universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

Trading Frictions and Arbitrage in the Bitcoin Market

verfasst von / submitted by

Yelena Holzer BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2022 / Vienna, 2022

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 913
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Master Economics/Masterstudium Volkswirtschaftslehre
Betreut von / Supervisor:	Univ.-Prof. Dr. Nikolaus Hautsch

Abstract

This thesis studies trading frictions and arbitrage in the Bitcoin market by analyzing Bitcoin orderbooks of seven fiat-to-crypto exchanges and network data. In order to provide systematic empirical evidence on the extent to which liquidity and fees contribute to the bitcoin price differences, a measure of the minutely-level, instantaneous arbitrage profit is constructed by optimally accounting for the prevailing orderbook depth and fee schedule. In the sample period from March 2018 to September 2018, more than 90% of the cross-market bitcoin price differences between the largest fiat-to-crypto exchanges Bitfinex, Bitstamp, Gemini, Coinbase Pro and Kraken do not correspond to arbitrage opportunities after adjusting the prices for the proportional taker fees. Only exchange CEX.IO exhibits arbitrage opportunities that could produce sizable and longer-lasting profits, which persist for 53 minutes, on average. Bitcoin price differences might be related to market liquidity, because liquidity can affect the efficacy of arbitrage. In the reverse direction, bitcoin price differences could affect liquidity. These ideas are explored at the exchange level by investigating the dynamic relations between liquidity and bitcoin price differences using vector autoregressions. The results show that fiat-to-crypto exchanges are heterogeneous with regard to the strength of the relation between price differences and liquidity. Nevertheless, shocks to the bitcoin price differences predict future liquidity on some exchanges. In the reverse direction, the evidence from the impulse response functions suggests that liquidity plays an important role in moving the bitcoin prices toward more efficient levels only in CEX.IO, which is a relatively illiquid market with substantial and persistent price differences.

Abstract

Diese Arbeit untersucht Handelsfriktionen und Arbitrage auf dem Bitcoin-Markt, indem sie Bitcoin-Orderbücher von sieben Fiat-zu-Krypto-Börsen und Netzwerkdaten analysiert. Um systematische, empirische Beweise dafür zu liefern, inwieweit Liquidität und Gebühren zu den Bitcoin-Preisunterschieden beitragen, wird ein Maß für den sofortigen Arbitragegewinn auf minütlicher Ebene konstruiert, indem die vorherrschende Orderbuchtiefe und der Gebührenplan optimal berücksichtigt werden. Im Stichprobenzeitraum von März 2018 bis September 2018 entsprechen mehr als 90% der marktübergreifenden Bitcoin-Preisunterschiede zwischen den größten Fiat-zu-Krypto-Börsen Bitfinex, Bitstamp, Gemini, Coinbase Pro und Kraken nicht den Arbitragemöglichkeiten nach Anpassung der Preise unter Berücksichtigung der anteiligen Abnehmergebühren (taker fees). Nur die Börse CEX.IO weist Arbitragemöglichkeiten auf, die beträchtliche und länger anhaltende Gewinne erzielen könnten, die im Durchschnitt 53 Minuten andauern. Bitcoin-Preisunterschiede könnten mit der Marktliquidität zusammenhängen, da Liquidität die Wirksamkeit von Arbitrage beeinflussen kann. Umgekehrt könnten Bitcoin-Preisunterschiede die Liquidität beeinträchtigen. Diese Ideen werden auf Börsenebene untersucht, indem die dynamischen Beziehungen zwischen Liquidität und Bitcoin-Preisunterschieden mithilfe von Vektorautoregressionen untersucht werden. Die Ergebnisse zeigen, dass Fiat-zu-Krypto-Börsen hinsichtlich der Stärke des Zusammenhangs zwischen Preisunterschieden und Liquidität heterogen sind. Dennoch sagen Schocks der Bitcoin-Preisunterschiede die zukünftige Liquidität an einigen Börsen voraus. Umgekehrt deuten die Resultate aus den Impulsantwortfunktionen darauf hin, dass die Liquidität nur bei CEX.IO, einem relativ illiquiden Markt mit erheblichen und anhaltenden Preisunterschieden, eine wichtige Rolle dabei spielt, die Bitcoin-Preise in Richtung effizienterer Niveaus zu bewegen.

Contents

List of Figures

List of Tables

Nomenclature

$\Delta_{it}^{A(B)}$ Difference in ask (bid) prices for minute t and exchange i

$\delta_{t,1}^{ij}$ Instantaneous first level price difference between exchange i and j at time t

$\delta_t^{ij}(Q)$ Instantaneous unit price difference of the Q-th bitcoin at time t and exchange i

$\tilde{\mathcal{A}}_t^i$ Set of active sell orders at exchange i

$\tilde{\mathcal{B}}_t^i$ Set of active sell orders at exchange i

$\mathcal{A}_t^i$ Ask-side depth profile $(p_k^{A,i}, q_k^{A,i})_t)$ for exchange i, where a quantity $q_k^{A,i}$ is a sum of all quantities q_{x^i} of orders x^i, submitted at time t_{x^i}, that correspond to the same price p_k^i - price at the level k

$\mathcal{B}_t^i$ Bid-side depth profile $(p_k^B, q_k^B)_t)$ for exchange i, where a quantity $q_k^{B\,i}$ is a sum of all quantities q_{x^i} of orders x^i, submitted at time t_{x^i}, that correspond to the same price p_k^i - price at the level k

$\mathcal{L}_t^i$ Limit Order Book or a set of active orders in a market i for a minute t

ϕ Likelihood of arbitrage opportunity

$\pi_t^{ij}(Q)$ Total instantaneous arbitrage profit from trading Q bitcoins for exchange pair ij at time t and exchange i

$\tilde{\delta}_{t,1}^{ij}$ Instantaneous taker fee-adjusted first level price difference between exchange i and j at time t

$\tilde{\pi}_t^{ij}(Q)$ Instantaneous arbitrage profit adjusted for the exchange-specific taker fees at time t and exchange i

$\widehat{\sigma}_{it}^2(h)$ Spot variance estimator in minute t and exchange i, on day $T-1$ based on bandwidth h

a_t^i Best ask price, defined as the highest stated price among active ask orders $\mathcal{A}_t^i$ at time t and exchange i

$AvgD_t^i$ Average quoted depth, defined as an average quantity that the trader can trade at the best prices at time t and exchange i

b_t^i Best bid price, defined as the highest stated price among active buy orders $\mathcal{B}_t^i$ at time t and exchange i

k Price level in the limit order book

m_t^i Mid-price - the average between the best bid and best ask price at time t and exchange i

OI_t^i Imbalance between buy and sell orders at time t and exchange i

$P^{eff}(Q, \mathcal{B}_t^j, f_j)$ and $P^{eff}(Q, \mathcal{A}_t^i, f_i)$ Selling and buying prices of Q bitcoins after subtracting and, respectively, adding exchange-specific taker fees at time t and exchange i

$P^{nom}(Q, \mathcal{A}_t^i(\mathcal{B}_t^i))$ Nominal buying (selling) price of Q bitcoins at at time t and exchange i

p_{x^i} Price at which order x is submitted at exchange i

$PD_{it}^{(B/A/M)}$ Cross-sectional average hourly price (first level (bid/ask/midquote) price) in hour t and exchange i

PS_t^i Proportional bid-ask spread at time t and exchange i

Q Amount of bitcoins traded

q_{x^i} Quantity at which order x is submitted at exchange i

QS_t^i Quoted bid-ask spread or the difference between the best ask price and the best bid price at time t and exchange i

t_{x^i} Submission time of order x at exchange i

x^i Order $(p_{x^i}, q_{x^i}, t_{x^i})$ submitted at time t_{x^i} with price p_{x^i} and size q_{x^i} at exchange i

$dwfee_i$ USD withdrawal fee at exchange i

$wfee_i$ Bitcoin withdrawal fee at exchange i

ADR American Depositary Receipt

AIC Akaike's Information Criterion

AML Anti-Money-Laundering

API Application Programming Interface

ARIMA Autoregressive–Moving-Average Model

CET Central European Time

DLC Dual Listed Companies

FOC Fill-or-Kill Order

GARCH Generalized autoregressive conditional heteroskedasticity model

IRF Impulse Response Function

KYC Know-Your-Customer

LOB Limit Order Book

SC Schwarz Information Criterion

SEPA Single Euro Payments Area

UN United Nations

VAR Vector Autoregression

1 Introduction

In recent years, cryptocurrencies have been a growing asset class which have received much attention from investors and academics due to their innovative design. One concern with any cryptocurrency as a decentralized settlement system is a double spending problem - the risk that the same single digital token can be spent more than once (Chochan, 2017). Satoshi Nakamoto proposed a solution to the double spending problem in his paper titled "Bitcoin: A peer-to-peer Electronic Cash System", which was published in October 2008. The first bitcoin software was released by Satoshi Nakamoto in January 2009. Thus, bitcoin became the earliest decentralized digital asset and payment system that used cryptography and information technology to control its creation and government without the need for intermediaries. Today more than 50 million active investors worldwide trade bitcoin and the total market capitalisation of bitcoin has reached 1 Trill. US dollars (Pound, 2021).

Prices of bitcoin are expressed over online cryptocurrency exchanges which allow customers to trade bitcoin for other assets, such as fiat money or other digital currencies. Similar to traditional trading venues, the cryptocurrency exchange's matching engine pairs up market buy and sell orders with asks and bids containing in an order book. However, according to Makarov and Schoar (2020), a distinctive feature of the cryptocurrency market is an absence of any such regulations, as, for example, the US Securities and Exchange Commission (SEC)'s National Best Bid and Offer (NBBO) regulation, that ensures that traders receive the best possible price by comparing quotes on multiple exchanges. Therefore, arbitrage activity becomes of even greater importance to the cryptocurrency market (Makarov and Schoar, 2020). Generally, the arbitrage process in the Bitcoin market is different from traditional financial market in three ways. First, the cryptocurrency exchanges operate 24 hours a day, 7 days a week with available and free of charge pricing data. In contrast to Bitcoin market, the price differences of ADRs (American Depositary Receipts) may arise because opening hours of exchanges may not overlap and arbitrageur can not exploit price differences at any time. Second, bitcoin is mutually interchangeable across exchanges. The mechanics of arbitrage is therefore different from the arbitrage process in ADRs, which are not fully fungible and, the arbitrageur has to convert the ADRs at a cost (Gagnon and Karolyi, 2010). Third, bitcoin is completely identical across exchanges. This is different from arbitrage in DLC (Dual Listed Companies), which are traded in different countries, while retaining their separate legal identity. Since the prices should move in lockstep, the arbitrage strategies involve keeping an open long and short position until the prices converge (Rosenthal and Young, 1990; Froot and Dabora, 1999; Rosenthal et al., 2009).

The law of one price, which states that identical goods should sell for identical prices, is enforced by arbitrage - it brings prices to fundamental value. However, empirical research has identified significant mispricing across bitcoin exchanges. For instance, Kroeger and Sarkar (2017) identify persistent and statistically significant bitcoin price differences from January 2012 to August 2016, and report that the average absolute price difference was around 2%, but the maximum ranged between 17% and 41%. Makarov and Schoar (2020) document large and recurring deviations in bitcoin prices between December 2017 and February 2018, and show that the bitcoin price differences are larger for exchanges across different geographical regions (20% between the US and Korean exchanges) than within the same region (below 2% for the US exchanges).

Persistent bitcoin price deviations across exchanges suggest constraints to arbitrage that prevent the convergence of prices to fundamental values (see, e.g., Shleifer (2000), Barberis and Thaler (2003)). In the traditional financial markets, empirical research has identified transaction and holding costs, capital and short-sale constraints, fundamental and idiosyncratic risk as potentially important impediments to arbitrage[1]. Hautsch et al. (2020) shows that a new type of market friction - the blockchain-related settlement latency imposes limits to arbitrage by exposing cross-market arbitrageurs to the risk of adverse price movements. This market friction is an inherent feature of the blockchain as a from of distributed ledger technology. The ownership of bitcoin is validated through the blockchain - the sequential database of information that is secured by methods of cryptographic proof (Nakamoto, 2008). Then, in a situation where inventory holdings and margins are exhausted, and the arbitrageur is forced to physically transfer the asset between markets, the settlement latency implies limits to arbitrage as it exposes arbitrageurs to price risk (Hautsch et al., 2020).

The contribution of this thesis is twofold. First, a comprehensive summary of stylized facts for cryptocurrency exchanges, including measures of liquidity, bitcoin price volatility, and bitcoin price differences between exchanges, is provided. Furthermore, a measure of arbitrage profits is constructed and documented. Second, it is attempted to formally discern whether the liquidity is related to the bitcoin price deviations on the market level for each exchange. Hereby, a large dataset of minute-by-minute orderbook snapshots of seven largest cryptocurrency exchanges that feature BTC versus USD trading is used. Exchanges, maintaining an API (application

[1]See, for example, Roll et al. (2007) on transaction costs, Pontiff (2006) on holding costs, Ofek et al. (2004) on short-sale constraints, Shleifer and Vishny (1997) on capital constraints.

programming interface) feed, typically provide two kinds of bitcoin's price data: the entire trade history (completed transactions, tick data) and limit order books. Previous empirical research[2] measured price differences and arbitrage profits using tick data, i.e. prices and amounts of completed transactions. Then, the arbitrage profits are calculated with the prices at which these amounts could have been purchased and sold. However, even if the trader could exploit the price difference, it does not indicate that the arbitrage opportunity still exists, and the order book depth is large enough to continue arbitrage. Limit order books, in turn, contain information on the prices and available market depth to capture the full magnitude of arbitrage opportunities that were available to exploit. Furthermore, the arbitrage profit, constructed in this thesis, is maximized by optimally accounting for the prevailing orderbook depth and fees schedule. This approach is particularly advantageous, because the optimal trading quantity not only efficiently determines whether the order book depth is sufficient to cover withdrawal fees, but also the full extent of the arbitrage opportunities and thus violations of the law of one price. In this manner, this thesis provides systematic empirical evidence on the extent to which such market frictions as liquidity and fees contribute to the bitcoin price differences.

A measure of the arbitrage profit, constructed in this thesis, evaluates the magnitude of the profit the trader could achieve from the arbitrage strategy, where the arbitrageur holds a positive balance of bitcoins and US dollars on both exchanges and simultaneously buys and sells bitcoins across two exchanges. The arbitrageur is not forced to physically transfer the asset between markets immediately, but only when it is necessary to replenish the balance. Importantly, the results characterize arbitrage profits in a situation where inventory holdings on an exchange are not exhausted, or short-selling, trading at margin, borrowing from hodlers is feasible, and all associated costs and risks are negligible. Hautsch et al. (2020) estimate arbitrage bounds in a scenario where arbitrageurs are fully exposed to settlement latency and document that the arbitrage bounds due to settlement latency contain up to 91% of the observed price differences, adjusted for transaction costs. In this thesis, the analysis shows that the bitcoin price differences and arbitrage profits constitute two groups. In the sample period, the relative large and persistent bitcoin price deviations existed only on two cryptocurrency markets in the sample - CEX.IO and bitFlyer. Specifically, the average absolute difference in the first level bid and ask prices on CEX.IO is 0.64% and 0.68%, the maximum difference in bid prices - 4.95% and 5.66%, respectively. There are 64% and 71.26% of hours in the sample period on which the bid and ask price deviations of CEX.IO are positive, which indicates that most of the

[2] See, for example, Makarov and Schoar (2020) and Kroeger and Sarkar (2017).

time CEX.IO trades bitcoin at a premium. Bitcoin price difference on bitFlyer are smaller in magnitude - 0.21% and 0.18% for the first bid and ask prices, respectively. The price deviations are also less persistent than on CEX.IO - bitFlyer often alternates between trading bitcoin at a premium and a discount. For the remaining exchanges in the sample, bitcoin price differences are typically more episodic and much less pronounced, and they regularly switch signs. The average absolute price difference ranges from 0.12% to 0.20%, the maximum - between 0.53% and 2.49%. However, the magnitude of bitcoin price differences is still higher compared to price deviations in other markets and assets. For example, Gagnon and Karolyi (2010) report that the mean price difference for ADRs (American Deposit Receipts) is 0.049%, Wenxin et al. (2018) observed mean daily deviations for the covered interest parity that fluctuate from 0.06% to 0.19% basis points annualized.

