitrage opportunities between the Bitcoin and fiat currencies, then the variation in the arbitrage index of Ethereum to Bitcoin should be significantly lower. This could be due to inapplicability of capital restrictions to cryptocurrencies.

Figure 2-13. Arbitrage index for Ethereum-Bitcoin pair across all short-listed exchanges



(Source: Catalyst library)

The graph below shows the dynamics of the arbitrage index of ETH-BTC pair across all the exchanges in the short-list (not dividing by geographical regions). The approach was similar to the previous subsections, prices was weighted by value and aggregated at the daily level. One difference is that, because there were not many exchanges which provided an opportunity to trade bitcoin to Ethereum but rather bitcoin to fiat, and Ethereum to fiat pairs, to compute the ETH to BTC cross rate, local exchange rates were used. For instance, to compute the ETH-BTC exchange rate on Bitflyer, the approach was to divide the exchange rate of ETH to Japanese yen by the exchange rate of BTC to Japanese yen.

As pointed out earlier, the arbitrage index should be static and equal to 1 if there are no frictions and the markets are efficient. The difference is clear compared to the previous cases. There are several jumps, but the range is very different from the BTC to fiat, for example, at most its value reached 1.025. But most of the time index was lower than 1.005 boundary, which corresponds to almost no arbitrage opportunities (the level of transaction costs is roughly 0.5%). The mean value is actually even smaller and is equal to 1.0043, median 1.0037. To compare, the Kimchi premium during this period was reaching its peaks, about 50%. This demonstrates once again that there are much less arbitrage opportunities between various cryptocurrencies than between cryptocurrencies and fiat currencies.

2.4 Order flow

This section is dedicated to study of how and why arbitrage opportunities appear. The current literature suggests the significance of net order flows for price moves on regular financial markets, like stock or bond markets.[35] While prior study assigns net order flow price pressure to

[35] See Evans and Lyons (2002) study ForEx markets, Brandt and Kavajecz (2004) study government bond markets in the United States, Deuskar and Johnson (2011) study the market of futures contracts on S&

price discovery, it is far less obvious what are the fundamentals for cryptocurrencies and if there are traders with more data than others. However, in this subsection it is demonstrated that in crypto markets there is also a strong connection between net order flow and prices.

A usual well described in literature way of estimating the impact of net order flow is to regress asset returns on the signed trading volume during the same period. One of the complications we encounter is that there are a lot of Bitcoin exchanges, on which is traded at the same time, and the prices vary a lot from one to another, as proved in the previous section. Thus, when investors are going to buy Bitcoin or other cryptocurrency, they might consider prices on several exchanges, where the asset is traded. As a consequence, a regression of returns on signed volume on each exchange may offer a biased image of the real effect of net order flow on prices. In order to adapt the situation of various crypto exchanges, on each of them signed volume traded is divided into a common element and an idiosyncratic, exchange-specific element:

$$s_{it} = \overline{s_i} + \beta_i^s s_t^* + \widehat{s_{it}}, \quad (1)$$

$$E(s_t^*) = 0, \quad E(\widehat{s_{it}}) = 0, \quad E(\widehat{s_{it}} s_t^*) = 0.$$

Where s_{it} is signed volume on $i-th$ exchange, s_t^* is the common component for all exchanges, $\widehat{s_{it}}$ is an exchange specific component, and $\overline{s_i}$ is the idiosyncratic component mean. Return for each exchange is decomposed in the same fashion:

$$r_{it} = \overline{r_i} + \beta_i^s r_t^* + \widehat{r_{it}}, \quad (2)$$

$$E(r_t^*) = 0, \quad E(\widehat{r_{it}}) = 0, \quad E(\widehat{r_{it}} r_t^*) = 0.$$

Where r_{it} is computed as log return on $i-th$ exchange, $r_{it} = \ln(\frac{p_{i,t}}{p_{i,t-1}})$, r_t^* is the common return component for all exchanges, $\widehat{r_{it}}$ is an exchange specific return component, and $\overline{r_i}$ is the idiosyncratic return component mean. Equations (1) and (2) can be estimated either jointly using canonical correlation analysis or individually using factor analysis, assuming that:

$$E\left(\widehat{s_{it}}\widehat{s_{jt}}\right) = 0, \quad E\left(\widehat{r_{it}}\widehat{r_{jt}}\right) = 0, \quad for\ all\ i \neq j$$

Both models are linear and both estimate common factors as a linear combination of input data:

$$s_t^* = \sum_i w_i^s (s_{it} - \overline{s_i}),$$

$$r_t^* = \sum_i w_i^r (r_{it} - \overline{r_i}),$$

where $w_i^s\ and\ w_i^r$ are the factor weights, for which:

$$\sum_i w_i^s \beta_i^s = 1, \quad \sum_i w_i^r \beta_i^r = 1.$$

To fix the scale of the common factor in volume, we make the sum of factor loadings equal to 1. $\sum \beta_i^s = 1$.

To fix the scale of the common factor in returns, we restrict the sum of factors to be equal to 1. $\sum w_i^r = 1$.

By using this approach, we make common factor in returns similar to portfolio. For illustration, if β_i^s is the amount of bitcoins, bought on $i - th$ exchange, and since both $\sum \beta_i^s$ and $\sum_i w_i^s \beta_i^s$ are equal to one, than the amount of bitcoin bought is increased by one, thus the common component in volume is increased by one. Therefore, in the regression

$$r_t^* = \lambda s_t^* + \varepsilon_{t,}$$

λ measures the price pressure of the aggregate order flow.

Although the bitcoin price across different exchanges may vary over some period of time, as it is demonstrated in the previous analysis, it should be expected that the prices are cointegrated between any two exchanges. Also, linear combinations of bitcoin prices across exchanges should be cointegrated. Thus, the limitation that the amount of factor weights, w_i^r is equal to 1, enables us to divide the bitcoin price on each cryptocurrency exchange into a common element and a idiosyncratic deviation from the common element:

$$p_{it} = p_t^* + \widehat{p_{it}}; \quad p_t^* = \sum_i w_i^r p_{it}.$$

In contrast to p_t^* , each $\widehat{p_{it}}$ is a bounded process. Since we use the log-prices $\widehat{p_{it}}$ is the percentage deviation from the weighted average price across exchanges. If any of $\widehat{p_{it}}$ processes were unbounded, it would imply arbitrary large arbitrage opportunities.

2.4.1 Data

Both models are estimated on the data from 9 exchanges, chosen at liquidity analysis part of this thesis, on different timeframes (5 minute, 1 hour, 1 day). To get data on the signed trading volume, we need to use limit order book snapshots and updates from these exchanges, to understand if the trade is sell initiated or buy initiated. It limits the time interval used for research, because it is rather hard to obtain high-frequency data, especially, if one needs data from several exchanges and for a long period of time. Part of the database was collected by parsing data from particular exchanges, the other another part was taken from the open library on the GitHub, called CryptoIntegrity[36] (code for parsing was also partially taken from this library). The resulting time interval analyzed in this chapter is February – May 2019.

[36] https://github.com/CryptoIntegrity/

The tables below describe the results of this estimate. In the first part of the table #prices and volumes at 5-minute timeframe are used to estimate the factor loadings, weights and R-squared of the factor analysis of signed volume. As expected, for exchanges with the highest trading volume and liquidity, the factor loadings are also the highest. The top-3 highest loadings are assigned to Bitfinex, Coinbase Pro and Kraken. The common element explaining significant part of dispersion in idiosyncratic signed volume. The same approach is repeated at two other timeframes, and the output is summarized in the Table 2-11.

Table 2-11. Results of a factor analysis applied to the signed volume data

	Bitfinex	**Coinbase Pro**	**Kraken**	**Bitstamp**	**Gemini**	**Liquid**	**Bitflyer**	**Binance**	**Bithumb**
5 minute frequency									
loadings	0.145	0.092	0.271	0.237	0.061	0.035	0.062	0.024	0.073
weights	1.076	0.992	1.057	1.172	1.091	0.864	0.938	0.936	0.986
R-squared	0.777	0.597	0.658	0.548	0.299	0.238	0.317	0.242	0.280
1 hour frequency									
loadings	0.228	0.016	0.222	0.076	0.133	0.098	0.101	0.101	0.026
weights	1.144	0.996	1.141	1.160	1.090	0.806	0.938	0.864	0.972
R-squared	0.839	0.527	0.654	0.478	0.340	0.269	0.328	0.254	0.262
1 day frequency									
loadings	0.297	0.154	0.211	0.133	0.023	0.019	0.071	0.075	0.018
weights	1.092	1.000	1.112	1.156	1.091	0.868	1.029	0.975	0.985
R-squared	0.865	0.710	0.609	0.444	0.241	0.249	0.319	0.252	0.247

The results obtained demonstrate that a higher percentage of dispersion is explained by the common component of signed volume at lower frequencies; R-squared grows up to 86% for Bitfinex on daily data. The factor loading coefficients remain comparatively stable, and the loadings on the less liquid and less integrated exchanges also are less on lower frequencies.

Table 2-12. Results of the factor analysis applied to the log-returns data

	Bitfinex	**Coinbase Pro**	**Kraken**	**Bitstamp**	**Gemini**	**Liquid**	**Bitflyer**	**Binance**	**Bithumb**
					5 minute frequency				
loadings	1.195	1.064	1.013	1.004	0.999	0.990	0.942	0.982	1.019
weights	0.332	0.160	0.019	0.173	0.002	0.023	0.134	0.072	0.084
R-squared	0.774	0.674	0.703	0.561	0.325	0.348	0.388	0.355	0.404
					1 hour frequency				
loadings	1.150	1.102	0.966	0.983	0.997	1.019	0.973	1.000	1.016
weights	0.338	0.161	0.063	0.091	0.162	0.048	0.066	0.030	0.041
R-squared	0.841	0.732	0.735	0.658	0.359	0.401	0.402	0.397	0.441
					1 day frequency				
loadings	1.133	1.033	0.974	1.041	0.992	0.966	0.964	1.056	1.049
weights	0.371	0.125	0.090	0.140	0.046	0.096	0.030	0.085	0.017
R-squared	0.923	0.767	0.756	0.749	0.432	0.435	0.498	0.418	0.510

Similar approach is used for analyzing common component in log-returns. The results are displayed at the Table 2-12 above. As mentioned in the previous subsection, the only distinction is that in returns analysis there's a restriction on exchanges weights to sum up to 1 involved. We begin with the log-return data of 5-minute frequency. It is noticeable that common component in log-returns explains even higher portion of dispersion, than in signed volume. The common element in log-returns, even at the 5-minute frequency, accounts for about 77 percent of returns on Bitfinex. The lowest R-squared is on the Kraken exchange, equal to 32.5 percent. The R-squared is lower for exchanges in Japan and Korea, compared to the US exchanges. Generally, the pattern is consistent through different timeframes. As for the trading volume case, R-squared increases for lower frequencies data. On Bitfinex the common element explains 92% of price dispersion at daily frequency and 84.1% at hourly frequency. This is straightforward, because the correlation between hourly and daily log-returns is exceptionally high.

2.4.2 Common components and order flow

Table 2-13 below shows that a very big percentage of the common element in returns is explained by the common element in signed trading volume. The first column of the table reports coefficients and t-stats of regression of common component in log-returns on common component in trading volume, without lagged values. The regression coefficient for the first case is equal to $8.8 * 10^{-4}$ with T-stat equal to 80.06 and R-squared equal to 0.54, which demonstrates a very powerful connection between the common element in log-returns and the trading volume. That can be interpreted as, if buy volume increases buy 10^4 bitcoins (ceteris paribus), the bitcoin price will increase by 8.8%. Columns (2) and (3) suggest about the persistence of the cost effect. Regressions with lagged values help to assess the persistence of the price move. In the next columns we add 1 lag variable of trading volume, and then 5 lags. Negative coefficients suggest that some of the price pressure is temporary in the common component. Almost half of the impact on the price reverses during the following 5 periods.