Nevertheless, not every price difference corresponds to an arbitrage opportunity. Therefore, the instantaneous arbitrage profits are quantified in order to show if the magnitude of price differences and order book depth is large enough to cover the fees, stemming from exchanges and the blockchain network. Similar to the bitcoin price difference, the bitcoin arbitrage profits constitute two groups. First, for all exchange pairs except CEX.IO-related ones, 50%-80% of the minutely observed instantaneous first level cross-market price differences do not constitute arbitrage opportunities. In other words, for more than a half of the sample period, the markets are not crossed, i.e. the bid price at one market is lower than the ask price of the other market. Then, for all exchange pairs, excluding pairs involving CEX.IO and bitFlyer, more than 98% of the first level cross-market price differences do not constitute arbitrage opportunities after adjusting the prices for the proportional taker fees. Thus, in the sample period, the bitcoin price differences between the largest fiat-to-crypto exchanges Bitfinex, Bitstamp, Gemini, Coinbase Pro and Kraken can be almost fully explained by the exchange-specific taker fees. Indeed, the average magnitude of the positive first level price differences is 0.01%-0.1% (excluding CEX.IO and bitFlyer), which is below the taker fees ranging from 0.15% to 0.26%.

An exception is the CEX.IO- and bitFlyer-related exchange pairs with CEX.IO and bitFlyer on the sell-side. On CEX.IO - the most profitable exchange in the sample, around 60% of the first level price differences are positive and associated with 10%-35% of the minutes on which arbitrage profits are positive (when CEX.IO serves as a sell-side exchange). CEX.IO as a sell-side market is profitable for arbitrage on average 7 hours per day with average arbitrage profit from $700 to $1000 per opportunity. Price deviations that correspond to arbitrage opportunities on CEX.IO persist, on average, 53.13 minutes, the maximum duration reaches 5 days. For the remaining exchanges, the average duration of the taker fee-adjusted first level price difference is much less persistent and ranges from 7.27 to 15.19 minutes, the maximum duration - from 14.17 minutes to 21.80 hours. In comparison to traditional markets, bitcoin price deviations are more persistent than in the foreign exchange (FX) market with 1.554 seconds per day of the average duration of arbitrage opportunities (Foucault et al., 2017), but comparable to the ADR market with the average duration of price deviations of 12.41 minutes (Rösch, 2021).

Another finding is that the liquidity in the form of the order book depth plays a lesser role in explaining bitcoin price differences than taker fees, but still reduce the amount of profitable arbitrage opportunities. After optimally accounting for the liquidity at the relevant price levels of the corresponding order books, the USD and BTC withdrawal fees reduce the frequency of the profitable arbitrage opportunities by up to 50%. It means that at least in 50% of cases of positive first-level price differences, the order book depth is large enough to cover the fees.

The quantifying arbitrage show that when a profit-maximizing volume is traded, the arbitrage opportunities produce sizable profits for the trader, especially on CEX.IO, given it's significant and persistent price deviations. Also other exchanges with a low frequency of arbitrage opportunities in comparison to CEX.IO, exhibited several periods in which the arbitrage opportunities were quite persistent (12-21 hours) with moderate profits ($20-$400 per arbitrage opportunity). This suggests the existence of additional market frictions as liquidity and transaction costs can not fully explain all arbitrage opportunities. In case the restriction of the same price of bitcoin, as a homogeneous asset with the same payoffs, is violated, two conditions must be satisfied. First, there must be arbitrage limits that potentially explain why price differences persist (Shleifer and Vishny, 1997; Gromb and Vayanos, 2010). Second, there must be reasons why bitcoin prices diverge in the first place, that is, reasons for bitcoin prices to deviate from fundamental value (Gromb and Vayanos, 2010). Following Gromb and Vayanos (2010), two examples were given in the discussion on the second condition. These examples describe demand and supply shocks that arise because of institutional frictions such as counterparty risk and fiat illiquidity risk, and that had significant and long-term price effects. To investigate the first condition of limits to arbitrage and to explore the idea that bitcoin price differences might be related to the frictions associated with liquidity, the intraday (hourly) and daily horizon time-series interaction between price differences and bid-ask spread is modelled by vector autoregressions for each exchange. On the empirical side, this thesis is closest to the study of Roll et al. (2007), who examined the the dynamic relation between the futures-cash basis and stock market liquidity. When the market is illiquid and a price difference is sufficiently large, arbitrageurs may struggle to close the price gap

(Roll et al., 2007). Also, according to Kyle (1985), arbitrageurs may not instantly provide the large volume of trades, required to close the price difference, as they may choose to avoid a large price impact by spreading their arbitrage trades out over time. If price differences indicate price pressures, then inventory problems, induced by associated order imbalances may have a contemporaneous as well as a persistent effect on liquidity (Stoll, 1978a; Amihud and Mendelson, 1980; O'Hara and Oldfield, 1986; Chordia et al., 2002). On the other hand, if arbitrage opportunities are due to information (according to Foucault et al. (2017) "toxic" arbitrage opportunities), then liquidity providers, exposed to the risk of trading at stale quotes, charge wider bid-ask spreads to cover adverse selection risk (Copeland and Galai, 1983). Guided by this literature, the intraday (hourly) and daily horizon time-series relations between bitcoin price differences and liquidity is explored by formulating exchange-specific vector autoregressions (VAR), since there are reasons to expect bivariate causality, explained above. To control for the activity in the Bitcoin network, the blockchain network fees and the number of transactions waiting for verification are added to the VAR specification as exogenous variables. The next robustness check employs average quoted depth as an alternative measure of liquidity that accounts for a quantity dimension of liquidity, which is related to the second property of a liquid market stated by Kyle (1985) - deep enough market for a small price impact. Controlling for fluctuations in liquidity and in price differences driven by volatility dynamics is performed by including bitcoin price volatility in the VAR system.

The first empirical finding is that cryptocurrency exchanges are heterogeneous with respect to the strength of the relationships between price differences and liquidity. In line with Roll et al. (2007), impulse functions for three out of six exchanges (CEX.IO, Coinbase Pro and Kraken) reveal that, at the hourly level, shocks to the price differences are significantly informative in predicting future shifts in spreads. Another result, which is consistent with Roll et al. (2007), is that shocks to spreads are informative in forecasting shifts in the price differences only at CEX.IO, the relatively illiquid exchange with substantial and longer-lasting price differences, suggesting that liquidity affects arbitrageurs more in the relatively less liquid market. This effect can be captured, however, only at a daily level, the hourly horizon is not long enough for this effect to be significant at CEX.IO.

This thesis is structured as follows. Section 2 presents the structure of the cryptocurrency market, fee schedule and exchanges used in the sample. In Section 3, using bitcoin order book data, measures of liquidity, bitcoin price volatility and bitcoin price differences are introduced and quantified. In Section 4, a measure of the arbitrage profit is constructed. In Section 5, the factors which might cause bitcoin price differences and contribute to their persistence are discussed. Section 6 focuses on liquidity as a market friction, hence, intertemporal associations between liquidity and bitcoin price differences are explored by using the vector autoregression technique. Finally, Section 7 concludes.

2 Cryptocurrency Exchange Market Structure

In order to study potential risks and costs that can impede the effectiveness of arbitrage in cryptocurrency market, one needs to investigate how the market is organized. Cryptocurrencies can be acquired by several ways: a successful miner finding a new block is rewarded with newly created bitcoins (Böhme et al., 2015). The second way is to exchange bitcoins for fiat currencies and other cryptocurrencies at exchange platforms.

There are two types of cryptocurrency exchanges: *centralized* and *decentralized* exchanges. Most of established exchanges are centralized. To be centralized means that there is a trusted third party that monitors and secures assets on behalf of the buyer and seller. Centralized exchanges serve as financial intermediaries between the traders and make profits from the bid-ask spreads and by collecting the fees. In order to trade at a centralized exchange, it is required to use wallets provided by the trading platform. For example, when a new bitcoin wallet (software program, where bitcoin address is saved) is created, an address is also randomly generated. In order to send bitcoin from a particular address, the transaction must be ratified by the sender using his or her "private key", similar to the password, which fits in a certain way with that address. The private key is not known to the public. When a wallet is created on a centralized cryptocurrency exchange, the private keys to the assets are held by the exchange on centralized servers. In this way, the traders do not have access to the private keys of their wallets and their funds are held by the company running the exchange. This puts all traders' trust in the hands of a third party, since centralized operations carry a range of risks for customers that could lead to a loss of assets such as external hacks, incompetency of the exchange operators or management, bankruptcy, unexpected account freeze or closure of the exchange due to government interference (Goldenberg, 2018). Since 2011, there have been at least 56 known cyberattacks on centralized cryptocurrency exchanges (Russolino and Jeong, 2018).

Centralized exchanges match orders in the order book and transactions are executed through a centralized server to speed up trading process. Typically, the bitcoin transactions' verification takes time – from 5 to 10 minutes at least, and up to several hours, but a centralized exchange settles all trades immediately, i.e. all transactions that happen *inside* the centralized exchange are off-chain and not recorded in the bitcoin blockchain. The only transactions that get recorded in the blockchain are withdrawals of bitcoins to external wallets or deposits of bitcoins to the exchange's wallet from external wallet. As a result, centralized exchanges are more liquid and faster in terms of speed of trading. Furthermore, most of them are user-friendly and offer large functionality. Mainly for these reasons centralized exchanges dominate the cryptocurrency market. However, their customers lose all the benefits of a decentralized value-transfer network for which the bitcoin payment system was invented (Goldenberg, 2018). Benefits include greater transparency of ownership, high security, reduced costs of networking and verification, no market power and data access by an intermediary (Catalini and Gans, 2017).

Unlike centralized trading platforms, *decentralized* exchanges fully utilize the blockchain technology. They do not immediately process trade, but match people behind the orders, and the whole matching process is operated and maintained exclusively by software. As a result, compared to centralized exchanges, trading operations are cheap: very small fees or no fees are necessary. Furthermore, decentralized exchanges are invulnerable to government interference: government cannot close the exchange and cannot force the exchange to collect information of their users. Traders have their personal wallets, that is, a decentralized exchange does not hold any private keys and funds. Cryptocurrency purchases are made in-person or online through cash deposits on banks, PayPal transfers, wire transfers and other transfer methods. The transactions are performed trustlessly and securely since they are recorded on the blockchain, however, high security comes at the cost of longer waiting time for actual bitcoin and fiat transactions to complete (Goldenberg, 2018).

Now, decentralized exchanges are less popular, as they often lack simple user interfaces and offer limited functionality. Because of these two shortcomings, trading volume and liquidity of decentralized exchanges is lower as compared with centralized ones. In fact, 99% of cryptocurrency transactions go through centralized exchanges (Goldenberg, 2018).

2.1 Fee Schedule

In practice, the arbitrageur has to incur a number of transaction fees in cryptocurrency market. The fee structure, which will be described below, is common not only for the exchanges from the sample, but also for other centralized cryptocurrency exchanges. The customers must pay certain types of fees for the services of centralized cryptocurrency exchanges. Typically, the fees depend on the kind of transaction and type of service. Setting aside derivatives order execution, trading at margin and short-selling, there are, in total, three types of fees: *fiat deposit* and *withdrawal fee*, paid as a fixed amount or a percentage from transfer amount, *bitcoin withdrawal fee* (including miners' fee), paid as a fixed amount in BTC, and *trading fees*, paid as a percentage from trading amount. Fees vary from exchange to exchange, but, in general, same type of fees do not differ significantly. An exception may be fiat deposit and withdrawal fees: the amount can be fixed or payable as a percentage of the size of deposit or withdrawal. Table 10 summarizes all the above fees for the exchanges participating in the analysis. The fee schedules are taken from the exchanges' websites.

First, there are *deposit fees*. Traders transfer funds into an account on an exchange, which may entail two type of fees – a fee associated with transfer method and a deposit fee to the exchange. The size of the fiat deposit fee to the exchange usually depends on the currency unit and the type of transfer, for example, the international wire transfer, SEPA (Single Euro Payments Area) bank transfer, PayPal and other methods. As for cryptocurrency, most, if not all exchanges make it free to deposit.

Second, when the funds finally arrive on the exchange account, the trader can start trading operations. Most exchanges charge *trading fees* that traders pay upon the execution of a trade. Trading fees are usually expressed as percentages of the trading amount. Many exchanges have developed a maker-taker trading fee structure. *Maker fees* are paid when the trader adds liquidity to the exchange's order book by placing a limit order under the ticker price for buy and above the ticker price for sell. *Taker fees* are paid when the trader removes liquidity from the order book by placing a market order (Bitfinex, 2018). The range of maker-taker fees is usually between 0.1% – 0.3%, whereas maker fees are always smaller than taker fees to encourage participants to maintain market depth in the order book. Typically, the larger the transaction is, the smaller are the proportional taker and maker fees. Some exchanges provides brokerage services for a fee. In the exchange sample of this study, only CEX.IO provides a brokerage service. It is based on Fill-or-Kill (FOK) orders and guarantees that its users will receive not less than the agreed amount of cryptocurrency for the agreed price. This is a simplified way for new users to buy bitcoin: beginners need to choose an amount of cryptocurrency they want to buy, and the price is calculated, proceed to the payment stage. Payments can be made using a bank card. The CEX.IO broker trades at a premium – the exchange takes a 7% fee for the service which is charged from the amount of fiat currency (CEX.IO, 2018).

Third, the trader may wish to withdraw his or her funds from the account. Same as for a fiat deposit fee, a *fiat withdrawal fee* is paid to the exchange and depends on the currency unit and the transfer method. When a trader withdraws bitcoins from an exchange, the exchange will need to make an on-chain transaction (i.e. the transaction has to be recorded on the blockchain) and pay transaction fees to the miners. *Bitcoin withdrawal fees* as well as the miner's fees are expressed and paid in bitcoins. Exchanges set a bitcoin withdrawal fee themselves, looking at the average bitcoin transaction fees and the network load. When many people are transacting, the miners give priority to the transactions with higher fees. Exchanges may increase the transaction fees to get faster the confirmation from the blockchain network. At the same time, some exchanges make a profit by charging a percentage of a total transfer value, so the traders pay two kinds of fees: the withdrawal fees to the exchange and the transaction fees to the miners (Peh, 2018). However, most exchanges don't reveal the amount paid as the withdrawal fee to the exchange.

All bitcoin withdrawal fees displayed on the exchanges' websites are fixed values, i.e. they do not depend on the quantity of bitcoins traded. Certain trading platforms, such as Bitstamp and Coinbase Pro (previously GDAX), do not charge traders to withdraw bitcoins, i.e. they cover the miners' fees associated with bitcoin withdrawal (bitcoin on-chain transaction).

3 Bitcoin Order Books

As discussed in the previous chapter, cryptocurrency exchanges may serve as direct peer-to-peer exchanges or as trading platforms similar to traditional currency and stock markets. The focus of this study is on cryptocurrency exchanges, operating limit order books, which are used by most traditional markets (Gould et al., 2013). In this respect, cryptocurrency trading platforms with LOBs exhibit similar features as conventional exchanges. However, cryptocurrency exchanges operate 24 hours, seven days a week, and without opening and closing auctions, which constitutes a notable difference to traditional markets.

The notation below builds upon the mathematical description of limit order books introduced by Gould et al. (2013). The limit order book or the LOB of exchange i is denoted by $\mathcal{L}_t^i$ and defined as a list of active orders in the market for a minute t, where an order x^i submitted at time t_{x^i} with price p_{x^i} and size q_{x^i} is a commitment to sell (respectively, buy) up to q_{x^i} units of the traded bitcoins at a price no less than (respectively, no greater than) p_{x^i}. Orders are submitted by a trader with an anonymous key that uniquely identifies him or her. The active orders in an LOB $\mathcal{L}_t^i$ can be partitioned into the set of active buy orders $\tilde{\mathcal{B}}_t^i$, and the set of active sell orders $\tilde{\mathcal{A}}_t^i$. Trade-matching algorithm of an exchange uses the LOB to pair buy and sell orders: if a buy (sell) concurs a sell (buy) order on price, then the matching occurs immediately. If not, the order becomes active until either it is fully filled, or it is cancelled. The exchange's matching algorithm matches the order against the current order book and executes it at the best price available. In the case where there is more than one buyer (seller) with the same bid (ask) price, the order which arrives earlier will be executed first.

When several orders x^i at time t contain the same price, they are referred as a price level k. In the bid-side and ask-side depth profiles $\mathcal{B}_t^i$ and $\mathcal{A}_t^i$ of the LOB of exchange i, where

$$\mathcal{B}_t^i = (p_k^B, q_k^B)_t^i \tag{1}$$

and

$$\mathcal{A}_t^i = (p_k^A, q_k^A)_t^i, \tag{2}$$

each pair of price $p_k^{A,i}$ and quantity $q_k^{A,i}$ corresponds to the price level k. A quantity $q_k^{A,i}$ is a sum of all quantities q_{x^i} of orders x^i, submitted at time t_{x^i}, that correspond to the same price $p_k^{A,i}$:

$$q_k^B := \sum_{x^i \in \tilde{\mathcal{B}}_t | p_{x^i} = p_k^B} q_{x^i}. \tag{3}$$

The ask-side depth $q_k^{A,i}$, available at price $p_k^{A,i}$ at time t and exchange i, is defined similarly using $\tilde{\mathcal{A}}_t^i$. The bid-side (ask-side) depth profiles entries are sorted in price descending (ascending) order such that the corresponding best quotes appear in the top, since the order with the best price will be filled first. In this way, the bid-side depth profile $\mathcal{B}_t^i$ at time t is a set of all ordered pairs $(p_k^{B,i}, q_k^{B,i})_t$, and the ask-side depth profile $\mathcal{A}_t^i$ is a set of all ordered pairs $(p_k^{A,i}, q_k^{A,i})_t$.

The first level of the LOB shows the *best bid price*, defined as the highest stated price among active buy orders $\mathcal{B}_t^i$ at time t and exchange i:

$$b_t^i := \max_{k \in \mathcal{B}_t^i} p_k = p_{1,t}^{B,i}, \tag{4}$$

and the *best ask price*, defined as the lowest stated price among active sell orders $\mathcal{A}_t^i$ at time t and exchange i:

$$a_t^i := \min_{k \in \mathcal{A}_t^i} p_k = p_{1,t}^{A,i}. \tag{5}$$

The average between the best bid and best ask price determines the *mid-price* at time t and exchange i:

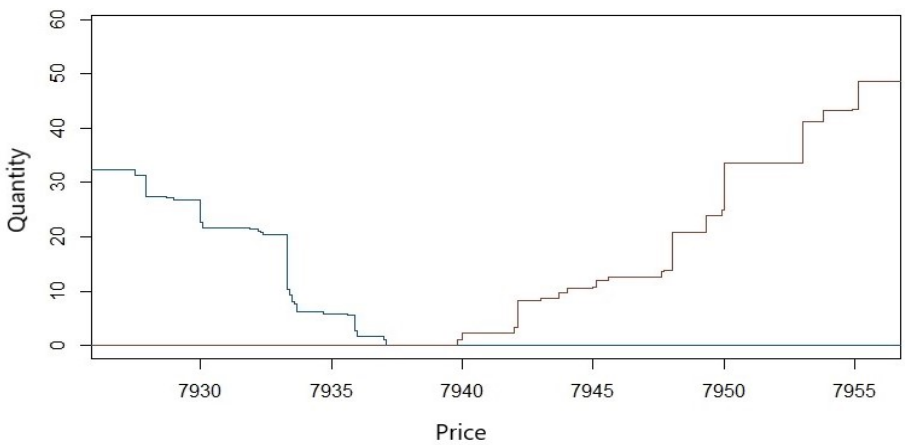
$$m_t^i := \frac{a_t^i + b_t^i}{2}. \tag{6}$$

3.1 Bitcoin Order Book Data

The data sample consists of the limit order books $\mathcal{L}_t^i$ for each minute t and each exchange i with corresponding bid-side and ask-side depth profiles $\tilde{\mathcal{B}}_t^i$ and $\tilde{\mathcal{A}}_t^i$ for the period from 02:00:00, March 29, 2018 to 01:00:00, September 5, 2018. The limit order books have been extracted by querying application programming interfaces (APIs) of seven cryptocurrency exchanges that feature BTC versus USD trading. Specifically, the first 25 levels of bid and ask prices with corresponding volumes were retrieved once in a one-minute interval $t \in \{1, \ldots, T\}$. Each order book snapshot is time stamped down to the minute defined by Central European Time (CET) for consistency. Since the BTC/USD currency pair leads by trading volume and order book depth, this study focuses on the BTC/USD market.

After the limit order books were retrieved, the bid-side and ask-side depth profiles $\mathcal{B}_t^i = (p_k^B, q_k^B)_t^i$ and $\mathcal{A}_t^i = (p_k^A, q_k^A)_t^i$ were calculated, where quantity q_k^i is a sum of all quantities q_{x^i} of orders x^i, submitted at time t_{x^i}, that correspond to the same price p_k^i. All subsequent computations will be made based on the sorted order book sequences $\mathcal{B}_t^i$ and $\mathcal{A}_t^i$. Figure 1 visualizes the 25-level LOB of Bitfinex at some instant time, where the x-axis is bitcoin unit prices in USD and the y-axis is cumulative order depth available at each price level k.

Figure 1. An example LOB



Notes: This Figure visualizes the 25-level LOB of Bitfinex at some instant time, where the x-axis is bitcoin unit prices in USD and the y-axis is cumulative order depth available at each price level k. Bids (buyers) are on the left, asks (sellers) are on the right.

3.1.1 Exchange Sample

Table 10 in the Appendix contains key characteristics of the cryptocurrency exchanges considered in the analysis. The sample covers the following seven exchanges: Bitfinex, bitFlyer, Bitstamp, Kraken, Coinbase Pro (formerly GDAX), Gemini, and CEX.IO.

The cryptocurrency exchanges in the sample are the largest exchanges by trading volume, allowing customers to deposit and withdraw fiat currency. The market shares by exchange for all currency pairs are represented by a pie chart in Figure 17 in the Appendix, and includes other exchanges not considered in the analysis. According to the source of the market share data bitcoinity.org (2020), in the sample period, all exchanges together make a significant share of the Bitcoin market - around 70%.

The exchange sample contains only centralized cryptocurrency exchanges. All exchanges in the sample are domiciled in G10 countries and have some level of KYC (Know-Your-Customer) and AML (Anti-Money Laundering) compliance. Therefore, the operational risk of the exchanges in the sample is relatively lower than that of most cryptocurrency exchanges, as only 32% of trading platforms in cryptocurrency market in the US and EU conduct full identity checks of their customers (Golstein, 2018). All exchanges except bitFlyer have USD or EUR as a base currency, and all exchanges except Bitfinex are available both in the US and in Europe. In 2017, Bitfinex restricted US citizens from its service, therefore, using the exchange became illegal for the US citizens. As for August 15, 2018, all US citizens and US corporations have been prohibited from using the services of Bitfinex (Bitfinex, 2018).

Despite the fact the considered exchanges are based in different countries, no exchange is a subject to cross-border capital controls, i.e. they are all based on the same jurisdiction and no documents are necessary to submit to authorities in order to prove the reasons for the transfers. For example, bitFlyer, a Japanese exchange, offers trading of BTC to Japanese Yen, so JPY is a base currency on this exchange - the main fiat currency on the exchange that typically coincides with the exchange's geographic focus (Makarov and Schoar, 2020). bitFlyer is nevertheless licensed to operate across the US and Europe. However, bitFlyer's BTC/USD market share remains small relative to other exchanges in the sample (bitcoinity.org, 2020).

The exchanges do not exhibit a substantial heterogeneity in terms of taker fees in the sample period. Only Gemini stands apart and increases its taker fee to 1% on April 2, 2018. Before this date, Gemini charged 0.25%, which is comparable to other exchanges. Coinbase Pro also increases the taker fee from 0.25% to 0.3% on April 5, 2018. However, there is a heterogeneity in terms

Table 1. Characteristics of the Cryptocurrency Exchanges

Exchange	Bitfinex	bitFlyer	Bitstamp	CEX.IO	Gemini	Coinbase Pro	Kraken
company location	Hong Kong	Japan	United Kingdom	United Kingdom	United States	United States	United States
founded	2012	2014	2012	2013	2015	2016	2011
base currency	USD	JPY	USD/EUR	USD	USD	USD/EUR	USD/EUR
margin account	yes	yes	no	yes	no	yes	yes
business account	yes	yes	no	yes	yes	yes	yes
short sale	yes	no	no	no	no	yes	yes
BTC confirmations	3	6	3	3	3	3	6
taker fee	0.2%	0.15%	0.25%	0.25%	0.25-1%	0.25-0.3%	0.26%
BTC Withdrawal fee	0.0004	0.0004	0	0.0005-0.001	0.002	0	0.0005
USD Deposit fee	0.10%	0	0.05%	0	0	$10	0
Minimum USD Withdrawal fee	0.1%	$10	0.09%	$50	0	$25	$5
Order Books	221452	221381	220520	220290	220077	221240	221903

Notes: This table gives the exchange-specific characteristics for the time period from 02:00:00, March 29, 2018 to 01:00:00, September 5, 2018. bitcoin withdrawal fees are quoted in BTC. USD Deposit/Withdrawal fees are bank transfer deposit/withdrawal fees, or credit card deposit/withdrawal fees if bank transfer is not available. The data was taken from the official exchanges' websites.

of BTC and USD withdrawal fees. Whereas the majority of exchanges impose a fixed amount of USD withdrawal fee, Bitfinex and Bitstamp charge percentages of the size of withdrawal. The size of the fixed USD withdrawal fee differs substantially across the exchanges and range from 0$ on Gemini to 50$ on CEX.IO. As for BTC withdrawal fee, the exchanges typically charge a fixed amount from 0.0005 to 0.002 BTC. As an exception, Bitstamp and Coinbase Pro do not charge traders to withdraw bitcoins. During the sample period, only CEX.IO increases the BTC withdrawal fee from 0.0005 to 0.001 BTC on April 14, 2018.

Besides withdrawal fees, the exchanges differ by requirements with respect to the number of block confirmations before they proceed to process BTC deposits. Specifically, the exchanges require a certain number of bitcoin deposit confirmations, which is a number of times bitcoin deposit must be confirmed by the network prior to being considered valid. For example, three required confirmations mean that the incoming transaction is three blocks "deep" in the blockchain. "Deep" transactions are more secure, since the probability that a miner will succeed in "double spent" will decrease as the number of blocks that need to be replaced increases. According to Yermack (2017), the header of each block contains a cryptographic hash of the previous block, which itself includes a hash function derived from its predecessor, therefore the data in any given block cannot be altered without the ripple effect of changing the cash codes of all subsequent blocks. It is a difficult task to find valid hashes for all subsequent block headers up to the latest, so a thief or forger seeking to change old transactions will face the insurmountable problem (Yermack, 2017).

Availability of business accounts shown in Table 10 indicates that the exchange is equipped with advanced trading features and provides large investors and institutions with such preferential conditions as lower deposit/withdrawal and trading fees, high or no limits for fiat currency withdrawals and various transfer methods. Furthermore, a business account provides institutional investors with the opportunity to hold inventories. A few exchanges allow short sales, among them Bitfinex, Coinbase Pro and Kraken, which offer short sale option to their large customers. Instead, some exchanges offer the arbitrageur to establish a negative position in bitcoin by trading at margin, which is similar to short sales but does not allow for physical settlement (Makarov and Schoar, 2020). Trading at margin is associated with the fees paid separately for opening a position (opening fees) with margin borrow limits depending on the verification level (type of account) and for maintaining a position for a certain period of time (rollover fees).

3.2 Measures of Market Liquidity, Order Imbalance and Volatility

In this section, measures of liquidity, order imbalance and volatility, which are used in the analysis, are provided. In context of arbitrage, liquidity, as one of the cryptocurrency exchange characteristics, is of particular interest. The limit order book accumulates active orders and provides liquidity to the market. According to Kyle (1985), one of the key properties of a liquid market are (a) tightness ("the cost of turning around a position over a short period of time"), and (b) depth ("the size of an order-flow innovation required to change prices a given amount"). The necessary condition for a "tight" market is a small difference between the bid and ask prices, which measured by a quoted bid-ask spread. In the context of a LOB, the *quoted bid-ask spread* for exchange i is a difference between the best ask price and the best bid price:

$$QS_t^i := a_t^i - b_t^i. \tag{7}$$

The *proportional bid-ask spread* is defined as:

$$PS_t^i := \frac{a_t^i - b_t^i}{m_t^i}, \tag{8}$$

where m_t^i is a mid-price defined in Equation 6.

Market microstructure theory establishes three type of costs that a LOBs bid-ask spread PS_t^i covers: order processing costs or the cost of waiting (Demsetz, 1968), inventory holding costs, modelled in Stoll (1978b) and adverse information costs, which received a greater emphasis by Glosten and Milgrom (1985). Demsetz (1968) argued that the determination of spread is dominated by waiting costs. A trader with a limit order sets his or her sell (buy) price higher (smaller) than buy (sell) price to earn a profit for supplying immediacy, but the execution of the order is neither immediate nor certain, whereas a market order guarantees immediate execution and certainty. In this respect, the bid-ask spread can be considered as a measure of how highly the market evaluates the cost of delayed execution versus the cost of immediacy.

Whereas the bid-ask spread measures the cost dimension of liquidity (Demsetz, 1968), the *average quoted depth* measures its quantity dimension and related to the second property (b) of a liquid market stated by Kyle (1985) – deep enough market for a small price impact. The average quoted depth for exchange i is an average quantity that the trader can trade at the best prices:

$$AvgD_t^i := \frac{q_{1,t}^{A,i} + q_{1,t}^{B,i}}{2}, \tag{9}$$

where $q_{1,t}^{A,i}$ and $q_{1,t}^{B,i}$ are the quantities associated with the best ask and the best bid price from the ask-side $(p_k^{A,i}, q_k^{A,i})_t$ and bid-side $(p_k^{B,i}, q_k^{B,i})_t$ depth profiles respectively. Too many orders of a particular type – either buy or sell on exchange make it impossible to match the orders. The limit order book imbalance measures the imbalance between supply and demand by a ratio of limit order volumes between the bid and ask side, and it is calculated as:

$$OI_t^i := \frac{V_{bid(t)} - V_{ask(t)}}{V_{bid(t)} + V_{ask(t)}}, \tag{10}$$

where $OI_t^i \in [-1, 1]$, and $V_{bid(t)}$ and $V_{ask(t)}$ are weighted average volumes at certain levels of the limit order book (number of levels can be chosen), using exponentially decreasing weights:

$$V_{bid(t)} = e^{-0.5(0)} \cdot q_{1,t}^{B,i} + e^{-0.5(1)} \cdot q_{2,t}^{B,i} + e^{-0.5(2)} \cdot q_{3,t}^{B,i} + ... + e^{-0.5(24)} \cdot q_{25,t}^{B,i}. \tag{11}$$

The formula is proposed by Rubisov (2015). Volumes $q_{l,t}^{B,i}$ are specified quantities from the LOB, where the bid depth corresponds to the bid-side depth profile $(p_k^{B,i}, q_k^{B,i})_t$ of the LOB $\mathcal{L}_t^i$ at time t. The volume $V_{ask(t)}$ is defined similarly using the corresponding quantities $q_{l,t}^{A,i}$ from the ask-side depth profile $(p_k^{A,i}, q_k^{A,i})_t$ of the LOB $\mathcal{L}_t^i$. In this way, the quantities at deeper levels of the order book are considered, but the deeper the level, the lesser weight will be given to the quantities at this level.