Table 2-13. Results from time series regression of the common component of returns on common component of signed trading volume and its lagged values

	5 minute frequency λ*10^4 (%)			**1 hour frequency** λ*10^4 (%)			**1 day frequency** λ*10^4 (%)		
s_t^*	8.8	9.9	10.1	6	6.6	6.6	3.6	3.9	4
	(80.06)	(86.19)	(88.05)	(35.12)	(39.7)	(40.41)	(16.92)	(19.93)	(18.96)
s_{t-1}^*		-3.1	-2.6		-2.1	-2		-1.1	-1.1
		(-36.54)	(-32.24)		(-16.53)	(-15.67)		(-4.05)	(-3.62)
s_{t-2}^*			-0.8			-0.4			0
			(-11.68)			(-3.71)			(-0.2)
s_{t-3}^*			-0.5			-0.1			-0.1
			(-7.56)			(-1.22)			(-0.76)
s_{t-4}^*			-0.4			-0.3			-0.3
			(-6.88)			(-3.00)			(-1.71)
s_{t-5}^*			-0.3			-0.1			0.3
			(-5.24)			(-1.33)			(1.57)
R^2	0.54	0.61	0.62	0.72	0.74	0.75	0.82	0.83	0.85

The following columns show results for hourly and daily timeframes. For the hourly timeframe, the price impact of the trading volume is still extremely substantial, but slightly lower than for the previous timeframe. Interpretation is if one buys 10^4 bitcoins during an hour, then the price is predicted to grow by about 6%. Similarly, about half of the price impact is temporary, and disappears during the next 5 periods. At last, for the daily timeframe the coefficient drops to 3.6, suggesting a lower price effect overall. Mean reversion also decreases, the coefficient for the second and third lags are almost 0. If someone would buy 10^4 bitcoins during a day, the bitcoin price is forecasted to grow by 3.6%. Generally, it is demonstrated that rather high fraction of the common log-return element is explained by the common element of the signed trading volume in all timeframes, and price reversion after the impact becomes less pronounced at low frequencies than it is at greater frequencies.

2.4.3 Idiosyncratic price pressure

To estimate the idiosyncratic price pressure, we use the VAR model of the idiosyncratic part of signed trading volume and the deviations of particular exchanges from the common component:

$$\widehat{s_{it}} = \sum_{s=1}^{\tau} b_{i,s}\widehat{s_{i,t-s}} + \gamma_i \widehat{p_{i,t-1}} + u_{i,t},$$

$$\widehat{p_{it}} = \sum_{s=1}^{\tau} a_{i,s}\widehat{p_{i,t-s}} + \lambda_i \widehat{s_{i,t}} + v_{i,t},$$

where

$$E(v_{i,t}) = E(u_{i,t}) = 0, \quad E(v_{i,t}v_{i,s}) = E(u_{i,t}u_{i,s}) = 0, \qquad for\ all\ s \neq t, \quad E(v_{i,t}u_{i,s}) = 0.$$

These equations are estimated as a system using ordinary least squares method and the outcome is reported in the table below. After removing the common element from each price, the residual value left is the idiosyncratic price element. The particular exchange differences from the common element are then estimated as a function of signed trading volume and its lagged values. Lagged values of idiosyncratic trading volume are not included in the model, since the estimation of the first equation suggests that lags of the volume have weak correlation with the current residual signed trading volume.

Table 2-14. Results from the regression of idiosyncratic component in trading volume of the particular exchanges on the deviation of the price from the component and 3 lags of the trading volume of that exchange.

	Bitfinex	**Coinbase Pro**	**Kraken**	**Bitstamp**	**Gemini**	**Liquid**	**Bitflyer**	**Binance**	**Bithumb**
					5 minute frequency				
$\gamma_i * 10^{-2}$	-1.12	0.86	-1.61	-0.22	-0.45	-0.15	-0.49	-0.23	0.05
	(-2.4)	(4.12)	(-15.41)	(-5.19)	(-2.66)	(-7.23)	(-7.78)	(-5.35)	(7.12)
b_{1i}	0.09	0.17	0.07	0.14	0.06	0.23	0.15	0.15	0.21
	(5.58)	(17.72)	6.58	(11.29)	(6.18)	(21.31)	(17.55)	(13.42)	(32.29)
b_{2i}	0.06	0.08	0.03	0.04	0.02	0.07	0.05	0.04	0.09
	(6.52)	(9.24)	(2.74)	(3.91)	(2.85)	(8)	(6.76)	(4.71)	(14.23)
b_{3i}	0.05	0.09	0.02	0.04	0.02	-0.07	0.06	0.06	0.08
	(5.19)	(9.82)	(2.32)	(5.34)	(3.04)	(7.87)	(7.9)	(7.01)	(12.34)
R^2	0.03	0.06	0.02	0.04	0.01	0.08	0.04	0.03	0.08

Note: The T-stats (in brackets) are computed using heteroscedasticity robust standard errors.

$\widehat{s_{it}} = \gamma_i \hat{p}_{i,t-1} + b_{1i}\hat{s}_{i,t-1} + b_{2i}\hat{s}_{i,t-2} + b_{3i}\hat{s}_{i,t-3} + \varepsilon_{it}$ – equation estimated in Table 2-14

Table 2-15. Results from the regression of the deviation of the price from the common component on each exchange on three lags of the difference of the price from the common component of the price and exchange-specific component of the signed volume.

	Bitfinex	**Coinbase Pro**	**Kraken**	**Bitstamp**	**Gemini**	**Liquid**	**Bitflyer**	**Binance**	**Bithumb**
					5 minute frequency				
$\lambda_i * 10^4 (\%)$	2.86	17.35	41.54	5.76	8.37	4.35	17.13	22.66	32.1
	(16.49)	(22.93)	(27.66)	(9.18)	(14.35)	(6.58)	(22.26)	(14.00)	(25.13)
a_{1i}	0.6	0.63	0.62	0.55	0.59	0.79	0.83	0.6	0.83
	(48.44)	(16.29)	(40.09)	(56.57)	(34.45)	(26.36)	(40.51)	(61.34)	(50.95)
a_{2i}	0.23	0.19	0.19	0.23	0.24	0.15	0.12	0.21	0.12
	(17.07)	(5.59)	(11.51)	(21.47)	(13.5)	(5.55)	(4.8)	(21.32)	(6.45)
a_{3i}	0.16	0.19	0.16	0.2	0.16	0.05	0.04	0.19	0.05
	(12.84)	(5.51)	(9.52)	(21.89)	(11.19)	(2.59)	(1.65)	(18.78)	(3.64)
R^2	0.97	0.97	0.94	0.94	0.89	0.98	0.99	0.98	0.99

Note: The T-stats (in brackets) are computed using heteroscedasticity robust standard errors.

$\widehat{p_{it}} = \lambda_i \hat{s}_{i,t-1} + a_{1i}\hat{p}_{i,t-1} + a_{2i}\hat{p}_{i,t-2} + a_{3i}\hat{p}_{i,t-3} + \varepsilon_{it}$ – equation estimated in Table 2-15

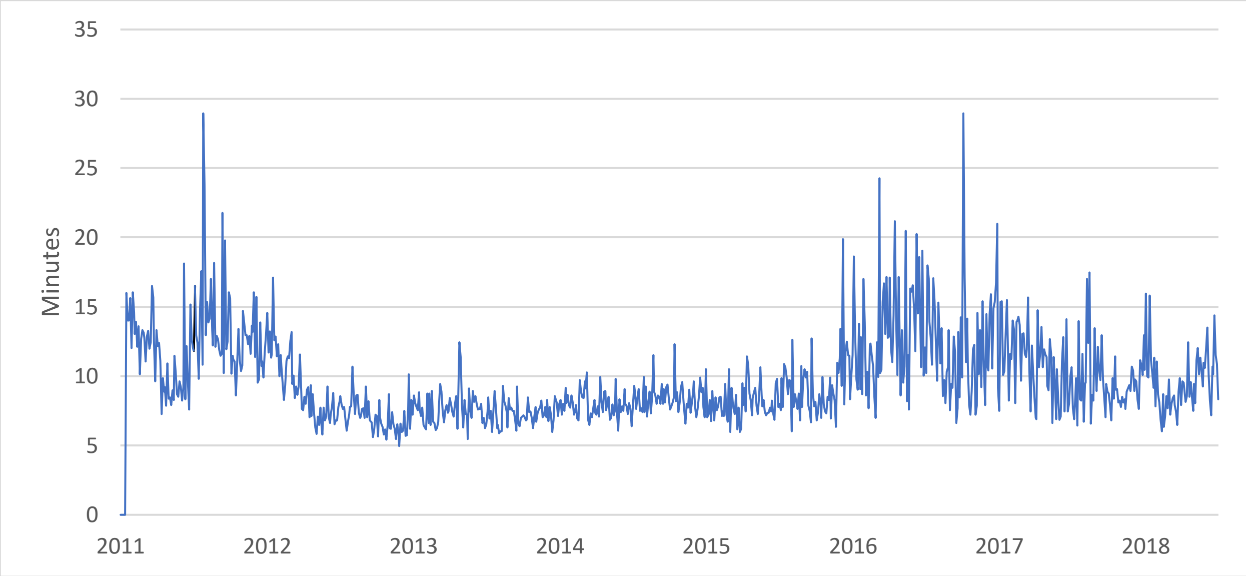
Every column of the table above shows the results of estimation of the idiosyncratic price component for each exchange. Coefficients on the lags of $p_{i,t}$ values suggest that that $p_{i,t}$ is an extremely persistent process of mean-reversal. All of the coefficients are positive for the previous 3 lags. The summed value of three lags is almost equal to 1 for each exchange. If the bitcoin price on any particular exchange moves up from the average price on the other exchanges, then the log-returns during the following periods are forecasted to be below the log-returns on other exchanges. But the reversion to the common element is slow. For Bitflyer or Bithump, which are both Asian exchanges, the coefficients on lagged variables are especially high. This supports the previous findings that on these exchanges arbitrage opportunities persist for longer.

The idiosyncratic price pressure is considerably large on nearly all crypto exchanges compared to the price pressure estimated for the common element. If the exchange has comparatively low liquidity, it is expected also to have high price pressure. Bitfinex and Bitstamp are the two exchanges for which the projected coefficient on the individual price effect is lower than that on the common. This explained by their size and high liquidity. However, there is a limitation for the chosen approach to model idiosyncratic price pressure. If price is either significantly higher or lower on a particular exchange, traders may change the exchange they use, and overall the way the trade. The exchange-specific portion of the signed trading volume could therefore change endogenously, and this type of relationship is not captured by the chosen model. However, our findings indicate that the idiosyncratic portion of the signed trading volume plays a significant role in explaining the price deviations from the common component on a particular exchange.

2.4.4 Arbitrage trading strategies

This section is dedicated to outlining the approaches to the arbitrage trading on cryptocurrency markets and to the possible risks and costs that may negatively impact trading efficiency. Consider the case discussed above - Kimchi premium, the bitcoin price in Korea is above the average price in the United States. This scenario would be a risk-free arbitrage if there were no frictions. Trader could purchase bitcoins in the United States, send them to Korea, sell bitcoin to Korean Won, then exchange KRW for USD, and make a transfer back to the United States. This theoretical arbitrage trading is impossible in reality because the nature of bitcoin transactions assumes that it takes some time to record the transaction on the blockchain (about 10 minutes on average, see Figure 2-14). In addition, it usually takes cryptocurrency exchanges from several hours to few days to transfer fiat currency. Obviously, arbitrage opportunity will not last for that time. As a default, an arbitrage trader must purchase bitcoin on the exchange with lower price and sell it on the exchange with a higher price at the same moment in time in order to lock in the arbitrage.

Figure 2-14. Median confirmation time of bitcoin transaction in minutes (averaged weekly)



(Source: Blockchain.com)

Optimally, the trader would like to short the bitcoin on the exchange with a higher price, for example Korea, and long bitcoin in the United States. Then the trader could send the bitcoin to the United States and lock in the risk-free profit. Though, this arbitrage trading approach is frequently not viable, as there are not so many cryptocurrency exchanges that enable short selling. In case of impossibility of short selling, the trader may switch to two alternative arbitrage trading approaches.