The bitcoin price volatility is estimated by a kernel method, proposed by Kristensen (2010). The estimator is a kernel-weighted version of the standard integrated volatility estimator. For each exchange i and minute t, the spot volatility $(\sigma_t^i)^2$ is estimated by:

$$(\widehat{\sigma}_{it})^2(h) = \sum_{l=1}^{t} K(l - t, h)(m_{i,l} - m_{i,l-1})^2, \tag{12}$$

where $K(l - t, h)$ is an one-sided Gaussian Kernel smoother with bandwidth h, and m is a mid-price, defined in Equation 6. The bandwidth h was chosen according to the method proposed by Kristensen (2010): to choose h such that the Integrated Squared Error is minimized:

$$\widehat{ISE}_T(h) = \sum_{i \in I_T} [(m_{i,l} - b_{i,l-1})^2 - \widehat{\sigma_{i,T,l}}^2(h)]^2. \tag{13}$$

Table 2 show that the exchanges in the sample are heterogeneous in terms of liquidity and order imbalance. Although Coinbase Pro is not the largest exchange in the sample by trading

Table 2. Measures of Liquidity, Order Imbalance and Volatility

Exchange	Quoted Spread (USD)	Quoted Spread, (bp)	Average Quoted Depth (BTC)	Median Quoted Depth (BTC)	Order Imbalance	Absolute Order Imbalance	Daily Spot Volatility	Daily Market Share
Bitfinex	0.25	0.35	7.23	8.79	0.02	0.50	2.87	0.41
bitFlyer	15.74	21.16	0.34	0.38	-0.32	0.53	2.67	0.002
Bitstamp	4.59	6.24	1.54	1.59	0.08	0.44	3.15	0.13
Coinbase Pro	0.05	0.06	11.09	16.21	-0.01	0.70	2.72	0.13
Gemini	1.98	2.74	2.47	2.55	0.07	0.48	2.78	0.04
CEX.IO	12.70	17.25	0.62	0.73	0.02	0.37	2.67	0.007
Kraken	2.50	3.37	1.53	1.79	-0.01	0.67	2.74	0.067

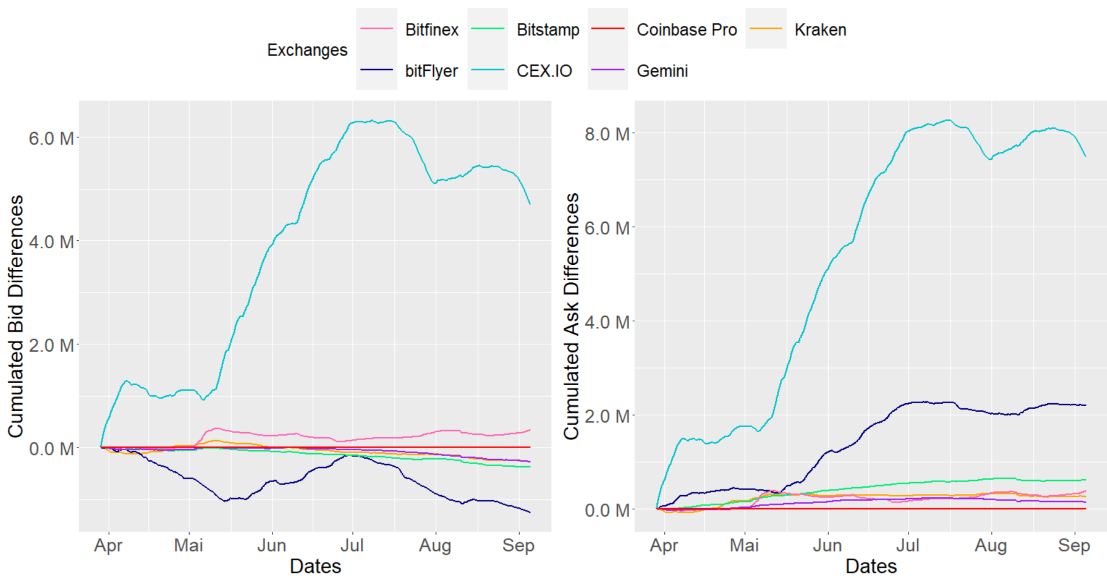
Notes: The Table reports the averages of measures of liquidity, order imbalance, volatility and market share for the time period from 02:00:00, March 29, 2018 to 01:00:00 September 5, 2018 for each exchange. Quoted Spread (USD) is the average quoted bid-ask spread in USD, defined in Equation 7. Quoted Spread (bp) is the average proportional bid-ask spread in basis points, defined in Equation 8. Average quoted depth (BTC) is defined in Equation 9. Order imbalance is the average ratio of limit order volumes between the bid and ask side, as it is defined in Equation 10. Daily spot volatility is the average daily volatility, computed by converting minute-level volatility estimates to the daily level. The minute-level volatility is estimated following Kristensen (2010), and converted to the daily level by multiplying it with $\sqrt{1440}$, which is the square root of the number of minutes on a trading day. The daily volume data to compute market share is obtained from bitcoinity.org (2020). The market share is the average daily volume (BTC) of a given exchange for BTC/USD currency pair weighted by the sum of daily volumes for BTC/USD currency pair of other exchanges, available on bitcoinity.org (2020).

volume, it is the most liquid exchange with the highest mean of the time-weighted quoted depth and the smallest time-weighted quoted bid-ask spread among all exchanges in the sample. The least liquid exchanges in the sample are also the smallest by BTC/USD trading volume - bitFlyer and CEX.IO exhibit the smallest average quoted depth and the largest time-weighted bid-ask spread. As it will be seen later, the highest and most persistent bitcoin price difference will be observed on these two exchanges. Nevertheless, comparing to traditional markets, the bid-ask spread on bitFlyer and CEX.IO is similar to spreads in small-cap stocks traded on Nasdaq - the time-weighted average spread for small-cap stocks traded for 2008 and 2009 is 38.06 basis points (bp). The average bid-ask spread in large-cap stocks traded on Nasdaq market is comparable to the remaining exchanges in the sample, and range between 4.44 bp and 7.20 bp on high volatility days (see Table 6 in Brogaard et al. (2014)). As for volatility, it does not differ substantially across cryptocurrency markets. Since bitcoin is identical across exchanges, the volatility estimates were expected to be similar. In contrast to spread, daily volatility of bitcoin is higher than the average daily volatility of S&P 500 index during the same period, which is 0.65% (Hautsch et al., 2020).

3.3 Price Differences across Cryptocurrency Markets

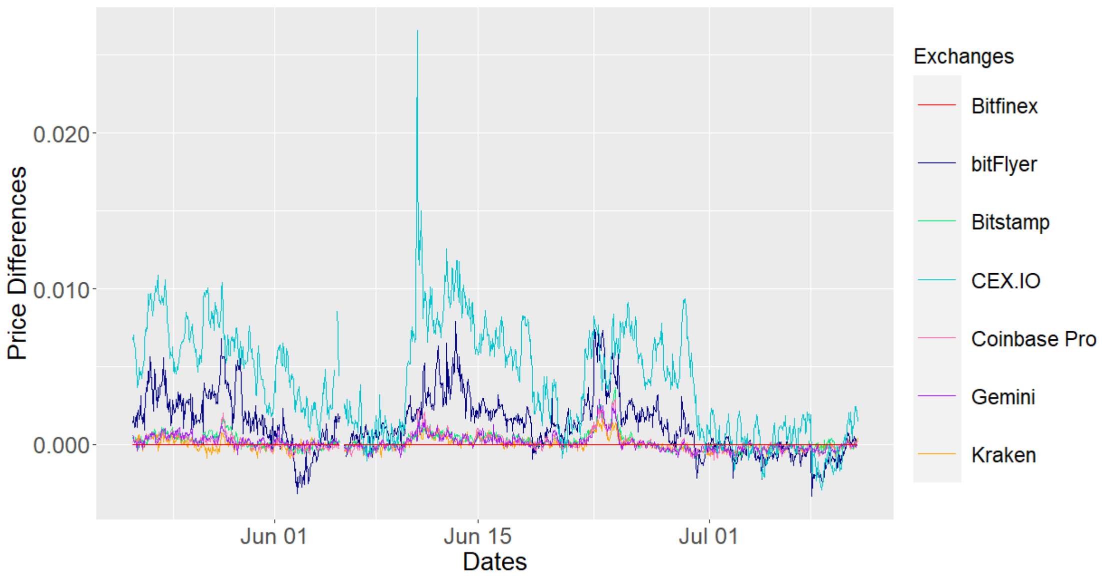
Figure 2 visualizes the cumulative sum of the minutely level differences in bid and ask prices of bitcoin over time and across exchanges. The differences in bid prices Δ_{it}^{B} are calculated by subtracting the best bid price b_t of Coinbase Pro from the best bid price of exchange i at time t. The differences in ask prices Δ_{it}^{A} are calculated analogously using the best ask prices a_t of respective exchanges. Since Coinbase Pro is one of the largest, oldest and advanced exchanges, its first level prices were chosen as a benchmark. By summing the values up over time, the time series of the minute-level price differences have been smoothed, making it easier to distinguish and compare multiple overlapping plots in one graph. To compare price differences between exchanges, the slopes of the curves of the cumulated price differences are of interest.

Figure 2. Cumulated Price Differences across Exchanges over Time



Notes: This figure shows the cumulative sum of the minute-level bitcoin price differences Δ_{it} over the time period from 02:00:00, March 29, 2018 to 01:00:00, September 5, 2018, which makes the entire sample period. The differences in bid prices Δ_{it}^{B} are calculated by subtracting the best bid price b_t of Coinbase Pro from the best bid price of exchange i at time t. The differences in ask prices Δ_{it}^{A} are calculated analogously using the best ask prices a_t of respective exchanges. The best bid and best ask prices are defined in Equations 4 and 5, respectively.

Figure 3. Hourly Price Differences across Exchanges over Time



Notes: This figure represents hourly price differences Δ_{it}^{M}, calculated by subtracting the mid-price m_t of Coinbase Pro from the mid-price of exchange i at time t for the period from 05:00:00, May 22 to 5:00:00, June 11. The price differences are then normalized by the average of the hourly mid-prices of two respective exchanges. The mid-prices are defined in Equation 6.

In order to measure the dispersion in bitcoin prices in different markets, the hourly cross-sectional price difference is introduced and defined as a ratio between the average hourly bitcoin price (first level ask, bid or midquote price) in exchange i and the average price (first level ask, bid or midquote price) across all exchanges in hour t:

$$PD_{it}^{B(A)} := \frac{a_{it}(b_{it})}{\overline{a}_t(\overline{b}_t)} - 1, \tag{14}$$

where $\overline{a}_t(\overline{b}_t)$ is the average of the hourly first level ask (bid) prices of all other exchanges at time t. Table 3 provides summary statistics of $PD_{it}^{B(A)}$ the cross-sectional hourly bid (ask) differences between exchange i and all other exchanges in the sample (expressed in %). As already stated, the hourly measures of the ask, bid and midquote prices are computed by taking an average of the corresponding minute-level prices in a given hour t. The ask and bid price differences are considered separately, as it is interesting to check, whether the bid and ask prices deviate in the same direction, and whether there are differences in summary statistics. Figure 2 and

Table 3 show that bitFlyer's bid and ask quotes do not share a common dynamics: on average, it quotes lower bid prices than all exchanges in the sample, but higher ask prices than most other exchanges in the sample (except CEX.IO), although the deviation in absolute is nearly the same. This price behavior is different from the price behavior on other exchanges: their bid and ask price differences share a common dynamics and similar summary statistics.

Table 3. Summary statistics of the exchange-specific differences in the first level bid and ask prices

Exchange	Mean		Abs		StDev		Min		Max		%pos	
	Bid	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid	Ask
Bitfinex	-0.01	-0.09	0.18	0.20	0.24	0.25	-0.93	-1.14	1.01	0.92	52.44	35.60
bitFlyer	-0.11	0.05	0.21	0.18	0.26	0.26	-2.58	-0.88	1.10	2.49	29.04	57.32
Bitstamp	-0.06	-0.07	0.14	0.14	0.17	0.16	-0.90	-0.99	0.90	0.90	35.47	32.94
Coinbase Pro	-0.03	-0.11	0.13	0.16	0.17	0.18	-1.00	-1.07	2.04	1.97	46.34	24.86
Gemini	-0.05	-0.11	0.12	0.14	0.15	0.16	-0.91	-1.03	0.60	0.48	39.56	25.60
CEX.IO	0.31	0.43	0.64	0.68	0.78	0.79	-1.49	-1.41	4.95	5.66	64.00	71.26
Kraken	-0.06	-0.10	0.16	0.17	0.20	0.20	-0.91	-1.05	0.57	0.53	41.70	32.73

Notes: This table provides summary statistics of $PD_{it}^{B(A)}$ - the cross-sectional hourly averages of differences between the first level bid (ask) price on each exchange and the first level bid (ask) prices on all other exchanges over the time period from 02:00:00, March 29, 2018 to 01:00:00, September 5, 2018. The price differences are then normalized by the average of the hourly bid (ask) of all other exchanges. The columns present the mean, the mean of the absolute value, the standard deviation, the minimum, and the maximum value of the price differences (expressed in %) as well as the percentage of hours in the sample period on which the price deviation was positive.

Figure 2 and 3 show that the majority of exchanges in the sample, such as Bitfinex, Bitstamp, Coinbase Pro, Gemini and Kraken do not feature large price differences. Compared to the price differences relative to CEX.IO and bitFlyer, the flat slopes of the curves of cumulated price differences of these exchanges indicate smaller, less persistent price differences with regularly switching signs. Over the time period, the price differences have become more integrated, especially in the second half of the sample period, i.e. in July and August. The ask prices seem more integrated across exchanges than bid prices - both CEX.IO and bitFlyer exhibit lower bid prices than other exchanges. In the sample period, average absolute prices difference ranges between 0.12% and 0.8% (see Table 3), whereas Makarov and Schoar (2020), in their sample period between November 2017 and February 2018, report that, on exchanges that operate in the US and Europe, the average price discrepancies are around 2% with maximum of 6%. Kroeger and Sarkar (2017) documented daily price differences for the longer period: from July 2014 to September 2016, and showed that the average absolute price difference is around 2%, but the maximum ranges between 17% and 41%. These findings suggest a general trend of decreasing prices differences and increasing efficiency of cryptocurrency markets.

CEX.IO stands out from other cryptocurrency markets with the longer-lasting and larger price deviations. The cumulated bid and ask differences curves of CEX.IO in Figure 2 have the steepest positive and negative slopes, which indicate that CEX.IO exhibit the largest and most persistent price differences among all other exchanges in the sample. The positive slopes show that CEX.IO consistently trades at higher bid and ask prices relative to all other exchanges, the negative slopes - the other way around. The negative slopes are much shorter: CEX.IO quoted lower bid and ask prices than other exchanges for several days in April and June and two last weeks of July, but most of the time CEX.IO trades bitcoin at a premium. According to Table 3, CEX.IO features the highest mean, standard deviation, maximum of the bid and ask price differences. Specifically, the mean and the mean of absolute vales of the cross-sectional hourly average of differences between the first level bid prices on CEX.IO and all other exchanges is 0.31% and 0.64%, respectively. The maximum difference in bid prices between CEX.IO and other exchanges is almost 5%. The percentage of the positive price differences is also the highest among all other exchanges: there are 64% of hours in the sample period on which the bid price deviation of CEX.IO is positive.

Also bitFlyer's prices tend to frequently deviate from other exchanges in the sample. Unlike CEX.IO, where bitcoin consistently trades at a premium, bitFlyer often alternates between trading bitcoin at a premium and a discount. Interestingly, together with CEX.IO, bitFlyer tends to quote higher ask prices than other exchanges, but bid prices are the lowest among other exchanges. In May and June both CEX.IO and bitFlyer exhibit a common dynamics and quote higher bid prices than other exchanges, in the last two weeks of July and in the beginning of August - lower bid prices than other exchanges. According to the slopes of the cumulated bid and ask differences in Figure 2, bitFlyer exhibits larger price deviations than other exchanges, but smaller than CEX.IO. bitFlyer quotes higher ask prices than other exchanges, besides CEX.IO, in May and June, but revert to the common dynamics in July. After CEX.IO, bitFlyer features the second highest mean, maximum of the differences in ask prices and the second highest percentage of positive differences in ask prices among all other exchanges. The mean and percentage of positive differences in ask prices on bitFlyer is 0.05% and 57.32%, respectively. As for bid prices, in comparison to other exchanges, they are lower in April and in the first half of May, then the price deviation reverts its sign, and until July bitFlyer quoted higher bid prices than other exchanges, besides CEX.IO. Table 3 also shows that, overall, bitFlyer tends to quote lower bid prices: bitFlyer features the smallest mean, minimum of the differences in bid prices and the smallest percentage of the positive differences in bid prices among all other exchanges. The mean and percentage of positive differences in bid prices on bitFlyer is -0.11% and 29.04%, respectively.

In summary, only two exchanges, CEX.IO and bitFlyer tend to exhibit larger and persistent bitcoin price differences. However, there are realtively short periods, when the prices on Bitfinex tend to deviate substantially: it exhibits the highest bid and ask prices among all exchanges in the first week of May, and also tends to quote higher bid prices in July. Table 3 shows that Bitfinex features 52% of hours in the sample period of positive bid price deviations, which is the second highest value after CEX.IO. Overall, much flatter slopes of cumulated bid and ask differences in Figure 2 indicate that the prices on Bitfinex are not persistently at a discount or a premium as on CEX.IO and bitFlyer in this sample period.

The magnitude of bitcoin price differences is higher compared to price deviations in other markets and assets. For instance, Gagnon and Karolyi (2010) find that the mean difference for ADRs (American Deposit Receipts) is 4.9 basis points with a daily standard deviation for a given stock pair of 1.4%, even though the deviations can reach extreme values as large as -40% and 127%. Wenxin et al. (2018) observed mean daily deviations for the covered interest parity that fluctuate from 6 to 19 basis points annualized, with standard deviations from 4 to 23 basis points. Nevertheless, in section 4, it will be shown, if the magnitude of the price differences and order book depth is large enough to cover the fees, stemming from the exchange and the blockhain network. The arbitrage opportunities will be quantified and documented. The persistence of the first level cross-market price differences and arbitrage profits will be also documented and discussed in section 4.

4 Quantifying Arbitrage Profits

The Bitcoin market provides a unique framework to study arbitrage due to a number of features. Unlike the traditional markets, there are hundreds of non-integrated and independent exchanges at which cryptocurrencies are traded 24 hours a day, 7 days a week with available and free of charge pricing data. Bitcoin, the most actively traded cryptocurrency, is completely identical and fully fungible across exchanges, which is different from traditional asset pairs such as American Depository Receipts (ADRs) and Dual Listed Companies (DLC), previously studied in the literature. The law of one price implies that assets with identical payoffs must trade at the same price (Gromb and Vayanos, 2010). However, empirical research has identified significant mispricing across bitcoin exchanges. Kroeger and Sarkar (2017) studied bitcoin price differences from January 2012 to August 2016 and identified the existence of persistent and statistically significant differences between the prices in multiple exchanges. Makarov and Schoar (2020) showed that from December 2017 to February 2018 large and recurring deviations in bitcoin prices persisted for several days and weeks across different countries as well as within the same country. In this thesis, Figure 2 in Section 3.3 suggests that arbitrage opportunities may exist as one can observe substantial (up to 6%, as shown by Table 3) and persistent price differences at least for two cryptocurrency exchanges in the sample - CEX.IO and bitFlyer. Bitcoin price deviations across exchanges suggest constraints to arbitrage that limit its effectiveness in achieving market efficiency. Also, there must be reasons why bitcoin have diverging prices on some cryptocurrency exchanges in the first place (Gromb and Vayanos, 2010).

Referring to Gromb and Vayanos (2010), the constraints to arbitrage, which will be discussed more extensively in Section 5.3, may arise as capital constraints (Shleifer and Vishny, 1997), transaction costs (Roll et al., 2007), holding costs (Pontiff, 2006) and short sale constraints (Ofek et al., 2004). In the Bitcoin market, a new type of market friction emerges - the blockchain-related settlement latency. Hautsch et al. (2020) shows that, in a situation where inventory holdings and margins are exhausted, and the arbitrageur is forced to physically transfer the asset between markets, the settlement latency imposes limits to arbitrage by exposing cross-market arbitrageurs to the risk of adverse price movements. This market friction is an inherent feature of the blockchain as a from of distributed ledger technology. The ownership of bitcoin is validated through the blockchain - the sequential database of information that is secured by methods of cryptographic proof (Nakamoto, 2008). The Nakamoto's blockchain bundles together up to 1 MB volume of transaction data into a new block approximately every 10 minutes and arrange these blocks in chronological sequence using hash functions. There is no sponsor or gatekeeper controlling the addition of new blocks. Instead, the process was decentralized to all users of the blockchain network in an ongoing competition catalysed by the award of new bitcoins to the miner who successfully completes a block with the required hash (Yermack, 2017). Miners also receive rewards for the verification of transactions (adding transactions into a new block) in form of transaction fees specified by the senders, and average verification time is approximately 20 minutes (Blockchain, 2020). The transfer of bitcoin is not accomplished and the receiving wallet cannot pursue further transactions unless the transfer is not verified. Thus, in contrast to traditional markets, there is no trusted intermediary that guarantee the final delivery of bitcoin in cross-market trading. There are, however, feasible arbitrage strategies to circumvent the blockchain-related settlement latency. In Section 4.1 the mechanics of the different arbitrage strategies that exist in the Bitcoin market will be discussed in more detail.

4.1 Arbitrage Strategies

Before quantifying arbitrage profits, the mechanics of arbitrage strategies in the Bitcoin market needs to be outlined. There are many feasible arbitrage strategies in the Bitcoin market, all of them are nevertheless associated with uncertainty and frictions of different forms and magnitudes that might prevent to exploit bitcoin price differences to their full potential. When constructing the arbitrage strategy, one has to consider the second feature of the Bitcoin market - the blockchain related settlement latency. For example, the arbitrageur could execute the following strategy: the arbitrageur buys bitcoin on the exchange, where it is relatively cheap, transfers and sells it on the exchange where it is relatively expensive. The mechanics of this arbitrage strategy is illustrated in Figure 21 in Appendix. As discussed above, the transfer of bitcoins is not accomplished unless the transaction is not verified. The riskless arbitrage with this strategy is therefore not possible, because transferring bitcoins across exchanges takes on average 10 minutes (during the sample period, new blocks are added to the network on average every 9.7 minutes). Over this period, the expected arbitrage profit associated with the strategy in Figure 21 is uncertain. Specifically, the arbitrage opportunity might disappear, i.e. the price difference can decrease or even revert in another direction, due to price volatility of bitcoin.

There are nevertheless ways to take advantage of the bitcoin price difference immediately and to circumvent the blockchain-related settlement latency. A way to avoid transferring bitcoins across exchanges directly after the purchase, is by simultaneously buying bitcoins on the exchange where the price is low and selling it on the exchange where the price is high. The arbitrage strategy would be to hold a positive balance of bitcoins on both exchanges, so that the arbitrageur can trade simultaneously across the two exchanges whenever the difference in prices occurs. The bitcoin balance will decline on the high-priced exchange, where bitcoins will be sold, and increase on the exchange where the bitcoin price is low. To restore the bitcoin balance the arbitrageur needs to transfer bitcoin from the exchange's wallet with high quantity of bitcoins to the one with low quantity of bitcoins. In these way, the arbitrageur does not need to transfer funds between exchanges immediately, as she already has an inventory of the asset on both markets. Nevertheless, the strategy requires active price monitoring and funds reallocation, therefore the speed is crucial when executing this strategy, and the entire process must be automated. The price monitoring and trade execution could be implemented by the recently developed cryptocurrency trading software systems and bots, which are often open source projects. For example, the

CCXT - Cryptocurrency Exchange Trading JavaScript and Python Library - allows to connect to several exchanges and trade with cryptocurrency exchanges and payment processing services worldwide (Huang, 2019; Kroitor, 2022). Thus, this library can be used to built an arbitrage bot - a cryptocurrency arbitrage framework which monitors multiple exchanges, requests ask and bid prices, compares them and builds arbitrage strategies (Huang, 2019). The open source arbitrage bot Blackbird short sells (long buys) on a short (long) exchange simultaneously (Hamilton, 2020). By using this trading system, the arbitrageur can profit from the trade if prices on the two exchanges converge in the future. Many arbitrage bots are written as Github projects, for example Triangular Arbitrage-Binance, Crypto Arbitrage Framework and many others (Huang, 2020). Hence, there are accessible tools that can support simultaneous trading on the exchanges and exploit instantaneous price differences whenever they arise.

Parallel holdings allows to avoid transferring bitcoins across two exchanges, but still expose the arbitrageur to bitcoin price fluctuations. For example, the bitcoin price on an expensive exchange could fall below the bitcoin price on a low-priced exchange, when bitcoins are transferred to a high-priced exchange to replenish the bitcoin balance. Also, an exposure to the counterparty risk, associated with the centralized nature (see Section 2.1) of cryptocurrency exchanges could prevent arbitrageur from using this strategy. Although investors in the Bitcoin market may not prefer to keep their large inventory holdings at exchanges for a long time, some investors may keep the assets at exchanges for hedging needs or diversification benefits. For those, who want to mitigate price and counterparty risks described above, an alternative would be to trade bitcoin at margin, establish short positions on exchanges or borrow bitcoin from *hodlers*, people who hold bitcoins rather than selling it. Incorporating costs associated with short-selling, trading at margin and borrowing bitcoin from hodlers in the arbitrage profits is problematic and beyond the scope of this thesis due to the lack of information and complexity of the whole trading procedure and fees, paid separately for opening a position (opening fees) with margin borrow limits, depending on the verification level (type of account), and for maintaining a position for a certain period of time (rollover fees).

In this thesis, a measure of the arbitrage profit evaluates the magnitude of the arbitrage profit the trader could achieve from the arbitrage strategy described above - purchasing bitcoins with dollars at one exchange, and selling bitcoins for dollars at another exchange, all simultaneously at the relevant bid and ask prices. Importantly, the quantified arbitrage profit is achievable in situations, where inventory holdings on an exchange are not exhausted, and all associated costs and risks are negligible. Hautsch et al. (2020) estimate arbitrage bounds in a scenario where arbitrageurs are fully exposed to settlement latency: where inventory holdings on an exchanges are too risky and expensive, and short-selling is costly or not possible. For example, short selling is not always feasible, since only a few exchanges allow short sales, among them Bitfinex, Coinbase Pro and Kraken, which offer short sale option to their large customers. Instead, some exchanges offer the arbitrageur to establish a negative position in bitcoin by trading at margin (see Table 10). Trading at margin does not allow for physical settlement, so the arbitrageur becomes exposed to the convergence risk, since the profit is positive when prices on the two exchanges converge in the future (Makarov and Schoar, 2020). Hautsch et al. (2020) incorporated risks associated with bitcoin price fluctuations by including the spot volatility on the sell-side exchange in estimation of the arbitrage bounds. The authors document that the arbitrage bounds due to settlement latency contain up to 91% of the observed price differences, adjusted for transaction costs. The settlement latency is therefore shown to be an important market friction, that implies limits to arbitrage as it exposes arbitrageurs to price risk.

4.2 Existence of Arbitrage Opportunities

After providing an empirical evidence on the extent of bitcoin price differences, the focus of this section is on the magnitude of the arbitrage opportunity, that is, on the extent to which the the bitcoin price differences could be explained by exchange-specific liquidity and trading-related fees. To test the importance of these market frictions, the instantaneous arbitrage profit is maximized by optimally accounting for the available order book depth and trading-related fees. Thus, the trading quantity that maximizes the total arbitrage return is identified given the instantaneous bitcoin price difference and exchange-specific trading-related fees. This optimal trading quantity not only efficiently determines whether the order book depth is sufficient to cover the fees, but also the full extent of the arbitrage opportunity and thus violations of the law of one price.

Limit order books provide detailed information on the prices and quantities to capture the magnitude of arbitrage opportunities by quantifying potential arbitrage profits based on observed instantaneous price differences across exchanges and available market depth. To quantify instantaneous arbitrage between two exchanges, their 25-level limit order books are used on a minute-level. Given exchange pair ij, the arbitrage profit depends on the trading direction of arbitrage, i.e. on the exchange, where bitcoins are purchased or sold. Specifically, the ask-side depth profile $\mathcal{A}_t^i$, defined in Equation 2, of the order book $\mathcal{L}_t^i$ is considered on exchange i to buy bitcoins, and the bid-side depth profile $\mathcal{B}_t^j$, defined in Equation 1, of the order book $\mathcal{L}_t^j$ is considered on exchange j, where bitcoins are going to be sold. In this way, for the ease of notation, the trading direction of arbitrage is fixed to $i \to j$, where exchange i serves as a buy-side market, and exchange j - as a sell-side market. But the discussion below is analogous to the opposite trading direction $j \to i$.

Not every price difference corresponds to an arbitrage opportunity, just as not every arbitrage opportunity corresponds to a profit. A specific goal of this section is to formally characterize a scenario in which price differences would constitute arbitrage opportunities. At this stage, arbitrage profit from trading Q bitcoins on exchange pair ij is an instantaneous difference between nominal buying and selling prices of Q bitcoins at time t on exchange i and j respectively, *without* considering the fees (alternatively, assuming zero fees), stemming from exchanges and the blockchain network. The price, paid for Q bitcoins on exchange i is defined as a *nominal buying price* $P^{nom}(Q, \mathcal{A}_t^i)$, where $Q \in (0, \sum_{k=1}^{n} q_k^{A,i}]$, where n is a number of price levels in the order book. The nominal buying price depends on the quantity Q of bitcoins to be purchased and on the ask-side depth profile $\mathcal{A}_t^i$ of the LOB $\mathcal{L}_t^i$:

$$\begin{aligned} P^{nom}(Q, \mathcal{A}_t^i) = &\sum_{k=1}^{n} \mathbb{1}_{\{Q \geq \sum_{l=1}^{k} q_l^{A,i}\}} p_k^{A,i} q_k^{A,i} \\ &+ \sum_{k=1}^{n} \mathbb{1}_{\{\sum_{l=1}^{k-1} q_l^{A,i} < Q < \sum_{l=1}^{k} q_l^{A,i}\}} p_k^{A,i} [Q - \sum_{l=1}^{k-1} q_l^{A,i}] \end{aligned} \tag{15}$$

for $1 \leq l, k \leq n,\ l, k \in \mathbb{N}$ and $\sum_{l=1}^{k-1} q_l^{A,i} = 0$ for $k = 1$. The nominal selling price of Q bitcoins $P^{nom}(Q, \mathcal{B}_t^j)$ at exchange j is defined similarly using the bid-side depth profile $\mathcal{B}_t^j$ of the LOB $\mathcal{L}_t^j$. Then, the total instantaneous arbitrage profit from trading Q for exchange pair ij is defined as:

$$\pi_t^{ij}(Q) = P^{nom}(Q, \mathcal{B}_t^j) - P^{nom}(Q, \mathcal{A}_t^i) \tag{16}$$

where $Q \in (0, \min(\sum_{k=1}^{n} q_k^{A,i}, \sum_{k=1}^{n} q_k^{B,j})]$

Given limit order books of two exchanges at time t, the instantaneous *unit* price difference between two exchanges also depends on the trading quantity Q. Then, for two exchanges i and j at time t, the instantaneous unit price difference or the price difference of the Q-th bitcoin is defined as a function of Q:

$$\delta_t^{ij}(Q) := p_t^{B,j}(Q) - p_t^{A,i}(Q), \tag{17}$$

where $Q \in (0, \min(\sum_{k=1}^{n} q_{k,t}^{A,i}, \sum_{k=1}^{n} q_{k,t}^{B,j})]$, and $p_t^{B,j}(Q)$ and $p_t^{A,i}(Q)$ are the bid and ask prices corresponding to the price level of the Q-th bitcoin in the cumulative order depth of the bid-side

$(p_k^B, q_k^B)_t^j$ and ask-side $(p_k^A, q_k^A)_t^i$ depth profiles respectively. The unit the price difference $\delta_t^{ij}(Q)$ is also referred to as *marginal* profit at the Q-th bitcoin.