The first method is to get negative exposure in bitcoin by margin trading, which is almost the same as short selling, except that it doesn't allow for physical settlement. If the trader sticks to this method, he will be able to realize arbitrage profit only if the prices on the exchanges intersect in the future. The trader is therefore subject to the risk of convergence, which has been researched widely in the limits to arbitrage literature.[37] While theoretically exchange-wide prices might not intersect for a long period of time, graphs in the Panel 2 demonstrate that in reality, arbitrage opportunities on cryptocurrency markets existed for about several days and never lasted for more than two months.

Another method is to keep some amount of bitcoins on both exchanges and purchase and sell them at the same moment in time on the appropriate exchanges when the arbitrage opportunity appears. Of course, the trader's bitcoin balance will decrease on the exchange where bitcoin's price was larger (because that's where he sells bitcoin) and rise on the exchange where the price

[37] See for example, Shleifer and Vishny (1997), or Gromb and Vayanos (2002)

was smaller. The arbitrage trader then should send bitcoins from the elevated bitcoin balance exchange to the low balance exchange to equalize amounts across two exchanges. While this approach doesn't make an exposure to the risk of convergence, a major disadvantage of this approach is that the trader has exposure in bitcoin to fiat. To minimize that risk, the arbitrage trader could borrow bitcoin from people who have big amounts of cryptocurrency and are not planning to sell it in the near future, the so called hodlers[38]. Obviously arbitrage trading can suit hodlers themselves, because of their constant exposure to cryptocurrencies. The trader might also use futures on bitcoin to hedge the risk.[39]

In reality, the arbitrage trader will bear several transaction costs, but comparing to the possible profits, their magnitude is way too small to discourage traders from arbitrage. Firstly, the transaction must be registered in the blockchain to send bitcoins. This is the job of the bitcoin miners who provide transaction approval. On average this cost is around $10 per single transaction, although they peaked at $30 during high volatility periods. These costs are independent of the size of the transfer; therefore, they are relatively small for the potential arbitrageur. Second, almost all the exchanges have trading charges that significantly boost trading costs. These charges range from 0.25% of the traded quantity to 0.05%, but on some exchanges, it may be even free, if the trader adds liquidity to the market, instead of taking it away (limit orders instead of market orders). Some exchanges do not collect trade-based charges immediately but assign them in a specified month or week relative to volume traded. The exchange charges are of comparable size to the spreads, which are somewhere between 0.01% and 0.1% on average. Lastly, many exchanges require charges at withdrawal, ranging from 0.1% to 0.5%. However, all major exchanges claim that they provide favorable tailored fees for large clients, which are well below individual investors' fees. In short, for big operations, these charges are low. Ultimately, the summed-up trading fees for big traders should be between 0.25% and 0.50% or may be even less. Relative to the possible arbitrage yields, reported in the previous section, these costs are insignificant and thus cannot explain the existence of such arbitrage spreads.

Another aspect that could restrain traders' readiness to participate in arbitrage deals is the governance risk of exchanges. It occurs because the trader has to send his bitcoins to the exchange wallet and thus lose control of the cryptocurrency in favor of the exchange. Judging from several commonly published exchange exploits or simply dishonesty and fraud may result in substantial

[38] The term hodler is a peculiarity of the bitcoin market since one investor in bitcoin wrote in a post on the Bitcoin talk forum in 2013 while prices were dropping I AM HODLING. This has become a meme for Hold On for Dear Life
[39] Bitcoin futures started trading on CME and CBOE in the end of 2017. Average daily open interest is about 10000 bitcoins. See, for instance,
https://www.cmegroup.com/trading/equity-index/us-index/bitcoin_quotes_settlements_futures.html

losses for investors trading there.[40] A recent example is when the hackers stole $40 million worth from Binance exchange.[41] In this case, however, the head of the exchange promised to reimburse to the clients all the losses.

It seems doubtful, though, that all the above-mentioned issues might explain the discovered arbitrage spread. Doubts about an exchange's governance risk should influence its trading volume and potentially spread. But the exchanges chosen for analysis already show high liquidity and low spreads. It was also demonstrated that the spreads are much bigger across geographical regions. To explain this trend by exchange risk, one would have to suppose that the exchange risk is correlated within a region. However, it is not backed by the data as there is considerable heterogeneity in the exchanges' liquidity within a particular region, but nonetheless spreads between them are low.

At last, cross-border capital controls are a serious limit to arbitrage trading. As it was mentioned earlier, unless the trader wants to bet on future convergence in prices between Korean and United States exchange, then he would have to sell cryptocurrency in Korea for fiat currency and transfer it back from Korea to the United States. The government regulation of some countries makes it hard, especially for individual investors, to make cross-border deals. For instance, Korean residents and businesses moving an equivalent of more than $50,000 out of Korea in a single year, must submit documentation to officials showing their justifications for the transactions, and these transactions may not always be approved and permitted. Market studies as well as trading blogs indicate that these limitations are most likely binding for individual investors. However, quantifying how binding these limitations are for big organizations, which perform trading on various global economic markets is harder.[42]

There are several studies suggesting that these limitations can be avoided by big organizations. In a latest IMF working document, Chikako and Kokenyne (2011) discover that the efficiency of capital regulations in South Korea appears to be low, as money flows into and out of Korea as well as monetary policy efficiency do not seem to change considerably after capital regulations has been introduced. Also, industry researches indicate that there are forex dealers that assist organizations to transfer capital in and out of Korea. Capital bindings should not, therefore, place insuperable limitations on arbitrage activities across geographical areas, especially for big traders, they just contribute to the costs of trading. This is backed by the fact that arbitrage spreads

[40] For instance, in the hack of Mt Gox in 2014 850,000 bitcoins were stolen from customers and the company.

[41] See https://www.bloomberg.com/news/articles/2019-05-08/crypto-exchange-giant-binance-reports-a-hack-of-7-000-bitcoin

[42] A related constraint is that many retail investors face restrictions on which exchanges they can trade. For example, foreign nationals are typically prevented from opening up accounts and trading on local exchanges. But similar to capital controls large financial institutions should be able to bypass these restrictions and be able to operate across regions.

on the same cryptocurrency exchanges where we see large spreads to fiat currencies are much lower in two-way cryptocurrency trades. But even in this case the arbitrage opportunities do not persist longer than two months.

To conclude, the analysis shows that the past of bitcoin exchanges is tagged by reoccurring episodes of expanding arbitrage opportunities and some periods of exceptionally large spreads that last for about a month. It seems that most of the time, arbitrage traders are able to balance bitcoin prices across various exchanges.[43] But sometimes the arbitrage traders' capital appears to be getting overloaded by the noise traders, individuals, who drive up prices on some exchanges or get feared and make panic sells when negative news about cryptocurrencies appear. One interpretation of the obtained results might be that the capital involved in arbitrage trading on cryptocurrency markets moves slowly, and traders, because of the described restrictions and risks, cannot scale up their trading algorithms with the intensity of noise trader activity in a reasonable amount of time.

[43] Industry reports suggest that hedge funds and high frequency traders have been active across different cryptocurrency markets for several years.

Conclusion

In this paper, we have studied several issues of cryptocurrency trading connected to arbitrage. Firstly, it was shown that there are a lot of exchanges, which use wash-trading to inflate their reported trading volumes. The results suggest that there are three groups of exchanges: some don't seem to inflate their trading volumes, some of them inflate volumes at some extent, but they also have rather high actual trading activity, and some exchanges have mostly artificial volumes of trading. The dishonest cryptocurrency exchanges are more or less known to the market participants, but the added value in this paper is that we measured and quantified the real liquidity and market depth.

After that, focusing on bona fide exchanges only, an assessment of arbitrage opportunities was conducted. It was shown that there are significant arbitrage opportunities across different exchanges that open up repeatedly and often persist for long periods of time (up to 2 months). Importantly, these arbitrage opportunities are much larger across geographical areas than within the same country. This implies that cryptocurrency exchanges within specified areas appear to be much greater interconnected than across areas. The arbitrage opportunities are not limited to Bitcoin, the pattern is very much the same for other considered cryptocurrencies, Litecoin and Ethereum. But there are significantly less arbitrage opportunities between crypto-to-crypto trading pairs than between crypto-to-fiat pairs.

To understand how these price differences across exchanges develop, we analyze the relationship between net order flows and prices. We decompose signed volume and returns on each exchange into a common component and an idiosyncratic, exchange-specific component. The common component of signed volume explains about 75% of the returns' dispersion using 5-minute data, and up to 87% for daily data. The exchange-specific residuals of signed volume are significant at explaining variation in exchange-specific residuals of returns at 5-minute and hour level. We also show that when the price on any exchange deviates above (below) from the average price on other exchanges then subsequent returns on this exchange are predicted to be lower (higher) than the returns on other exchanges.

Further, we discussed possible strategies of arbitrage trading, specific to cryptocurrency market and attempted to assess limits to arbitrage. The arbitrage trader bears several transaction costs, among which miners fee, required to send bitcoin, trading fees on exchanges and deposit/withdrawal fees. But comparing to the possible profits, their magnitude is way too small to discourage traders from arbitrage. The total trading fees for big traders should lie somewhere between 0.25% and 0.50% or may be even less. Relative to the possible arbitrage yields, these costs are insignificant and thus cannot explain the existence of such arbitrage spreads. Another aspect that could restrain traders' readiness to participate in arbitrage deals is the governance risk of exchanges. It occurs because the trader has to send his bitcoins to the exchange wallet and thus lose control of the cryptocurrency in favor of the exchange. Judging from several commonly published exchange exploits this may result in substantial losses for investors trading there. A recent example is when the hackers stole $40 million worth from Binance exchange. It seems doubtful, though, that all the above-mentioned issues might explain the discovered arbitrage spread. Doubts about an exchange's governance risk should influence its trading volume and potentially spread. But the exchanges chosen for analysis already show high liquidity and low spreads.

To conclude, the analysis shows that the past of bitcoin exchanges is tagged by reoccurring episodes of expanding arbitrage opportunities and some periods of exceptionally large spreads that last for about a month. It seems that most of the time, arbitrage traders are able to balance bitcoin prices across various exchanges. But sometimes the arbitrage traders' capital appears to be getting overloaded by the noise traders, individuals, who drive up prices on some exchanges or get feared and make panic sells when negative news about cryptocurrencies appear. One interpretation of the obtained results might be that the capital involved in arbitrage trading on cryptocurrency markets moves slowly, and traders, because of the described restrictions and risks, cannot scale up their trading algorithms with the intensity of noise trader activity in a reasonable amount of time.

Bibliography

Abadi, Joseph and Brunnermeier, Markus Konrad, *Blockchain Economics* (December 2018). NBER Working Paper No. w25407.

Ametrano, Ferdinando M., Hayek. *Money: The Cryptocurrency Price Stability Solution* (August 13, 2016).

Aste, T. *Cryptocurrency market structure: connecting emotions and economics. Special Issue of Digital Finance on Cryptocurrencies*. (2018). ISSN 2524-6186.

Babaioff, Moshe & Dobzinski, Shahar & Oren, Sigal & Zohar, Aviv. *On Bitcoin and Red Balloons*. (2011).

Babich, Volodymyr and Hilary, Gilles. *Distributed Ledgers and Operations: What Operations Management Researchers Should Know About Blockchain Technology* (August 27, 2018).

Barker, J. *Triangular arbitrage with Bitcoin*. Texas A&M University. (2017)

Bech, Morten L. and Garratt, Rodney, *Central Bank Cryptocurrencies* (September 17, 2017). BIS Quarterly Review September 2017.

Bistarelli, S., A. Cretarola, G. Figa-Talamanca, I. Mercanti, and M. Patacca. *Is Arbitrage Possible in the Bitcoin Market?* In: M. Coppola, E. Carlini, D. (2019)

Buterin, V. *On Stake.* Ethereum Blog. (2014)

Camerer, C. *Bubbles and fads in asset prices*. (1989). Journal of Economic Surveys 3 (1): 3–41.

Chaim, P., and M. Laurini. *Is Bitcoin a bubble?* (March 2019).