There is *no arbitrage opportunity* in the trading direction $i \to j$, if the highest bid price b_t^j of exchange j is lower than the lowest ask price a_t^i of exchange i at time t:

$$\begin{aligned}\delta_t^{ij}(\min(q_{1,t}^{A,i}, q_{1,t}^{B,j})) &= p_t^{B,j}(\min(q_{1,t}^{A,i}, q_{1,t}^{B,j})) - p_t^{A,i}(\min(q_{1,t}^{A,i}, q_{1,t}^{B,j})) \\ &= b_t^j - a_t^i \\ &:= \delta_{t,1}^{ij} < 0.\end{aligned} \tag{18}$$

$\delta_{t,1}^{ij}$ can also be referred to as the first level cross-market price difference. Since the orders with the highest bid and the lowest ask quotes are filled first, the entries of $\mathcal{B}_t^j$ and $\mathcal{A}_t^i$ are sorted in price descending and price ascending orders respectively. Then, the price difference $\delta_t^{ij}(Q)$ between two exchanges is decreasing with each bitcoin Q, placed deeper in the limit order book. Consequently, the first level cross-market price difference $\delta_{t,1}^{ij}$ is also the maximum price difference for all $Q \in (0, \min(\sum_{k=1}^{n} q_{k,t}^{A,i}, \sum_{k=1}^{n} q_{k,t}^{B,j})]$:

$$\delta_{t,1}^{ij} = \max_Q \delta_t^{ij}(Q) \tag{19}$$

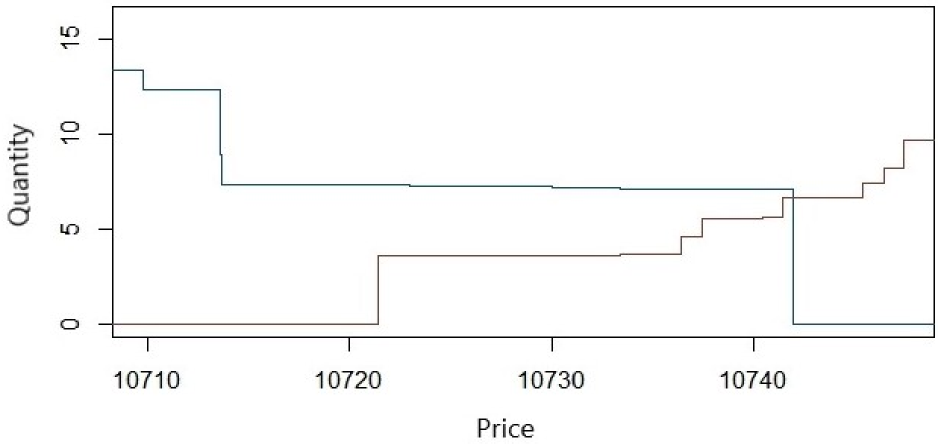
Since, as a function of Q, the negative price difference $\delta_t^{ij}(Q)$ between the exchanges decreases with Q, there is no such trading quantity Q of bitcoins that yield a positive arbitrage profit $\pi_t^{ij}(Q)$ in the trading direction $i \to j$. Thus, the first level cross-market price difference $\delta_{t,1}^{ij}$ constitutes a potential arbitrage opportunity in this trading direction only if it is positive:

$$\delta_{t,1}^{ij} > 0, \tag{20}$$

meaning that exchange i and exchange j are crossed, i.e. the best bid price on exchange j exceeds the best ask price of exchange i at time t. Condition 20 is a necessary, but not yet sufficient condition for existence of a positive arbitrage profit in the direction $i \to j$, whereas a negative $\delta_{t,1}^{ij}$ does not correspond to an arbitrage opportunity. A rational arbitrager would not exploit this price difference, therefore the arbitrage profit is set to zero. If taker fees are not zero, then the first level cross-market price difference $\delta_{t,1}^{ij}$, adjusted for the taker fees, may revert its sign to negative. In Section 4.4, more general price difference than $\delta_{t,1}^{ij}$, taker fee-adjusted first level price difference $\tilde{\delta}_{t,1}^{ij}$ is defined. However, $\tilde{\delta}_{t,1}^{ij} > 0$ is also not yet sufficient condition for a positive arbitrage profit, since BTC and USD withdrawal fees need to be considered. To determine, if an arbitrage opportunity constitutes a positive arbitrage profit, it must be examined if liquidity at both exchanges is sufficient to cover withdrawal fees (see Section 4.4 and 4.3).

Graphically, if the bid curve of exchange j intersects the ask curve of exchange i at time t, then the firs level price difference $\delta_{t,1}^{ij}$ is positive. Figure 4 illustrates the intersection of bid and ask curves of Bitfinex and Kraken, which indicates an open arbitrage opportunity in the trading direction Bitfinex → Kraken. The values on the x-axis are the unit prices $p_{k,t}^{B}$ and $p_{k,t}^{A}$, the y-axis is a corresponding cumulative order depth Q for each price level k. When the curves do not intersect, the first price difference $\delta_{t,1}^{ij}$ is negative, which means that there is no opportunity for arbitrage in this trading direction.

Figure 4. Arbitrage opportunity: asks on Bitfinex intersecting bids on Kraken



Notes: The Figure illustrates the intersection of bid (red) and ask curves (blue) of Bitfinex and Kraken on February 17, 2018, at 12:35:00. The values on the x-axis are the unit prices $p_{k,t}^{B}$ and $p_{k,t}^{A}$, the y-axis is a corresponding cumulative order depth Q for each price level k.

There are three mutually exclusive scenarios: (1) $\delta_{t,1}^{ij} > 0$, (2) $\delta_{t,1}^{ji} > 0$, (3) δ_t in both arbitrage directions is negative. The first two scenarios are co-determined because the quoted bid-ask spread $QS_t^{i(j)} = a_t^{i(j)} - b_t^{i(j)}$ is always positive: if $\delta_{t,1}^{ij} > 0$, then $\delta_{t,1}^{ji} < 0$, and vice versa.

4.3 Optimal Trading Quantity

A positive price difference alone does not guarantee a positive arbitrage profit and does not capture the full extent of arbitrage opportunity. An arbitrage opportunity might disappear, if the price difference and depth available behind the best prices is not sufficient to cover the fees, stemming from the exchanges and the blockchain network. In order to provide a full picture of arbitrage opportunities, a look deeper into the order book needs to be taken.

The Figure 4 shows that the unit price difference between two exchanges $\delta_t^{ij}(Q)$, where i is Bitfinex and j is Kraken, is still positive for the deeper orders, but decreasing with Q. Hence, the order book depth can be optimally accounted to maximize the arbitrage profit. In this respect, the full magnitude of arbitrage opportunity for exchange pair ij can be captured by identifying the trading quantity q^* that maximizes the arbitrage profit given the bid-side and ask-side depth profiles of corresponding exchanges at time t.

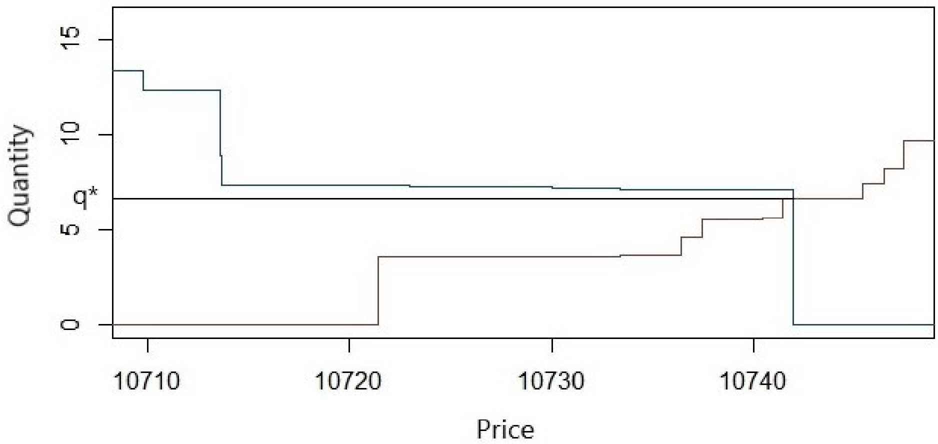
Let the condition $\delta_{t,1}^{ij} > 0$ be satisfied, which constitutes a potential arbitrage opportunity for exchange pair ij in the trading direction $i \rightarrow j$ at time t. Since the orders with the highest bid and the lowest ask quotes are filled first, the entries of $\mathcal{B}_t^j$ and $\mathcal{A}_t^i$ are sorted in price descending and price ascending orders respectively. Hence, the unit price difference $\delta_t^{ij}(Q)$, which is equivalent to the marginal profit of the Q-th bitcoin, decreases with Q. As a result, the total arbitrage profit $\pi_t^{ij}(Q)$, defined in Equation 16, grows with increasing quantity of bitcoins until the price difference $\delta_t^{ij}(Q)$ is zero or negative[3]. Thus, the quantity of bitcoins that minimizes the absolute price difference $\delta_t^{ij}(Q)$ or the marginal profit over Q maximizes the total arbitrage profit, and corresponds to the intersection of the bid and ask curves at time t:

$$\begin{aligned} \underset{Q}{\operatorname{argmin}} |\delta_t^{ij}(Q)| &= \underset{Q}{\operatorname{argmax}} [P^{nom}(Q, \mathcal{B}_t^j) - P^{nom}(Q, \mathcal{A}_t^i)] \\ &= \underset{Q}{\operatorname{argmax}}\, \pi_t^{ij}(Q) \qquad (21) \\ &:= q^*, \end{aligned}$$

where $Q \in (0, \min(\sum_{k=1}^{n} q_{k,t}^{A,i}, \sum_{k=1}^{n} q_{k,t}^{B,j})]$. Figure 5 shows how to determine the profit-maximizing trading quantity as the intersection point of the bid and ask curves.

[3]As a discrete function, the marginal profit at the quantity of bitcoins corresponding to the intersection of the bid and ask curves depends on the type of the curves' intersection. The types are illustrated by Figures 18, 19 and 20 in the Appendix.

Figure 5. Profit-maximizing quantity as an intersection point



Notes: The Figure shows the optimal trading quantity q^* in the intersection point of the bids on Bitfinex and and asks on Kraken on February 17, 2018, at 12:35:00.

In the numerical example shown by Figure 5, the profit-maximizing quantity is 6.7 BTC. It is not optimal to buy a quantity Q smaller than 6.7 BTC, since the price difference for the 6.7-th bitcoin $\delta_t(6.7)$ is still positive and yields a positive profit. For each bitcoin after the 6.7-th bitcoin, the difference between bid and ask prices is negative, resulting in a loss. Graphically, the value of the arbitrage profit corresponds to the area between the two curves.

4.4 Optimal Trading Quantity and Taker Fees

It needs to be investigated whether the trading quantity of bitcoins corresponding to the intersection of the bid ans ask curves is profit-maximizing after considering all the fees, stemming from exchanges and the blockchain network. According to the fee schedule of the centralized exchanges, the trader pays an exchange-specific *taker fee* (only market orders are considered) upon the execution of a trade. The fee is paid as a percentage of the trading volume, its amount therefore depends on the price and the trading quantity Q.

Effective prices, denoted by $P^{eff}(Q, \mathcal{B}_t^j, f_j)$ and $P^{eff}(Q, \mathcal{A}_t^i, f_i)$, are the selling and buying prices of Q bitcoins after subtracting and, respectively, adding exchange-specific taker fees. The effective price can be derived from the nominal price:

$$P^{eff}(Q, \mathcal{A}_t^i, f_i) = (1 + f_i)P^{nom}(Q, \mathcal{A}_t^i), \tag{22}$$

where the taker fees f_i are expressed as proportions. The effective price for selling Q bitcoins $P^{eff}(Q, \mathcal{B}_t, f_j)$ is defined analogously, but the fees f are obviously subtracted:

$$P^{eff}(Q, \mathcal{B}_t^j, f_j) = (1 - f_j) P^{nom}(Q, \mathcal{B}_t^j), \tag{23}$$

Finally, the *effective* first level cross-market price difference is the first level price difference, adjusted for the proportional taker fees, and defined as:

$$\tilde{\delta}_{t,1}^{ij} = (1 - f_j) b_t^j - (1 + f_i) a_t^i. \tag{24}$$

The taker fees shift the ask prices upward and the bid prices downward. Since the taker fees are proportional and the stated bid and ask quotes are the prices per unit, the taker fees linearly transform the bid and ask curves. The linear transformation corresponds to the rotation of the bid and ask curves: the bid curve becomes steeper and the ask curve becomes flatter by rotating both curves clockwise. It needs to be checked whether the quantity of bitcoins corresponding to the intersection of the shifted bid and ask curves is still profit-maximizing.

Proposition 1. Let the arbitrage profit for exchange pair ij at time t be denoted by $\pi_t^{ij}(Q)$, and let the effective prices P^{eff} be the buying and selling prices of Q bitcoins at exchange i and j after accounting for proportional exchange-specific taker fees f as defined in Equations 22 and 23.

The total effective arbitrage profit is an arbitrage profit adjusted for the takers fees, and can be written in the following way:

$$\tilde{\pi}_t^{ij}(Q) = P^{eff}(Q, \mathcal{B}_t^j, f_j) - P^{eff}(Q, \mathcal{A}_t^i, f_i), \tag{25}$$

where $\mathcal{B}_t^j$ and $\mathcal{A}_t^i$ are the bid-side and ask-side depth profiles in the order books $\mathcal{L}_t^j$ and $\mathcal{L}_t^i$ of exchange j and i at time t. The profiles are such that the condition of the positive taker fee-adjusted first level price difference $\tilde{\delta}_{t,1}^{ij} > 0$ is satisfied at time t, i.e. for exchange pair ij in the trading direction $i \to j$, there is an arbitrage opportunity after accounting for the taker fees.

Let $\mathcal{A}_{t,i}^{eff}$ and $\mathcal{B}_{t,j}^{eff}$ be the corresponding ask-side and bid-side depth profiles of exchange i and j at time t, where the ask and bid prices have been shifted by the exchange-specific taker fees f:

$$\mathcal{A}_{t,i}^{eff}(f_i) = ((1 + f_i) p_k^{A,i}, q_k^{A,i})_t, \tag{26}$$

and

$$\mathcal{B}_{t,j}^{eff}(f_j) = ((1-f_j)p_k^{B,j}, q_k^{B,j})_t. \tag{27}$$

The nominal prices $P^{nom}(Q)$ are defined as in Equation 15. Then there exists such $\mathcal{A}_{t,i}^*$ and $\mathcal{B}_{t,j}^*$, so that maximizing the total effective arbitrage profit $\tilde{\pi}_t^{ij}(Q)$, defined in Equation 25, by Q is equivalent to maximizing the profit:

$$\pi_{t,ij}^*(Q) = P^{nom}(Q, \mathcal{B}_{t,j}^*) - P^{nom}(Q, \mathcal{A}_{t,i}^*) \tag{28}$$

by Q, where $\mathcal{B}_{t,j}^* = \mathcal{B}_{t,j}^{eff}$ and $\mathcal{A}_{t,i}^* = \mathcal{A}_{t,i}^{eff}$. That is

$$\underset{Q}{\operatorname{argmax}}[P^{eff}(Q, \mathcal{B}_t^j, f_j) - P^{eff}(Q, \mathcal{A}_t^i, f_i)] = \underset{Q}{\operatorname{argmax}}[P^{nom}(Q, \mathcal{B}_{t,j}^{eff}(f_j)) - P^{nom}(Q, \mathcal{A}_{t,i}^{eff}(f_i))], \tag{29}$$

where $Q \in (0, \min(\sum_{k=1}^n q_{k,t}^{A,i}, \sum_{k=1}^n q_{k,t}^{B,j})]$ □

Proof. It is enough to consider only the bid part of the total effective arbitrage profit, i.e. the price $P^{eff}(Q, \mathcal{B}_t^j, f_j)$, because the derivation for the ask part $P^{eff}(Q, \mathcal{A}_t^i, f_i)$ are analogous using the definition of the effective price in Equation 22. Applying the definition of the effective prices in 23 and inserting the nominal price from Equation 15 gives

$$\begin{aligned}
P^{eff}(Q, \mathcal{B}_t^j, f_j) &= (1-f_j)P^{nom}(Q, \mathcal{B}_t^j) \\
&= (1-f_j)\sum_{k=1}^n \mathbb{1}_{\{Q \geq \sum_{l=1}^k q_l^{B,j}\}} p_k^{B,j} q_k^{B,j} \\
&+ (1-f_j)\sum_{k=1}^n \mathbb{1}_{\{\sum_{l=1}^{k-1} q_l^{B,j} < Q < \sum_{l=1}^k q_l^{B,j}\}} p_k^{B,j}[Q - \sum_{l=1}^{k-1} q_l^{B,j}] \\
&= \sum_{k=1}^n \mathbb{1}_{\{Q \geq \sum_{l=1}^k q_l^{B,j}\}} (1-f_j) p_k^{B,j} q_k^{B,j} \\
&+ \sum_{k=1}^n \mathbb{1}_{\{\sum_{l=1}^{k-1} q_l^{B,j} < Q < \sum_{l=1}^k q_l^{B,j}\}} (1-f_j) p_k^{B,j}[Q - \sum_{l=1}^{k-1} q_l^{B,j}] \\
&= P^{nom}(Q, \mathcal{B}_{t,j}^{eff}),
\end{aligned} \tag{30}$$

where $\mathcal{A}_{t,i}^{eff}$ and $\mathcal{B}_{t,j}^{eff}$ are defined in Equations 26 and 27. That is, the Equation 29 is satisfied, because

$$P^{eff}(Q, \mathcal{B}_t^j, f_j) - P^{eff}(Q, \mathcal{A}_t^i, f_i) = P^{nom}(Q, \mathcal{B}_{t,j}^{eff}(f_j)) - P^{nom}(Q, \mathcal{A}_{t,i}^{eff}(f_i)) \tag{31}$$

■

In the proof it can be observed that the effective price of selling Q bitcoins $(1-f_j)P^{nom}(Q,\mathcal{B}_t^j)$ is equivalent to the selling price of Q bitcoins, which is calculated based on the bid prices shifted by the taker fees downward:

$$(1-f_j)P^{nom}(Q,\mathcal{B}_t^j) = P^{nom}(Q,\mathcal{B}_{t,j}^{eff}(f_j)), \tag{32}$$

where $\mathcal{B}_{t,j}^{eff}(f_j) = ((1-f_j)p_k^{B,j}, q_k^{B,j})$. The same applies to the nominal buying prices $P^{nom}(Q,\mathcal{A}_{t,i}^{eff})$. This way, the quantity that maximizes the right hand-side of Equation 31 is equivalent to the quantity that maximizes its left-hand side. Graphically, the result shows that the optimal trading quantity of bitcoins corresponds to the intersection of the rotated bid and ask curves.