Chaum, D. *Blind Signatures for Untraceable Payments.* In: Chaum, D., Rivest R.L.

Cong, L., Li, Y., and N. Wang. *Tokenomics: Dynamic Adoption and Valuation*. (2018). SSRN.

Cretarola, A., G. Figa-Talamanca. *A Confidence-Based Model for Asset and Derivative Prices in the Bitcoin Market.* (2017)

D’Agostino, J. Altmann, and J. Banares. *Economics of Grids, Clouds, Systems, and Services.* GECON 2018. Lecture Notes in Computer Science, vol 11113. Springer, Cham.

Delmolino, Kevin & Arnett, Mitchell & Kosba, Ahmed & Miller, Andrew & Shi, Elaine. *Step by Step Towards Creating a Safe Smart Contract: Lessons and Insights from a Cryptocurrency Lab.* (2016).

Demirgüç-Kunt, Asli, Leora Klapper, Dorothe Singer, Saniya Ansar, and Jake Hess. 2018. Global Findex Database 2017: Measuring Financial Inclusion and the Fintech Revolution. Washington, DC: World Bank.

Dong, F., Z. Xu, and Y. Zhang. *Bubbly bitcoin.* (2018). SSRN.

Dwyer, Gerald P., and Margarita Samartín. *Why Do Banks Promise to Pay Par on Demand?* (2009). Journal of Financial Stability 5 (June): 147–169

Foley, Sean and Karlsen, Jonathan R. and Putnins, Talis J. *Sex, Drugs, and Bitcoin: How Much Illegal Activity Is Financed Through Cryptocurrencies?* (December 14, 2018). Review of Financial Studies.

Franco, P. *Understanding bitcoin*. (2014). Chichester, West Sussex: John Wiley & Sons Ltd.

Hafner. *Testing for Bubbles in Cryptocurrencies with Time-Varying Volatility*. (2018). SSRN.

Hautsch, N., Scheuch, Ch., and S. Voigt. *Limits to Arbitrage in Markets with Stochastic Settlement Latency*. (2018). CFS Working Paper, No. 616.

Hayes, A. *Cryptocurrency Value Formation: An empirical study leading to a cost of production model for valuing Bitcoin*. (2017). Telematics and Informatics.

Hileman, G., and M. Rauchs. *Global Cryptocurrency Benchmarking Study. (*2017).

Iwamura, Mitsuru and Kitamura, Yukinobu and Matsumoto, Tsutomu and Saito, Kenji. *Can We Stabilize the Price of a Cryptocurrency?: Understanding the Design of Bitcoin and Its Potential to Compete with Central Bank Money*. (October 25, 2014).

Karame, Ghassan & Androulaki, Elli & Capkun, Srdjan. *Double-spending fast payments in Bitcoin.* (2012).

Kasper, Dr. Jochen. *Evolution of Bitcoin - Volatility Comparisons with Least Developed Countries' Currencies.* (October 13, 2017).

Lee, S. *Bitcoin's Energy Consumption Can Power An Entire Country – But EOS Is Trying To Fix That.* (2018). Forbes.

Makarov, I., and A. Schoar. *Trading and Arbitrage in Cryptocurrency Markets*. (2018). The Journal of Financial Economics.

Marimon, Ramon & Nicolini, Juan Pablo & Teles, Pedro. *Money is an experience good: Competition and trust in the private provision of money.* (2012). Journal of Monetary Economics, Elsevier, vol. 59(8), pages 815-825.

Mishra, S., V. Jacob, and S. Radhakrishnan. *Energy Consumption – Bitcoin's Achilles Heel. (2018).*

Nakamoto, S. *Bitcoin: A Peer-to-Peer Electronic Cash System*. (2008).

Nasekin, S., and C. Chen. *Deep Learning-Based Cryptocurrency Sentiment Construction*. (2019).

Ong, B., L. Guo, and K. Lee. *Evaluating the Potential of Alternative Cryptocurrencies.* (2015). Elsevier, Academic Press. Pp. 81–135.

Pagnotto, E., and A. Buraschi. *An Equilibrium Valuation of Bitcoin and Decentralized Network Assets*. (2018). SSRN.

Reid, Fergal & Harrigan, Martin. *An Analysis of Anonymity in the Bitcoin System. Security and Privacy in Social Networks.* (2011).

Schilling, L., and H. Uhlig. *Some simple Bitcoin Economics*. (2018). Becker Friedman Institute for Research in Economics Working Paper No. 2018-21.

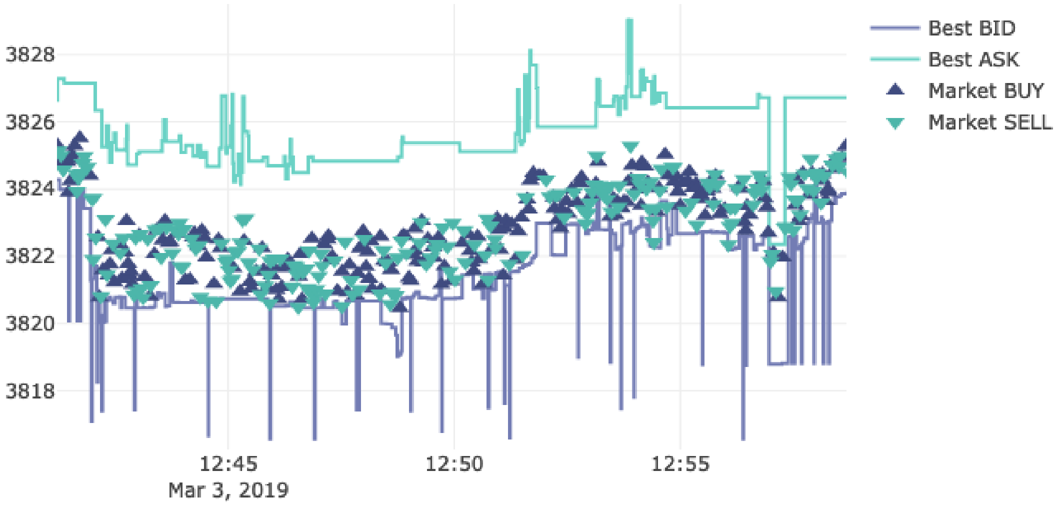
Sharma, R. *The Number Of Dark Pools In Cryptocurrency Trading Is Increasing*. (2018). Investopedia.

Wallace, Benjamin. *The rise and fall of Bitcoin*. (2011, November 23). Wired Magazine

World Bank. 2017. *World Bank Annual Report* 2017. Washington, DC: World Bank.

Appendix

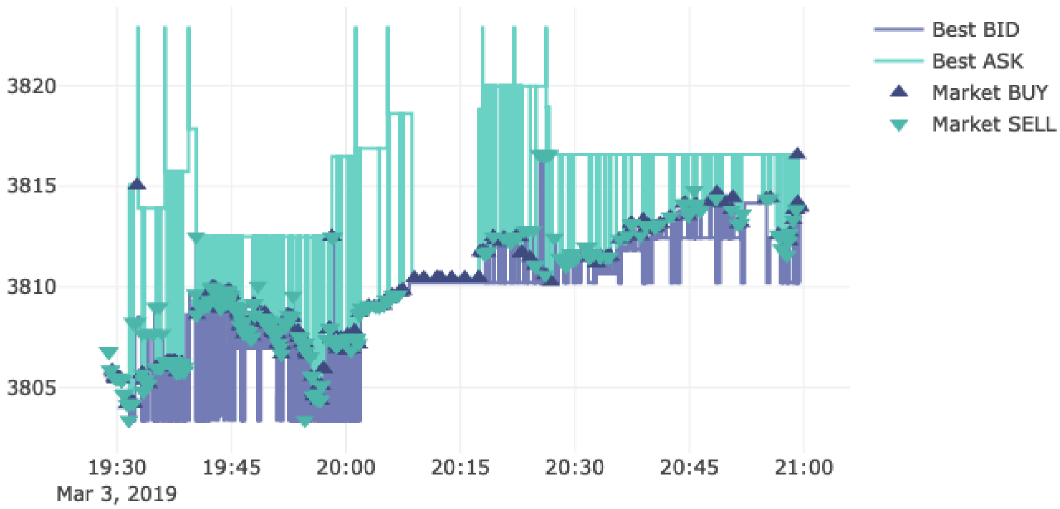
Figure 1. Identification of wash trading BTC_USDT pair on BIT-Z



(Source: computed using code from CryptoIntegrity, data gathered directly from exchange)

Bit-Z shows systematic in-spread trades without any change in bid-ask spread. Unstable bid and ask prices can also be signals of low liquidity.

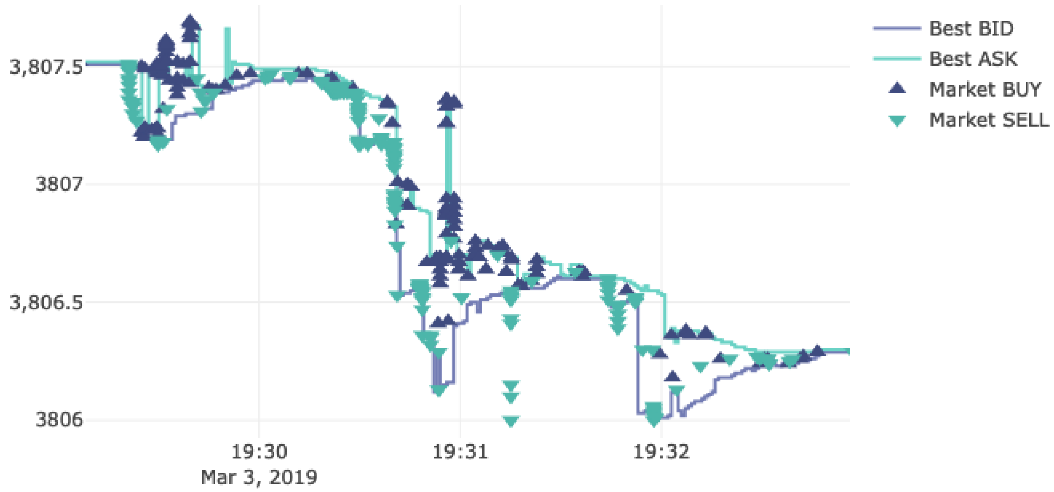
Figure 2. Identification of wash trading BTC_USDT on Coinbene



(Source: computed using code from CryptoIntegrity, data gathered directly from exchange)

Coinbene orderbook data shows wide bid ask spread with short lifespan limit orders which are almost always executed in-spread.

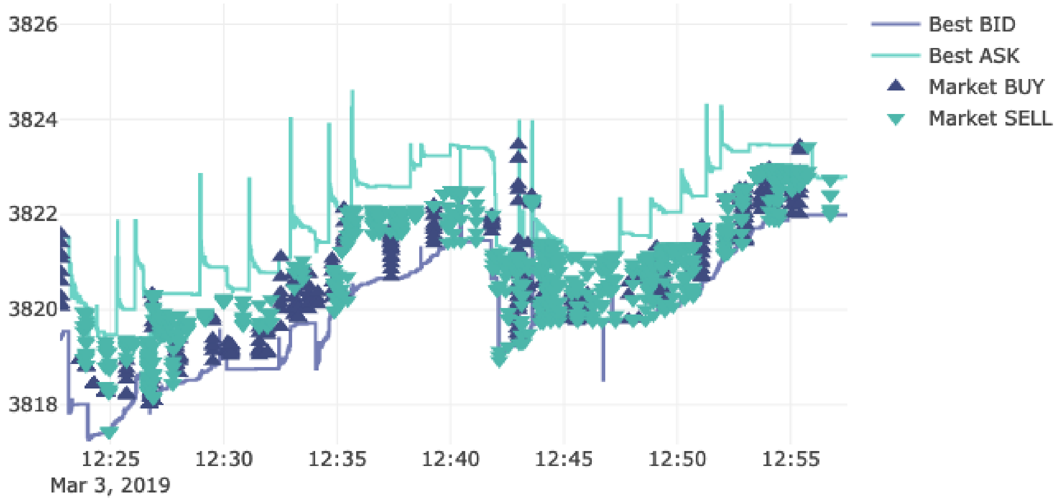
Figure 3. Identification of wash trading BTC_USDT on Huobi



(Source: computed using code from CryptoIntegrity, data gathered directly from exchange)

The data from Huobi does not let us make any conclusion because of the low frequency of updates, which hides true history of orders and price movements.

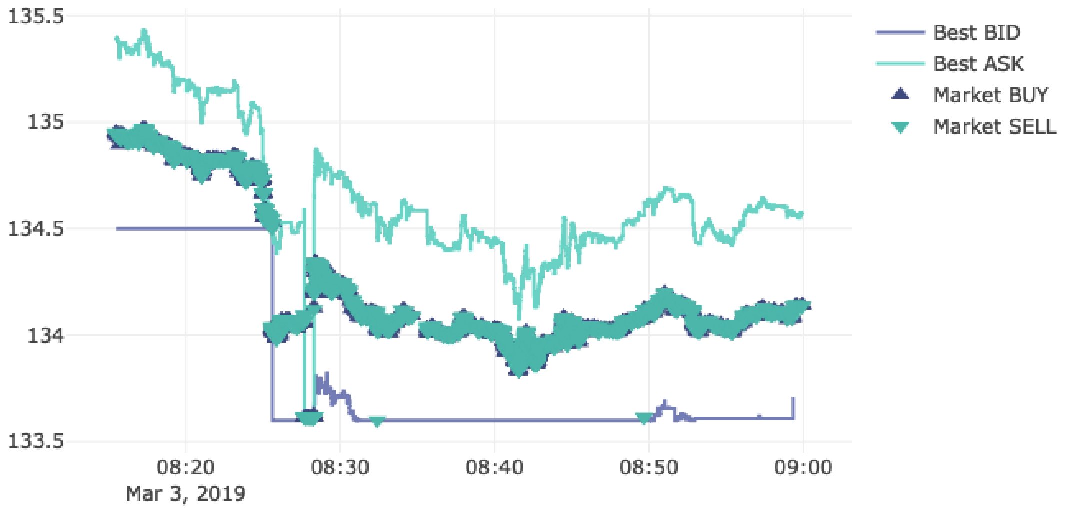
Figure 4. Identification of wash trading BTC_USDT on ZB



(Source: computed using code from CryptoIntegrity, data gathered directly from exchange)

ZB.com show both in-spread and out-of-spread trades.

Figure 5. Identification of wash trading ETH_USDT on BW



(Source: computed using code from CryptoIntegrity, data gathered directly from exchange)

BW gives an example of the simplest mid-market in-spread trades.

Summary

Cryptocurrencies have experienced a noticeable rise and received a lot of attention over the past few years. These digital currencies are based on blockchain that is able to provide payments verification without a centralized custodian. Bitcoin came into existence in 2009, launched by anonymous person or group of persons, called Nakamoto. Over the past 10 years since that time, the cryptocurrency market has changed a lot. A lot of new coins have appeared, some of them are quite different from Bitcoin from a technological point of view. The total number of cryptocurrencies at the moment is more than two thousand, according to the Coinmarketcap website. All these altcoins are traded on more than 200 digital exchanges across the world. The average daily trading volume of Bitcoin is more than $20 billion as of May 2019, and for the entire cryptocurrency market, this figure is $ 65 billion. However, according to some sources, these figures might be significantly exaggerated (SEC, 2019); this topic will be analyzed further in the paper. It is estimated that the amount of active traders exceeds 15 million, including both retail and institutional investors (such as DRW, Jump Trading, or Hehmeyer Trading).

While the cryptocurrency market is still young, it provides many opportunities for economic research. This paper touches the topics of cross-exchange arbitrage, the efficient market hypothesis, order book liquidity and wash-trading. First, in different geographical areas and jurisdictions, there are many non-integrated digital exchanges that operate in parallel. Most of them are almost unregulated and owned and managed privately. Most of these exchanges operate as ordinary stock markets, where traders place orders, and the exchange clears transactions based on a centralized order book. However, cryptocurrency exchanges also have many differences from stock markets. For example, there is no guarantee of the best price, as provided on traditional markets by the Securities Exchange Commission's NBBO rule. The National Best Bid and Offer rule helps retail investors who may not have the capacity to compare prices on several exchanges, providing them with the best price for the submitted order. The absence of such mechanisms implies that the comparison of prices on different markets and the subsequent selection of one of them for placing an order lies on the shoulders of market participants. Secondly, the stock exchanges are scattered around the world, and today the largest and the most liquid ones are located in Europe, Asia (Hong Kong, Japan, Korea) and the United States. But between many countries there are barriers to the free movement of capital, and the exchanges themselves in some cases do not allow foreign citizens to open an account. Such market segmentation creates opportunities for large players who have both a large amount of funds and opportunities for the movement of capital between countries.

Data

The main data for this research is minute timeframe prices data obtained from different sources, among which Enigma database,[44] using its specialized Python library, called Catalyst, and data obtained directly from exchanges using APIs. Catalyst is an algorithmic trading library for crypto assets. It has several features among which access to historical data of daily and minute resolution, possibility to backtest trading strategies, and provision of analytical tools regarding a particular strategy's performance. Generally, their data is viewed as reliable in this paper, but where available, confirmations of the quality of the data by comparing them with the data reported by the exchanges to bitcoincharts.com are provided. Also, data from an independent resource, coinmarketcap.com is used. It aggregates information on trading volumes by exchange and by coin.

The variables contained in the data are essentially open, high, low, close prices of the cryptocurrencies for a given timeframe and volume traded in this period on the given exchange. For shorter time periods (from February 2019 to May 2019) more frequent data was used: limit orderbook snapshots and its updates parsed directly from exchanges. Usually frequency of updates of orderbooks by exchanges lies somewhere between several times in one second or once in a second for some particular exchanges.

When it comes to choosing cryptocurrency exchanges, unfortunately, anyone, who would like to do a research or just trade cryptocurrencies, faces several problems. Wide audience of speculators and projects seem to care only about one metric regarding exchanges which is trading volume. Therefore, exchanges tend to overstate their reported trading volumes. Most of the cryptocurrency exchanges execute trades in a centralized database and can therefore trade the same coins back and forth between two bot accounts, not providing any real liquidity into the market. Thus, people or companies who want to trade cryptocurrencies receive biased picture of the real liquidity on exchanges, if they sort them just by volume traded (as it is done on the Coinmarketcap.com), and therefore, make wrong choices. Users join an exchange and think that they can buy or sell rather big amounts of cryptocurrency without affecting prices, which is apparently not the case, because there might be only a few bots trading on the platform. Of course, volume traded is not a bad measure overall, but it should be used together with the other metrics.

[44] https://enigma.co/catalyst

Liquidity

It is hard to overestimate the importance of liquidity and trading volume indicators a trader or investor. Non falsified, true trading volume is an insight into the intrinsic value of the asset. It can be used as a leading indicator of future market price of an asset by traders. In theory, market price of a bitcoin or any other cryptocurrency (or stock, commodity, etc.) is a result of interaction of supply and demand. Demand, in turn, can be driven by different factors. One of them is utility, or in other words how many people will use it in future. Another factor affecting demand is investment value of an asset, or how many people will buy it, hoping for future cash flows or market price increase. In short, the demand should reflect people’s interest in the underlying technology, its future prospects and its mass adoption among investors. The higher the interest, the higher the volume. Further, the trading volume is a function of orderbook liquidity among other factors. To formulate it mathematically:

$P = f(V, p_1, \dots, p_n)$, where P is market price, V is the volume, and p_i are some other parameters;

$V = f(L, k_1, \dots, k_n)$, where L is order book liquidity, k_i are some other parameters.

Things considered, understanding of an asset's true liquidity and volume traded is necessary to make a balanced investment decision. Although traditional economic markets have a set of techniques and regulations to avoid distortion of market data, cryptocurrency transactions remain susceptible to fraudulent activities that mislead shareholders and traders. In controlled financial markets, there are several fraudulent trading activities that are forbidden. Wash trading and disruptive activities like flipping, spoofing, quote stuffing are the most prevalent examples. This part of the research is an attempt to detect exchanges on wash trading as the cryptocurrencies world's most prevalent fraud.

The three key techniques — from the simplest to execute (and identify) to the hardest:

1. in-spread trades without limit orders;
2. in-spread trades with short-lived limit orders;
3. trades near bid and ask prices using short-lived limit orders.

The first of these practices is the most used one, especially among fresh new exchanges, and its mostly because it is easy to implement and almost riskless. An exchange that implements this method claims a trade while the order book is completely unchanged. This practice bears no risk because, even for fractions of seconds, there are no limit orders in the order book, so other traders cannot make counter orders to them. Only the exchange itself can sustain this kind of trading activity.

In-spread trades with short-lived limit orders. Anyone can create a limit order with price between bid and ask prices. Such activities make the spread on the instrument smaller and add liquidity to the market. This manipulative technique, on the other hand, implies that only orders with very short lifetime in the order book (milliseconds), so that a human is not only unable to meet them but also to just notice them in the order book. Not all of the short-lived limit orders are malicious, practically speaking. There are algorithmic traders and arbitrageurs who track all oncoming orders and rapidly meet those which bring them the chance to trade. But there is no justification for the dominance of transactions induced by such orders.

Trades near bid-ask caused by short-lived limit orders. Advanced wash trading algorithms are adjusted to generate sell orders close to the best bid offer and vice versa buy orders near the best ask to make the trading flow look more natural. But there is no intention of wash traders to genuinely trade with anyone. They set limit orders for fractions of a second and then immediately cover them with counter orders in order to minimize this risk. Naturally, a limit order of 1 BTC hit after, say, 50 milliseconds by a market order of 1 BTC can readily be categorized as a wash trade. The fakers are therefore trying to add distortion to their operations:

- divide limit order into several orders of different size;
- put aggressive orders of bigger or less size than the size of the limit;
- introduce noise to life of duration of limit orders;
- create aggressive orders with different lifetime (can be the case if the size of an aggressive order is bigger than the size of a resting one, then the not-filled part of the aggressive order becomes new resting order).

The algorithm calculates 3 metrics of liquidity for the most liquid trading instruments on every analyzed exchange:

- Handy liquidity defined as the cumulative volume in an order book at prices further from the mid-market (average of bid and ask) price by 0.5% or less. The higher the liquidity of an asset, the higher should be the handy liquidity.
- Bid ask spread, defined as:

$$\frac{(best\ ask - best\ bid)}{\frac{best\ ask\ +\ best\ bid}{2}} * 100$$

The higher the liquidity, the narrower the bid ask spread.

- Bid ask spread by 10 BTC, calculated as:

$$\frac{(ask -\ bid)}{\frac{(ask + bid)}{2}} * 100,$$

where ask and bid are weighted by aggregated volumes of 10 BTC average prices. This metric should be lower for liquid assets.

The algorithm examines market data collected using APIs of every exchange in the list. Data used is order books and trades, some of the exchanges also provide the updates of order books. The rules by which the system believes that transaction is suspicious (wash trade):

1. If a trade takes place at a price lower than the best bid or higher than the best ask price, it is classified as “out-of-spread trade”;
2. If a trade takes place at a price lower than the best ask or higher than the best bid price, it is classified as “in-spread trade”;
3. If trade happens at bid or ask price, it is referred to as normal trade.

Sum of the out-of-spread and in-spread trading volumes combined make up total artificial trading volume. The overall statistics are calculated as follows:

1. Share of total artificial volume and of total artificial trades (quantity) per trading pair per each exchange examined is calculated;
2. Arithmetic mean of the shares of total artificial trading volume and total artificial quantity of trades across all selected trading pairs of all exchanges examined is calculated.

The algorithm can identify fake trades when there is no wash trading performed (Type I error) due to probable problems with data gathering and time synchronization. Coinbase Pro was regarded as a benchmark for 100% organic trading flow, although the algorithm discovered minor anomalies.

Results

The summary of all kinds of instruments which produce suspicious (wash) trading are presented in Table 1 below.