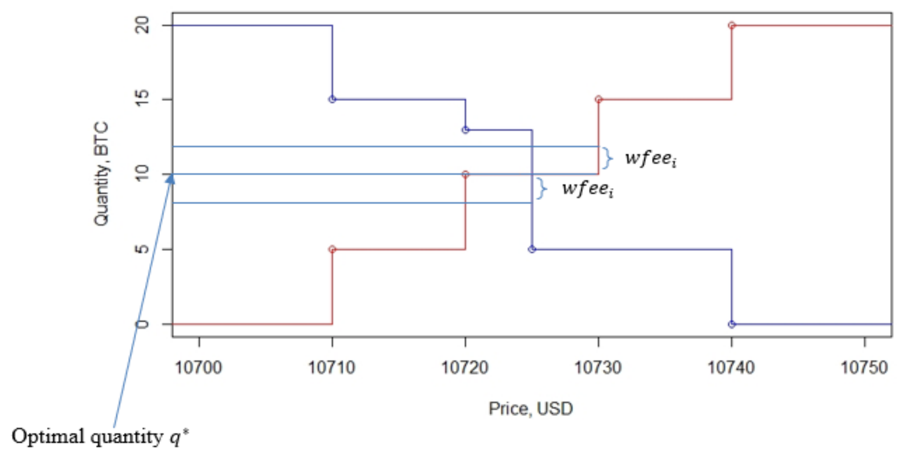
4.5 Optimal Trading Quantity and Bitcoin Withdrawal Fees

To exploit arbitrage opportunity, the trader has to transfer Q bitcoins from exchange i to exchange j. In the previous sections, the same amount of bitcoins Q was traded on both exchanges. However, since most of the exchanges charge a fee $wfee$ to withdraw bitcoins from the exchange's wallet, the trading amount is different. Bitcoin withdrawal fees are expressed and paid in bitcoins, and do not depend on the trading size. Although a measure of the arbitrage profit, constructed in the next section, evaluates the magnitude of the arbitrage profit the trader could achieve from simultaneous trading across two exchanges, the arbitrageur needs to transfer bitcoins from the exchange's wallet with high quantity of bitcoins (buy-side exchange) to the one with low quantity of bitcoins (sell-side exchange). The arbitrageur may wish or need to transfer bitcoins at any minute, therefore the bitcoin withdrawal fee will be considered every minute, that is, every minute the arbitrageur can sell only the bitcoin amount minus the withdrawal fee. In this way, the instantaneous profit measure will evaluate the minimum arbitrage profit available each minute after adjusting for the bitcoin withdrawal fees.

Without accounting for bitcoin withdrawal fees, it is profit-maximizing for the arbitrageur to buy q^* bitcoins which corresponds to the intersection of the bid and ask curves at time t. However, it needs to be investigated whether it is optimal to buy this quantity when a withdrawal fee is charged. The condition $\delta_{t,1}^{ij} > 0$ implies four different types of intersection of the bid and ask curves. Figure 6 provides a graphical illustration of the first type of intersection, Figures 18, 19 and 20 in Appendix illustrate the second, the third and the fourth types of intersection respectively. The first and second type of the curves' intersection are such that only a single amount of bitcoins corresponds to the intersection of the curves. The third type is a vertical tangency, i.e. the bid and ask curves have a set M of bitcoin quantities in their intersection, and its element q^* is defined as $\min_{m \in M} q_m$. The last kind of intersection is a horizontal tangency with the set of common points M, where the corresponding quantity of bitcoins is constant.

Regardless of the type of bid and ask curves intersection, the marginal profit of the q^*-th bitcoin $\tilde{\delta}_t(q^*)$ is positive. In the numerical example illustrated by Figure 6, $q^* = 10$ is the quantity that corresponds to the intersection of the taker fee-adjusted bid and ask curves. When bitcoins are withdrawn from the exchange i to transfer them to exchange j, $wfee_i$ is paid to exchange i, and the arbitrageur loses the part of the profit $(10725 - 10720) \cdot wfee_i$, because $10 - wfee_i$ instead of 10 bitcoins, can be sold on exchange j. Figure 6 shows that the marginal profit of selling the last 10-th bitcoin $\tilde{\delta}_t^{ij}(10)$ is greater than zero. In order to be able to sell 10 bitcoins on exchange j, the arbitrageur has to buy at least $wfee_i$ more bitcoins on exchange i. The Proposition 2 gives a condition under which it is worth buying extra $wfee_i$, i.e. $Q^B = q^* + wfee_i$ bitcoins for all four types of intersection of bid and ask curves, illustrated in Figures 6, 18, 19 and 20 in Appendix. In this respect, the Proposition 2 gives the condition under which it is profit-maximizing to trade Q^B bitcoins on exchange i.

Figure 6. Intersection in a point



Proposition 2. Consider the bid-side $\mathcal{B}_{t,j}^{eff}(f_j)$ and ask-side $\mathcal{A}_{t,i}^{eff}(f_i)$ depth profiles and the intersection point q^* of the corresponding taker fee-adjusted bid and ask curves. The profiles are such that the condition of the positive taker-fee adjusted first price difference $\tilde{\delta}_{t,1}^{ij} > 0$ is satisfied at time t, i.e. for exchange pair ij in the trading direction $i \to j$, there is an arbitrage opportunity after accounting for the taker fees. Let $wfee_i$ be a withdrawal fee of exchange i

in BTC. Then, it is profit-maximizing to buy $q^* + wfee_i$ bitcoins at exchange i and sell q^* at exchange j, if the bid price for the q^*-th bitcoin on exchange j is more than twice as high as its ask price on exchange i, i.e. $p^{B,j}(q^*) > 2p^{A,i}(q^*)$ at time t.

□

Proof. Since q^* is an intersection point of the bid and ask curves, the arbitrageur observes a positive price difference $\tilde{\delta}_t^{ij}(q^*)$ or, equivalently, a positive marginal profit at the q^*-th bitcoin. If the arbitrageur buys q^* bitcoins and pays a withdrawal fee $wfee_i$ at exchange i, the part of the arbitrage profit is lost, since the arbitrageur is not able to sell q^*, but $q^* - wfee_i$ bitcoins at exchange j. The lost part of the profit is defined as

$$\begin{aligned} \pi_t^{ij}(wfee_i) &= wfee_i \cdot \tilde{\delta}_t^{ij}(q^*) \\ &= wfee_i \cdot (p^{B,j}(q^*) - p^{A,i}(q^*)), \end{aligned} \tag{33}$$

where $p^{B,j}(q^*)$ and $p^{A,i}(q^*)$ are the bid and ask prices quoted at the price levels corresponding to the q^*-th bitcoin in the bid-side $\mathcal{B}_{t,j}^{eff}$ and ask-side $\mathcal{A}_{t,i}^{eff}$ depth profiles respectively. This part of the arbitrage profit can be regained, if the arbitrageur buys q^* and additional $wfee_i$ bitcoins, and pays a withdrawal fee $wfee_i$ at exchange i. Then the arbitrageur is able to sell q^* bitcoins at exchange j. The arbitrageur buys additional bitcoins, if the the cost of purchasing extra $wfee_i$ bitcoins is strictly smaller than the regained profit:

$$\pi_{t,ij}(wfee_i) - wfee_i \cdot p^{A,i}(q^* + wfee_i) > 0. \tag{34}$$

Inserting 33 and rearranging the inequality gives:

$$(p^{B,j}(q^*) - p^{A,i}(q^*)) - p^{A,i}(q^* + wfee_i) > 0. \tag{35}$$

According to the structure of the ask-side depth profile, asks are sorted in price ascending order. If the arbitrageur buys q^* and additional $wfee_i$ bitcoins, they either entirely correspond to the same price level as the q^*-th bitcoin, or entirely to the higher price levels, or to both. The ask price $p^{A,i}(q^* + wfee_i)$ is therefore always at least as high as the ask price $p^{A,i}(q^*)$. The inequality 35 can be therefore rewritten in the following way:

$$\begin{aligned} p^{B,j}(q^*) &> p^{A,i}(q^*) + p^{A,i}(q^* + wfee_i) \\ &\geq 2p^{A,i}(q^*). \end{aligned} \tag{36}$$

In the next step the lower boundary for $p^{B,j}(q^*)$ can be determined:

$$p^{B,j}(q^*) > 2p^{A,i}(q^*). \tag{37}$$

In order for the condition 34 to be satisfied, the inequality 37 must hold. Consequently, in order for the arbitrageur to optimally trade additional $wfee_i$ bitcoins, the bid price for the q^*-th bitcoin on exchange j must be more than twice as high as its ask price on exchange i.

■

To summarize, the arbitrageur is not able to fully exploit the price difference between two exchanges if a withdrawal fee $wfee_i$ is charged, because she loses a part of the arbitrage profit by selling $wfee_i$ bitcoins less than q^* bitcoins that she could profitably sell on exchange j. If the lost profit, associated with the unsold $wfee_i$ bitcoins is higher than 100%, then it is optimal to buy additional $wfee_i$ bitcoins on exchange i in order to be able to trade q^* bitcoins on exchange j. To get a profit higher than 100%, the bid price for the q^*-th bitcoin on exchange j must be more than twice as high as its ask price on exchange i, as it states by condition 37. Such a significant price difference $\tilde{\delta}_t^{ij}$ between two exchanges is not observed in the data sample (see Table 25 in Appendix). Therefore, the arbitrage profits are computed based on the optimal trading quantity q^* of bitcoins that corresponds to the intersection of the bid and ask curves. If the optimal trading quantity is below the withdrawal fee, then the trading quantity is set to zero.

Remark 1. The intersection of the bid and ask curves may be such that the withdrawal fee $wfee_i$, as the lost part of the arbitrage profit defined in Equation 33, corresponds to more than one price level of the bid-side and ask-side depth profiles.

Scenario 1. This case is illustrated by Figure 19 in Appendix. The withdrawal fee $wfee_i$ is divided into two parts on the bid-side: $wfee_{i,1}$ entirely corresponds to the same price level as the q^*-th bitcoin, the second part $wfee_{i,2}$ - to the price level below. Following the same steps as in the proof of Proposition 2, similar condition as inequality 37 is derived for the price of $wfee_i$, but involving two price levels:

$$p^{B,j}(q^*) \cdot wfee_{i,1} + p_k^{B,j} \cdot wfee_{i,2} > 2p^{A,i}(q^*) \cdot wfee_i, \tag{38}$$

where the price level k is such that $q^* - wfee_i < \sum_{l=1}^{k} q_l^{B,j} < q^*$ with $k = 1, ..., n$. The price level k is below the price level corresponding to the intersection point q^*. Then, the price level corresponding to the q^*-th bitcoin is $k+1$, i.e. $p^{B,j}(q^*) = p_{k+1}^{B,j}$. According to the structure of the bid-side depth profile, asks are sorted in price descending order. Therefore, the inequality 38 can be extended in the following way:

$$\begin{aligned} p_k^{B,j} \cdot wfee_i &= p_k^{B,j} \cdot wfee_{i,1} + p_k^{B,j} \cdot wfee_{i,2} \\ &\geq p^{B,j}(q^*) \cdot wfee_{i,1} + p_k^{B,j} \cdot wfee_{i,2} \\ &> 2p^{A,i}(q^*) \cdot wfee_i, \end{aligned} \tag{39}$$

since the bid price $p_k^{B,j}$ is at least as high as $p^{B,j}(q^*)$. Rearranging the last inequality gives:

$$p_k^{B,j} > 2p^{A,i}(q^*). \tag{40}$$

The inequality 40 means that in order for the arbitrageur to optimally trade additional $wfee_i$ bitcoins in scenario 1, the bid price corresponding to the price level k (the next level below the price level corresponding to the intersection point q^*) on exchange j must be more than twice as high as the ask price for the q^*-th bitcoin on exchange i.

Scenario 2. In this case, the withdrawal fee $wfee_i$ is divided into two parts on the bid-side and ask-side: $wfee_{i,1}$ entirely corresponds to the same price level as the q^*-th bitcoin, the second part $wfee_{i,2}$ - to the price level below. Following the same steps as in the proof of Proposition 2, similar condition as inequality 35 is derived for the price of $wfee_i$, but involving two price levels:

$$\begin{aligned} p^{B,j}(q^*) \cdot wfee_{i,1} + p_{k_j}^{B,j} \cdot wfee_{i,2} > p^{A,i}(q^*) \cdot wfee_{i,1} \\ + p_{k_i}^{A,i} + p^{A,i}(q^* + wfee_i), \end{aligned} \tag{41}$$

where the price level k_j is such that $q^* - wfee_i < \sum_{l=1}^{k_j} q_l^{B,j} < q^*$ with $k_j = 1, ..., n$ and the the price level k_i is such that $q^* - wfee_i < \sum_{l=1}^{k_i} q_l^{A,i} < q^*$ with $k_i = 1, ..., n$. The price levels k_i and k_j are the next levels below the price level corresponding to the intersection point q^*. According to the structure of the ask-side depth profile, asks are sorted in price ascending order. Therefore, the following inequality holds:

$$p^{A,i}(q^* + wfee_i) \geq p^{A,i}(q^*) \geq p_{k_i}^{A,i}. \tag{42}$$

Since the bid price $p_{k_j}^{B,j}$ is at least as high as $p^{B,j}(q^*)$, inserting the inequality 39 and rearranging the last inequality gives:

$$\begin{aligned} p_{k_j}^{B,j} \cdot wfee_i &> p^{A,i}(q^*) + p_{k_i}^{A,i} \cdot wfee_{i,1} + p_{k_i}^{A,i} \cdot wfee_{i,2} \\ &\geq 2p_{k_i}^{A,i} \cdot wfee_i. \end{aligned} \tag{43}$$

The final inequality gives the condition:

$$p_{k_j}^{B,j} > 2p_{k_i}^{A,i}. \tag{44}$$

Consequently, in order for the arbitrageur to optimally buy additional $wfee_i$ bitcoins in scenario 2, the bid quote at the price level k_j on exchange j must be more than twice as high as the ask quote at the price level k_i.