Table 1. Identified mechanisms of suspicious (wash) trading activity

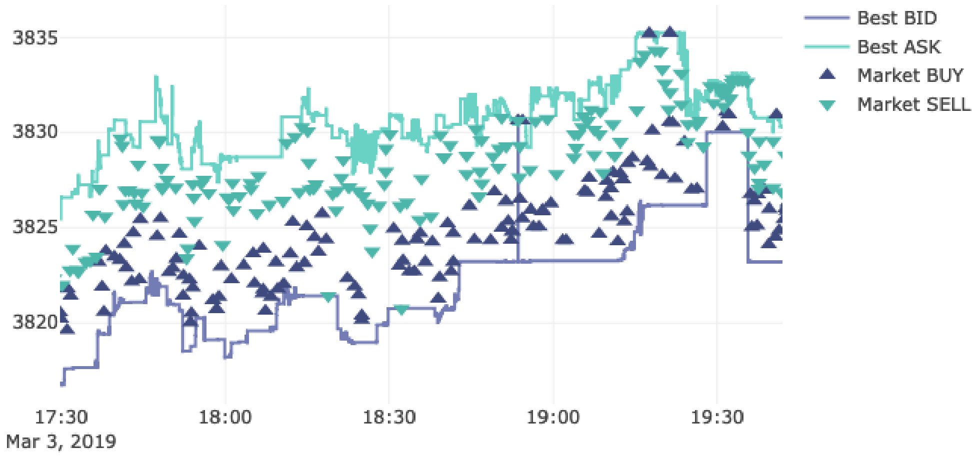
	In-spread trades without limit orders	**In-spread trades with short-lived limit orders**	**Trades near bid-ask caused by short-lived limit orders**
OKEx	●		●
Bit-Z	●		
ZB.COM	●		
CoinBene		●	
LBank	●		
Huobi Global	n/a	n/a	n/a
BW	●		
HitBTC	n/a	n/a	n/a
IDAX	●	●	
BitMart	●	●	
CoinTiger	●		

Complexity

lower *higher*

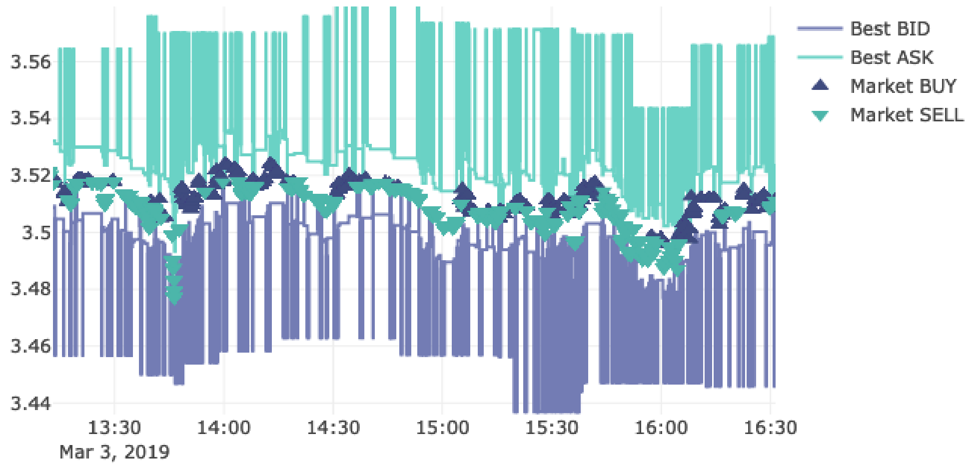
The graphs below demonstrate some cases of suspicious trading behavior with a brief remark on the chosen markets

Figure 1. Identification of wash trading on LBank exchange (BTC_USDT pair)



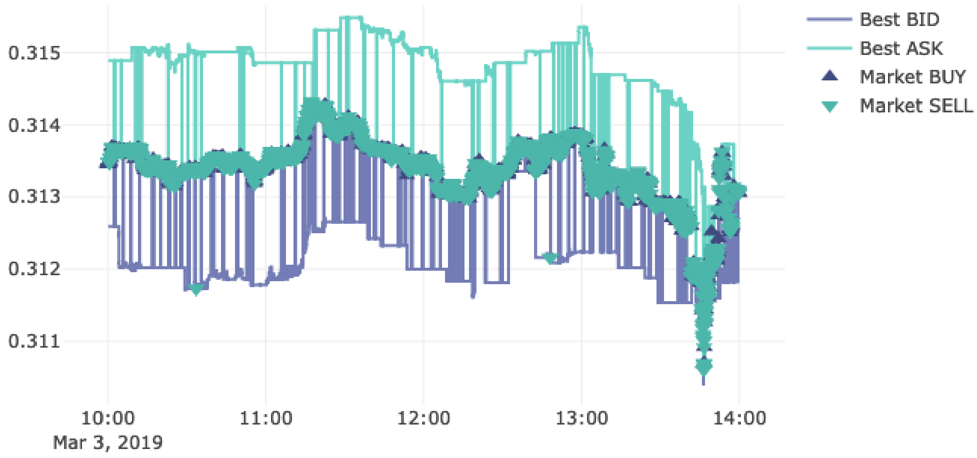
LBank orderbook data shows a lot of in spread-trades at random prices within the spread.

Figure 2. Identification of wash trading on Bitmart (EOS_USDT pair)



BitMart shows two patterns of wash trading, among which trades in-spread with limit orders and without limit orders.

Figure 3. Identification of wash trading on IDAX (XRP_USDT pair)



IDAX orderbook shows big spreads which is not a good sign in terms of liquidity, and the simplest wash trading activity - mid-market in-spread trades with short lifetime limit orders.

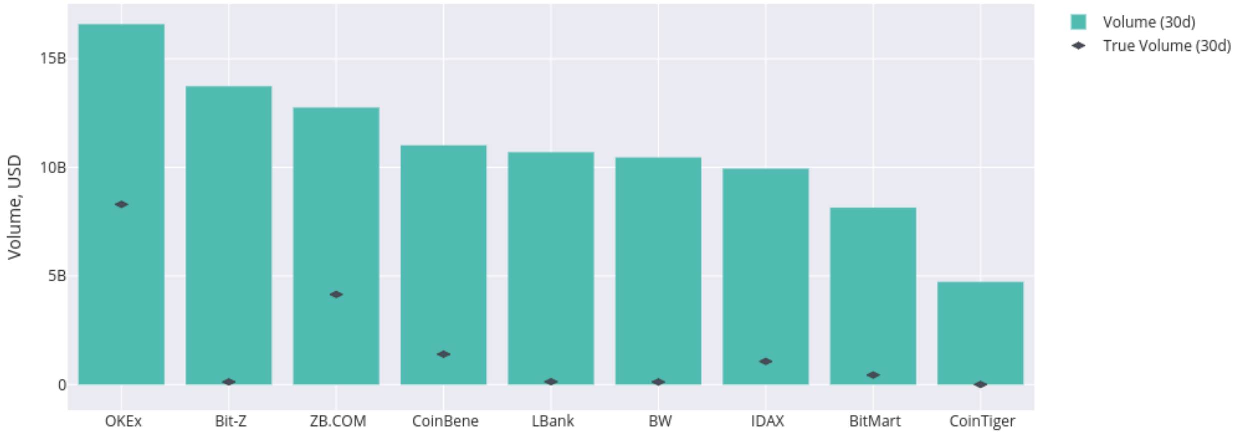
Table 2. Summary of liquidity by exchange. Green means high liquidity, red – low. Coinbase is regarded as a benchmark.

BTC/USDT	Handy liquidity mean, BTC	Spread 10 BTC, mean, %	Bid-ask spread, mean, %
Lbank	6.13	0.7	0.21
BitMart	24.96	0.25	0.08
BW	5.44	0.24	0.12
Bit-Z	5.94	0.27	0.11
CoinTiger	0.83	0.82	0.68
HitBTC*	4.65	0.02	0.09
LiveCoin	2.23	0.84	0.35
ZB	1.24	0.1	0.05
IDAX	9.24	0.19	0.1
Huobi*	5.86	0.02	0.01
Coinbene	3.72	2.18	0.21
OKEx	11.32	0.04	0.01
Coinbase	111.34	0.03	0.003

(note: HitBTC and Huobi are marked with a star, because there are issues with the data from these exchanges, thus, results regarding them are not truly reliable)

Interestingly, during the period studied BitMart had the greatest handy liquidity. All three measures had excellent estimated scores for OKEx. However, further assessment may be necessary as the liquidity presented may not be equivalent to the liquidity that is genuinely available. It is alleged that some fraudulent cryptocurrency exchanges display limit orders of special kind that other traders cannot meet. This is another limitation of this research.

Figure 4. Reported and true trading volume by exchange



Finally, Figure 2-84 represents the scale of trading volume manipulations by some cryptocurrency exchanges. This analysis has reduced the number of exchanges in the short list to 9, which are the most liquid and have significant trading volumes. In the next section, based on the data obtained from these exchanges, an analysis of arbitrage opportunities will be conducted.

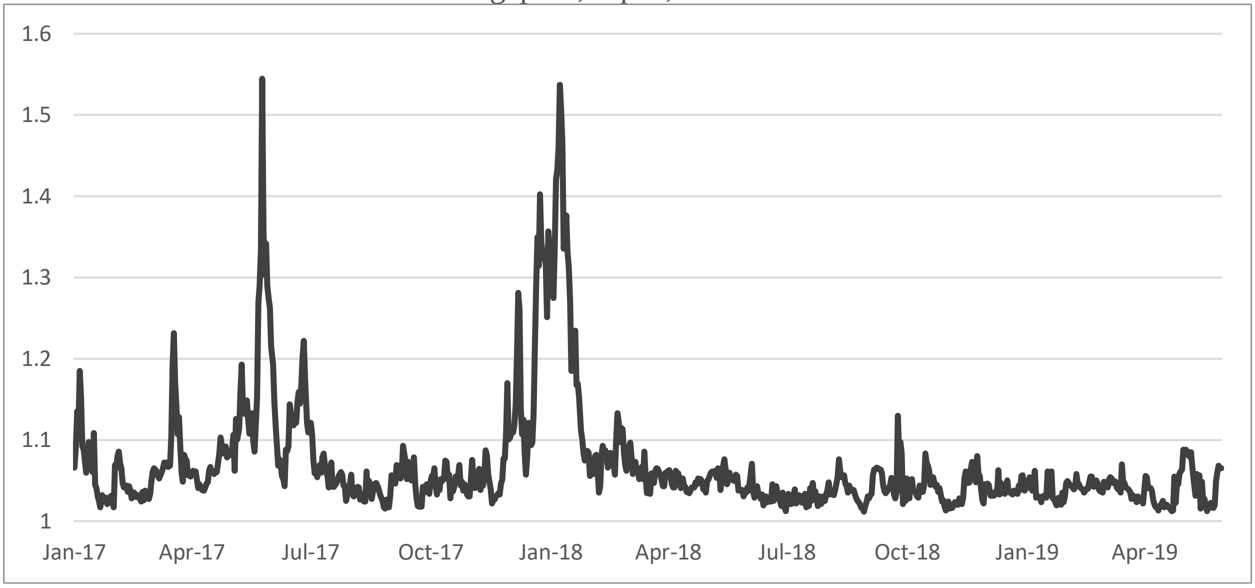
Arbitrage opportunities

Decentralization of cryptocurrency trading creates an interesting environment for studying exchange-wide arbitrage. The cross-correlations computed in the previous chapter indicated that cryptocurrency markets may not be completely effective. To quantify the price differences across exchanges at every time period, arbitrage index is computed. Its essence is that it measures the maximum price difference across given exchanges. To begin, the index is computed using minute timeframe data. It is calculated using volume-weighted price for every exchange in the sample. The index is computed by dividing max price across all short-listed exchanges for a given minute by min price. To clean the data from shocks, any movements of more than 10% between two neighboring transactions are excluded. At last, to decrease the effect of intra-day price fluctuation, we compute daily average of minute-level indexes.

Resulting index must be equivalent to 1 all the time if the markets are efficient and there are no arbitrage opportunities. Firstly, arbitrage index is computed using the exchanges from the previous parts, the sample is quite representative, there are exchanges from all over the world: United States, Europe, Hong Kong, Singapore, Korea, Japan, Australia. There are no exchanges from China, because many exchanges from there moved their headquarters to other countries due to strict regulations and ban on cryptocurrency exchanges in 2017 by Chinese government.[45]

[45] See https://www.reuters.com/article/us-china-bitcoin/chinas-okcoin-huobi-exchanges-to-stop-bitcoin-withdrawals-idUSKBN15P0HE

Figure 5. Arbitrage index. Regions included are United States, Europe, Korea, Hong Kong, Singapore, Japan, Australia.



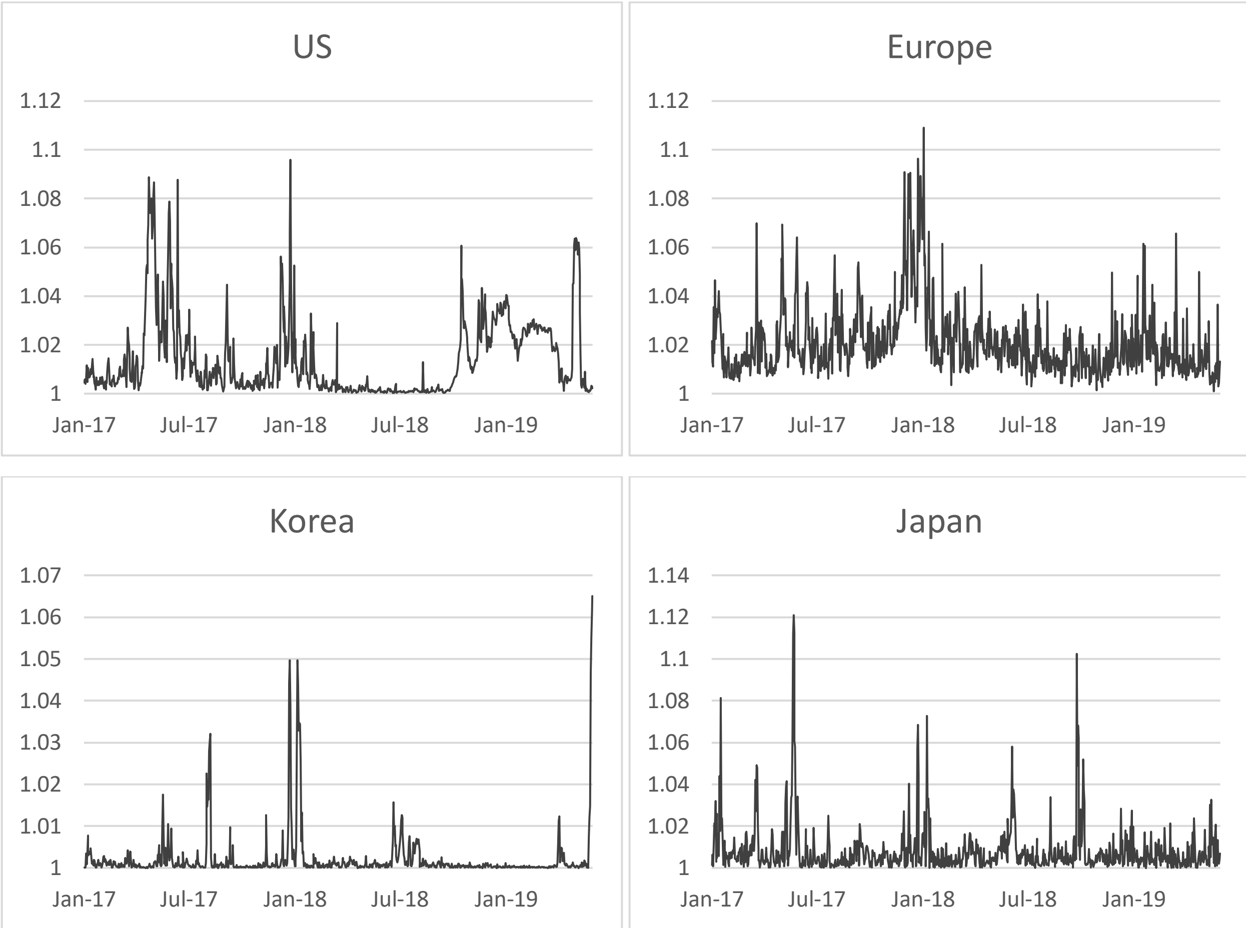
(Source: computed using data gathered from Catalyst library and particular exchanges API)

Figure 2-105 above indicates that the arbitrage index value is almost always above 1 for the period since January 2017 until May 2019. Average value of the index across this period is 1.07, median 1.05 and maximum value is 1.54. This implies that the distinction between the prices across various markets was 7% on an average day. Figure above also shows that important variation occurs during the analyzed period. It is noteworthy that there are several months in this period in which the arbitrage index remains constantly remains higher than 1.1, for instance December 2017 to February 2018.

It is not included in the graph, but the value for the index for the period since year 2014 was also computed. Interestingly, the arbitrage opportunities are much smaller during previous years, even if we take into account Chinese exchanges. Average of the index in 2016, for example, was 1.03 as opposed to 1.07 in 2017 and the highest peaks happen in June and December. These findings indicate that the excessive possibilities for between-exchanges arbitrage opened up when volatility (as well as prices) raised significantly in 2017.

To examine further what are the key drivers of the arbitrage opportunities, the arbitrage index is computed across geographical regions (instead of within regions). Firstly, the index is computed for every geographical region in which more than 1 cryptocurrency exchange is present in the data, these would be Europe, US, Japan and Korea. For every region stated above, there is data for about 3-4 major exchanges, except for Korea, which has only 2, with proved liquidity, and time earned reputation.

Panel 1. Arbitrage index within geographical regions.



(US exchanges: Gemini, Kraken, Bitstamp, Coinbase Pro, currency used - USD. Japanese exchanges: Zaif, Bitflyer and Quoine, currency - JPY. European exchanges: Kraken, Bitstamp and Coinbase: currency - EUR. Korean exchanges: Bithumb and Korbit. KRW)

In the Panel 1 above, the graph on upper-left represents the arbitrage index computed on data from 4 major US cryptocurrency exchanges from the beginning of 2017 until May 2019. The approach was the same as for the previous case with all exchanges in one sample. The graph still has several spikes almost reaching 1.1, but on average variation of price is much smaller. Average value of the index is only 1.014. Spikes are about the same time as in the previous case: May, June and December 2017, but range is about 1.06-1.08. Although the spread is much smaller, compared to the total index, it is still huge if we compare it with stock market spreads, for example.

Table 3. Descriptive Statistics for Arbitrage Index in different regions

Arbitrage Index	**Average**	**Median**	**Maximum**
United States	1.014	1.007	1.096
Europe	1.021	1.017	1.109
Korea	1.002	1.000	1.065
Japan	1.008	1.005	1.121
Overall	1.069	1.048	1.545

(Source: computed using data gathered from Catalyst library and particular exchanges API)

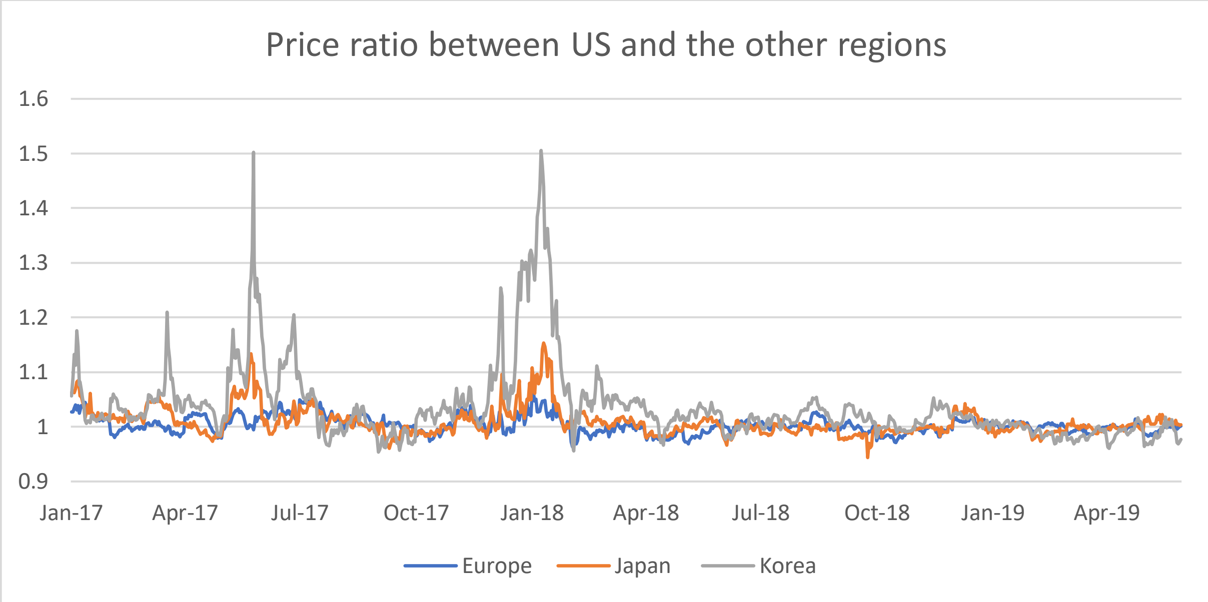
Then the same approach was repeated for the European cryptocurrency exchanges arbitrage index for the same period. Upper right graph shows that price variation within European exchanges is much lower than the arbitrage index as a whole. The average index is 1.021, median is 1.016. Compared to US the index for Europe has more spikes, and on average it is higher than for US. And again, a comparable image appears in Japan, where the arbitrage index on average is lower than for both US and Europe (1.008). In January and May 2017, and September 2018 there are brief intervals when the index moves to about 1.08. The peak in winter 2017 is similar to the trends in the other areas, but January 2017’s rise is peculiar to Japan only.

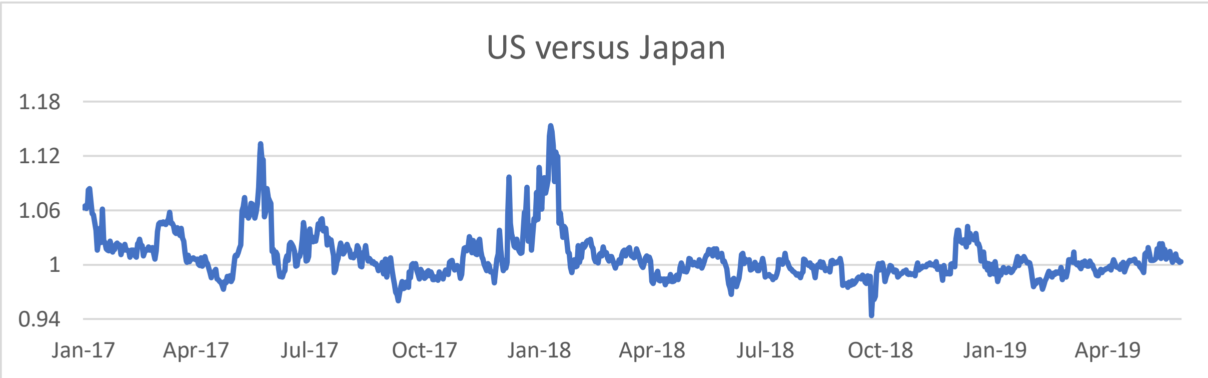
And lastly, we're looking at Korea, for which there is data for only two exchanges. Here the picture is different from other regions, we analyzed. The arbitrage index has also common peaks in September and in December 2017, but the range is much smaller than for other regions. At most it reached 1.05. Also, there is noticeable recent spike in May 2019, which is unique to Korean exchanges only. The index computed on Korean exchanges has lower average, mean and max values than for any other regions. Although, it might be due to lack of exchanges.

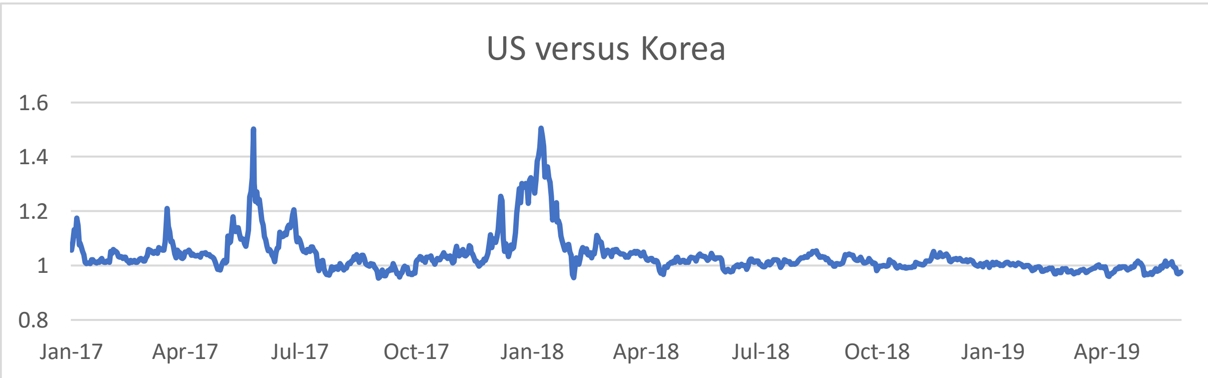
Overall, the findings indicate that the possibilities for arbitrage trades within geographical areas are much lower than between them. This implies that the most noticeable price variation is likely driven by country-wide price variation. The findings therefore indicate that crypto exchanges within specified areas appear to be much greater embedded than across areas.

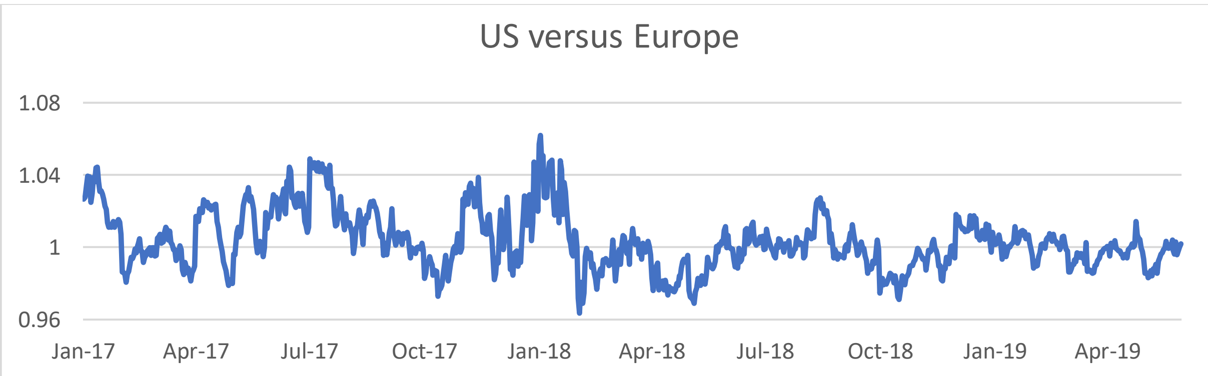
In order to verify the assumption that a substantial portion of the arbitrage spread is powered by price variation across geographic areas, the ratio of prices between United States exchanges and every other geographic region exchanges at minute timeframe is plotted in Panel 2 below. Ratios are calculated on the volume-weighted prices using minute timeframe and then averaged at the daily level. The graphs represent that prices on Korean crypto exchanges were two times 50%+ higher than on US exchanges and had the largest spread overall. The fact that Korean crypto exchanges had such a big premium during 2017 even has the special name - "Kimchi premium".

Panel 2. Price ratios across regions. These graphs plot the average price ratio between the price of bitcoin to USD across several regions from the beginning of 2017 until May 2019.







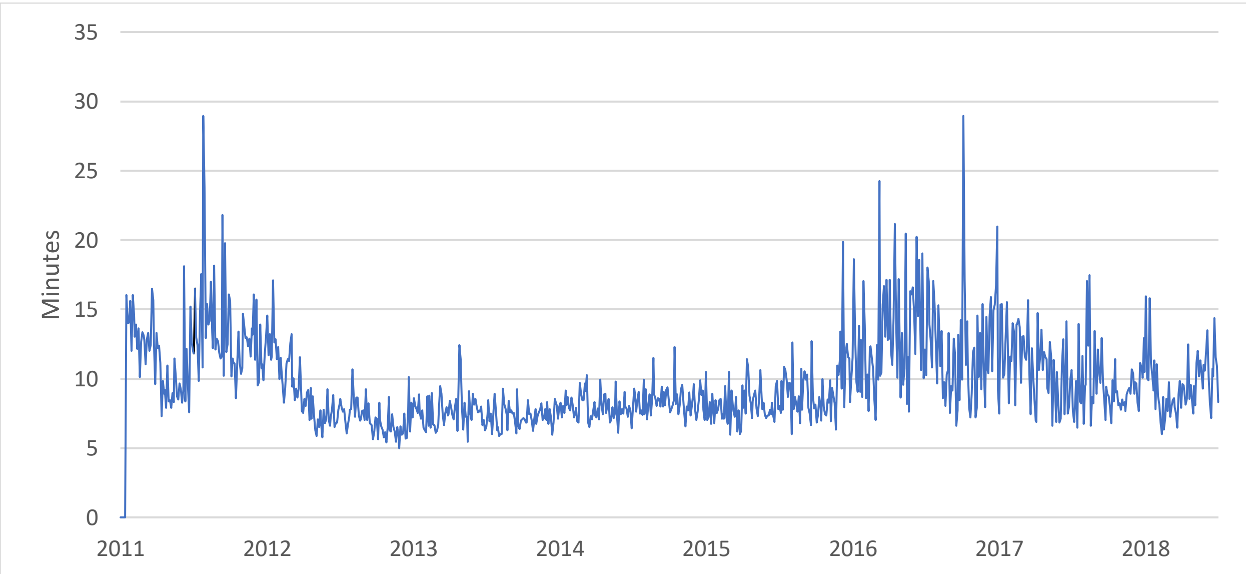


At the same time as when Korea had peak spread, Japan also had significant bitcoin price deviations from the US. But Japan's bitcoin price premium to the US had a peak of about 1.15. In January 2018 Europe, on the other hand, experienced the largest price distinction to the United States of around 6% above US rates. Compared to other areas, the price gap between the United States and Europe is low, which is not strange given that the same exchanges operate in both US and Europe. The findings indicate that a large percentage of the huge arbitrage spreads calculated for bitcoin are dictated by region-wide price variations. And in many cases, these variations persist over rather long time periods.

Arbitrage trading strategies

Consider the case discussed above - Kimchi premium, the bitcoin price in Korea is above the average price in the United States. This scenario would be a risk-free arbitrage if there were no frictions. Trader could purchase bitcoins in the United States, send them to Korea, sell bitcoin to Korean Won, then exchange KRW for USD, and make a transfer back to the United States. This theoretical arbitrage trading is impossible in reality because the nature of bitcoin transactions assumes that it takes some time to record the transaction on the blockchain (about 10 minutes on average, see Figure 2-14). In addition, it usually takes cryptocurrency exchanges from several hours to few days to transfer fiat currency. Obviously, arbitrage opportunity will not last for that time. As a default, an arbitrage trader must purchase bitcoin on the exchange with lower price and sell it on the exchange with a higher price at the same moment in time in order to lock in the arbitrage.

Figure 6. Median confirmation time of bitcoin transaction in minutes (averaged weekly)



(Source: Blockchain.com)

Optimally, the trader would like to short the bitcoin on the exchange with a higher price, for example Korea, and long bitcoin in the United States. Then the trader could send the bitcoin to the United States and lock in the risk-free profit. Though, this arbitrage trading approach is frequently not viable, as there are not so many cryptocurrency exchanges that enable short selling. In case of impossibility of short selling, the trader may switch to two alternative arbitrage trading approaches.

The first method is to get negative exposure in bitcoin by margin trading, which is almost the same as short selling, except that it doesn't allow for physical settlement. If the trader sticks to this method, he will be able to realize arbitrage profit only if the prices on the exchanges intersect in the future. The trader is therefore subject to the risk of convergence, which has been researched widely in the limits to arbitrage literature. While theoretically exchange-wide prices might not intersect for a long period of time, graphs in the Panel 2 demonstrate that in reality, arbitrage opportunities on cryptocurrency markets existed for about several days and never lasted for more than two months.

Another method is to keep some amount of bitcoins on both exchanges and purchase and sell them at the same moment in time on the appropriate exchanges when the arbitrage opportunity appears. Of course, the trader's bitcoin balance will decrease on the exchange where bitcoin's price was larger (because that's where he sells bitcoin) and rise on the exchange where the price was smaller. The arbitrage trader then should send bitcoins from the elevated bitcoin balance exchange to the low balance exchange to equalize amounts across two exchanges. While this approach doesn't make an exposure to the risk of convergence, a major disadvantage of this approach is that the trader has exposure in bitcoin to fiat. To minimize that risk, the arbitrage trader could borrow bitcoin from people who have big amounts of cryptocurrency and are not planning to sell it in the near future, the so called hodlers. Obviously arbitrage trading can suit hodlers themselves, because of their constant exposure to cryptocurrencies. The trader might also use futures on bitcoin to hedge the risk.

In reality, the arbitrage trader will bear several transaction costs, but comparing to the possible profits, their magnitude is way too small to discourage traders from arbitrage. Ultimately, the summed-up trading fees for big traders should be between 0.25% and 0.50% or may be even less. Relative to the possible arbitrage yields, reported in the previous section, these costs are insignificant and thus cannot explain the existence of such arbitrage spreads.

Another aspect that could restrain traders' readiness to participate in arbitrage deals is the governance risk of exchanges. It occurs because the trader has to send his bitcoins to the exchange wallet and thus lose control of the cryptocurrency in favor of the exchange. Judging from several commonly published exchange exploits or simply dishonesty and fraud may result in substantial losses for investors trading there. A recent example is when the hackers stole $40 million worth from Binance exchange. In this case, however, the head of the exchange promised to reimburse to the clients all the losses.

It seems doubtful, though, that all the above-mentioned issues might explain the discovered arbitrage spread. Doubts about an exchange's governance risk should influence its trading volume and potentially spread. But the exchanges chosen for analysis already show high liquidity and low spreads. It was also demonstrated that the spreads are much bigger across geographical regions. To explain this trend by exchange risk, one would have to suppose that the exchange risk is correlated within a region. However, it is not backed by the data as there is considerable heterogeneity in the exchanges' liquidity within a particular region, but nonetheless spreads between them are low.

At last, cross-border capital controls are a serious limit to arbitrage trading. As it was mentioned earlier, unless the trader wants to bet on future convergence in prices between Korean and United States exchange, then he would have to sell cryptocurrency in Korea for fiat currency and transfer it back from Korea to the United States. The government regulation of some countries makes it hard, especially for individual investors, to make cross-border deals. For instance, Korean residents and businesses moving an equivalent of more than $50,000 out of Korea in a single year, must submit documentation to officials showing their justifications for the transactions, and these transactions may not always be approved and permitted. Market studies as well as trading blogs indicate that these limitations are most likely binding for individual investors. However, quantifying how binding these limitations are for big organizations, which perform trading on various global economic markets is harder.

There are several studies suggesting that these limitations can be avoided by big organizations. In a latest IMF working document, Chikako and Kokenyne (2011) discover that the efficiency of capital regulations in South Korea appears to be low, as money flows into and out of Korea as well as monetary policy efficiency do not seem to change considerably after capital regulations has been introduced. Also, industry researches indicate that there are forex dealers that assist organizations to transfer capital in and out of Korea. Capital bindings should not, therefore, place insuperable limitations on arbitrage activities across geographical areas, especially for big traders, they just contribute to the costs of trading. This is backed by the fact that arbitrage spreads on the same cryptocurrency exchanges where we see large spreads to fiat currencies are much lower in two-way cryptocurrency trades. But even in this case the arbitrage opportunities do not persist for longer than two months.