The inequalities 38 and 43 from both scenarios can be easily extended, if $wfee_i$ corresponds to more than two price levels as illustrated in Figure 19 in the Appendix. The last price level that could be involved in computation of the bid and ask price of $wfee_i$ is the first price level of the bid-side and ask-side depth profiles. Given the bid-side $\mathcal{B}_{t,j}^{eff}$ and ask-side $\mathcal{A}_{t,i}^{eff}$ depth profiles of exchange i and j at time t, the first level cross-market price difference is the highest price difference between exchanges over all trading quantities Q:

$$\max_Q \tilde{\delta}_t^{ij}(Q) = p_1^{B,j} - p_1^{A,i} = \tilde{\delta}_{t,1}^{ij}, \tag{45}$$

where $Q \in (0, \min(\sum_{k=1}^n q_k^{A,i}, \sum_{k=1}^n q_k^{B,j})]$. According to the Table 25 in the Appendix, there is no exchange pair in the sample, for which the highest bid on exchange j was more than twice as high as the lowest ask on exchange i. Since the price difference $\tilde{\delta}_t^{ij}(Q)$ decreases with Q, the conditions 37, 40 and 44 are never satisfied for this data sample.

Remark 2. As mentioned above, Figure 19 also shows a vertical tangency - the third type of intersection, where the bid and ask curves have a set of points M in their intersection, and its element q^* was defined as $\min_{m \in M} q_m$. However, the optimal trading quantity of bitcoins could be defined on the whole M until the marginal profit is negative. If more than $\min_M q$ bitcoins are traded, the optimality condition for q^* depends on the number of price levels to which $wfee_i$ corresponds. Only two cases need to be distinguished: when $wfee_i$ entirely corresponds to the same price level as the intersection set M as in Proposition 2, and the case, when it also involves the price levels below as in Remark 1. In the first case, it is never optimal to buy additional

$wfee_i$ because, using the Equation 35 from the proof of Proposition 2, it gives:

$$\underbrace{(p^{B,j}(q^*) - p^{A,i}(q^*))}_{=0} - p^{A,i}(q^* + wfee_i) = -p^{A,i}(q^* + wfee_i) < 0. \tag{46}$$

The bid price and ask price for the quantities larger than $\min_{m \in M} q_m$ are equal, so the condition is never satisfied. In the second case, the discussion is the same as in Remark 1.

4.6 Construction of a Measure of the Arbitrage Profit

After verifying the accuracy of the optimal amount in the presence of fees, the potential arbitrage profit can be expressed in a concise way. Provided that the bid-side $\mathcal{B}^{eff}_{t,j}$ and ask-side $\mathcal{A}^{eff}_{t,i}$ depth profiles of two exchanges i and j are such that the first price difference is positive $\tilde{\delta}^{ij}_{t,1} > 0$, and q^* is an optimal (profit-maximizing) quantity of bitcoins as defined in Equation 21, corresponding to the intersection point of the bid and ask curves, then the arbitrage profit in the trading direction $i \to j$, is calculated by the formula:

$$\pi^{ij}_t(q^*, wfee_i, dwfee_{ij}) := P^{nom}(q^* - wfee_i, \mathcal{B}^{eff}_{t,j}) - P^{nom}(q^*, \mathcal{A}^{eff}_{t,i}) + dwfee_{ij}, \tag{47}$$

where $dwfee$ is the USD withdrawal fee difference:

$$dwfee_{ij} := dwfee_i - dwfee_j, \tag{48}$$

where $dwfee_i$ and $dwfee_j$ are USD withdrawal fees of exchange i and j respectively. In terms of USD withdrawal fee, this difference shows how advantageous arbitrage is in the trading direction $i \to j$. More precisely, it shows an advantage of withdrawing USD from exchange j, and not from exchange i. If $\tilde{\delta}^{ij}_{t,1} > 0$ and q^* is large enough to cover the taker and bitcoin withdrawal fees, there is already an advantage of a price difference between two exchanges in the trading direction $i \to j$. If, in addition, $dwfee_{ij}$ is positive, then, transfer of bitcoins to exchange j is even more favourable, because it is cheaper to withdraw USD from exchange j than from exchange i. If $dwfee_{ij}$ is negative, then the trading direction $i \to j$ is not advantageous in terms of USD withdrawal fee, and the corresponding value is deducted from the profit.

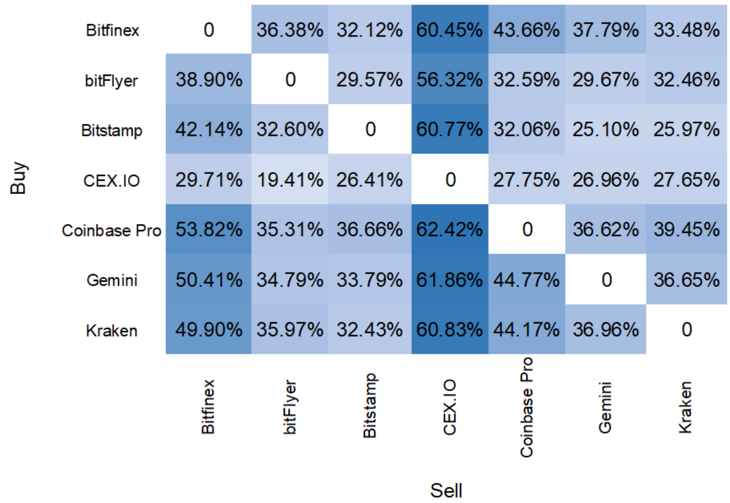
In summary, the potential arbitrage profit for exchange pair ij has the following four determinants: (i) $\delta^{ij}_{t,1}-$ the first level cross-market price difference in the trading direction $i \to j$; (ii) $\delta^{ji}_{t,1}-$ the first level price difference in the trading direction $j \to i$; (iii) liquidity at both exchanges: the bid-side depth available at the relevant prices p_k^b at time t: q_k^b, where the relevant price levels k are such that $\sum_{n=1}^{k} q_n^b \leq q^*$ with $k = 1, ..., n$, and the ask-side depth q_l^a, available at relevant prices p_l^a, where the relevant price levels l are such that $\sum_{n=1}^{l} q_n^b \leq q^*$ with $l = 1, ..., n$; (iv) the exchange-specific fees arising from transacting, such as proportional taker fees, BTC and USD withdrawal fees.

At least the first two conditions, as discussed above, are co-determined. The interplay of these determinants is necessary to ensure that the profit is sufficient to cover the fees.

4.7 Descriptive Analysis

Figure 22 in the Appendix plots the quantified instantaneous minute-level arbitrage profits, defined in the previous section, over time. Each plot shows the arbitrage profits in both trading directions, where the upper part of the graph corresponds to the direction indicated in the title of the plot, and the lower part of the graph corresponds to the opposite trading direction. The panels show the minute-level arbitrage profits $\pi_t^{ij}(q^*, wfee_i, dwfee_{ij})$ by exchange pair ij, based on optimal trading quantity and adjusted for the exchange-specific trading fees. As it was described in Section 4, arbitrage profits can be positive only in one trading direction $i \rightarrow j$, but they can be zero in both trading directions at the same minute for a given exchange pair ij, as determined by the first level price differences $\delta_{t,1}^{ij}$. The instantaneous first level price difference can be negative in both trading directions. If it is positive in the direction $i \rightarrow j$, then the first level price difference $j \rightarrow i$ is negative in the opposite direction, because it is co-determined by the fact that the bid-ask spread QS_t is always positive for each exchange. Since the negative first level price differences do not correspond to an arbitrage opportunity, they are set to zero. Table 25 in the Appendix provides summary statistics of the arbitrage profits $\pi_t^{ij}(q^*, wfee_i, dwfee_{ij})$, the first level price differences $\delta_{t,1}^{ij}$ and optimal quantities q^* for all 42 exchange pairs over the sample period. Figure 7 shows a heatmap of the frequencies of the positive instantaneous first level price differences $\delta_{t,1}^{ij}$, whereas Figure 8 displays a heatmap of frequencies of $\tilde{\delta}_{t,1}^{ij}$ - the positive instantaneous first level price differences, adjusted for the taker fees. Table 4 reports the detailed descriptive statistics of the persistence and the magnitude of arbitrage opportunities.

Figure 7. Frequencies of the positive first level price differences



Notes: The heatmap shows the exchange pair-specific percentages of minutes in the sample period on which the instantaneous first level price difference $\delta_{t,1}^{ij}$, based on the highest bid and the lowest ask quotes according to Equation 19, is positive. A positive first level price difference corresponds to a potential instantaneous arbitrage opportunity in a given trading direction. The sample period starts from 02:00:00, March 29, 2018 to 01:00:00.

Figure 8. Frequencies of the positive taker fee-adjusted first level price differences

Buy \ Sell	Bitfinex	bitFlyer	Bitstamp	CEX.IO	Coinbase Pro	Gemini	Kraken
Bitfinex	0	11.39%	0.36%	33.60%	0.38%	0.002%	0.08%
bitFlyer	6.97%	0	2.24%	30.83%	1.73%	0.03%	2.77%
Bitstamp	0.91%	5.65%	0	33.22%	0.21%	0.05%	0.12%
CEX.IO	13.45%	7.52%	9.75%	0	9.11%	0.25%	9.93%
Coinbase Pro	1.21%	5.78%	0.10%	32.92%	0	0.01%	0.07%
Gemini	0.01%	0.4%	0.04%	14.00%	0.08%	0	0.01%
Kraken	0.8%	7.65%	0.11%	33.63%	0.19%	0.03%	0

Notes: The heatmap shows the exchange pair-specific percentages of minutes in the sample period on which the instantaneous first level price difference $\tilde{\delta}^{ij}_{t,1}$, based on minute-level, taker fee adjusted bid and ask quotes according to Equation 24, is positive. The taker fees are given in Table 10.

Figure 9. Frequencies of the positive arbitrage profits

Buy \ Sell	Bitfinex	bitFlyer	Bitstamp	CEX.IO	Coinbase Pro	Gemini	Kraken
Bitfinex	0	5.16%	0.28%	33.60%	0.30%	0.002%	0.04%
bitFlyer	4.00%	0	1.18%	24.88%	1.02%	0.02%	1.48%
Bitstamp	0.89%	5.65%	0	33.20%	0.21%	0.05%	0.12%
CEX.IO	10.79%	3.40%	7.24%	0	6.89%	0.02%	7.40%
Coinbase Pro	1.21%	5.77%	0.10%	32.90%	0	0.01%	0.07%
Gemini	0.004%	0.20%	0.04%	10.66%	0.07%	0	0.003%
Kraken	0.61%	2.97%	0.08%	31.36%	0.19%	0.01%	0

Notes: The heatmap shows the exchange pair-specific percentages of minutes in the sample period on which the instantaneous arbitrage profit $\pi^{ij}_t(q^*, wfee_i, dwfee_{ij})$, defined in Equation 47, is positive.

According to Figure 9, more than 98% of the minute-level arbitrage profits are zero for all exchanges, except CEX.IO- and bitFlyer-related exchange pairs. As expected, Bitfinex, Kraken, Bitstamp, Coinbase Pro and Gemini exhibit very low frequencies of profitable arbitrage opportunities as their prices are more integrated (see Figure 2). As it was discussed in Section 4, a price difference between exchanges corresponds to a potential arbitrage opportunity, if the first level price difference between two exchanges, defined in Equation 19, is positive. In this way, there are two main reasons that the frequencies of arbitrage profits approach zero. The first reason is that the first level price differences δ_t at the best bid and ask prices, *not* adjusted for the taker fees, do not constitute arbitrage opportunities in 50%-80% of the minutes in the sample period (see Figure 7). An exception is the CEX.IO-related exchange pairs with CEX.IO on the sell-side – around 60% of the first level price differences are positive and associated with 10%-35% of the minutes on which arbitrage profits are positive.

The second factor is the exchange-specific proportional *taker fee*. Given a positive first level price difference between two exchanges, an arbitrage opportunity disappears when the price difference is not high enough to cover the taker fees. The heatmap in Figure 8 shows that for all exchange pairs, except CEX.IO- and bitFlyer-related exchanges, more than 98% of the first level price differences do not constitute arbitrage opportunities after adjusting for the taker fees. Indeed, Table 23 in the Appendix reports that the average magnitude of the positive first level price differences is 0.01%-0.1% (excluding CEX.IO and bitFlyer), which is below the taker fees ranging from 0.15% to 0.26%. The average, taker fee adjusted, first level price difference is around 0.01% (StDev = 0.05%), besides CEX.IO with 0.2% (StDev = 0.4%, see Table 4). The CEX.IO-related exchange pairs are again an exception - at most 50% of the positive first level price differences revert the sign after adjusting for the taker fees. Gemini stands apart with the highest taker fee of 1%, making the Gemini-related exchange pairs the least profitable exchange-pairs for arbitrage. Specifically, Gemini-related exchange pairs yield the smallest and the least frequent time-weighted average of arbitrage profits in both trading directions, compared to other exchange pairs. Although the frequency and the extent of the positive first level price differences, not adjusted for the taker fees, is similar to other exchanges, the arbitrage opportunities on Gemini yield a profit only in a couple of times during the observation period (see Figure 22). On average, there is only one profitable minute per day, as shown by Table 4.

Whereas the taker fees reduce the frequency of the profitable arbitrage opportunities by 50%-99%, the USD and BTC withdrawal fees - by up to 50%, after optimally accounting for the liquidity at the relevant price levels of the corresponding order books (see Section 4). It

Table 4. Summary statistic of the duration and magnitude of arbitrage opportunities

Sell-side Exchange	Mean	StdDev	p5	Median (p50)	p95	Max
Bitfinex						
Duration	15.19	61.20	2.08	4.83	26.57	1307.83
$\tilde{\delta}$	0.01	0.05	< 0.01	< 0.01	0.08	2.75
ϕ	3.78	10.80	< 0.01	0.09	23.29	71.60
bitFlyer						
Duration	13.24	40.26	2.00	4.33	24.80	743.83
$\tilde{\delta}$	0.02	0.09	< 0.01	< 0.01	0.09	2.98
ϕ	6.20	15.53	< 0.01	0.01	39.59	75.84
Bitstamp						
Duration	8.36	29.37	2.00	2.67	11.25	479.00
$\tilde{\delta}$	0.01	0.05	< 0.01	< 0.01	0.05	3.73
ϕ	2.04	6.08	< 0.01	0.01	15.17	35.81
CEX.IO						
Duration	53.13	359.16	2.00	4.50	37.67	7273.67
$\tilde{\delta}$	0.20	0.40	< 0.01	< 0.01	1.06	5.00
ϕ	29.16	38.71	< 0.01	3.40	95.98	99.93
Kraken						
Duration	9.27	32.53	2.02	2.92	14.62	494.83
$\tilde{\delta}$	0.01	0.04	< 0.01	< 0.01	0.05	3.05
ϕ	2.11	5.98	0.001	0.01	15.66	33.85
Gemini						
Duration	7.27	< 0.01	6.17	6.67	8.97	14.17
$\tilde{\delta}$	< 0.01	0.01	< 0.01	< 0.01	< 0.01	2.56
ϕ	0.07	0.42	< 0.01	< 0.01	0.17	4.63
Coinbase Pro						
Duration	14.45	54.66	2.00	3.17	20.67	640.83
$\tilde{\delta}$	0.01	0.05	< 0.01	< 0.01	0.04	3.38
ϕ	1.91	6.32	< 0.01	0.01	14.35	38.91

Notes: This table provides summary statistics of the following variables: $\tilde{\delta}$ is the taker fee-adjusted first level price difference in %, $Duration_t$, which denotes the average duration of the positive taker fee-adjusted first level price differences on day t (in minutes), and ϕ measures the likelihood of arbitrage opportunity, which is the number of arbitrage opportunities (when $\tilde{\delta} > 0$) on day t, divided by 1440 (the number of minutes in a day. The columns present the mean, the standard deviation, the 5% quantile, the median, the 95% quantile and the maximum values. The summary statistics are averaged across exchange pairs with a fixed sell-side exchange.

means that at least in 50% of cases, the order book depth is large enough to cover the fees. The smallest reductions are observable on CEX.IO-related exchanges, indicating higher price differences. Indeed, the highest magnitude and frequency of the profitable arbitrage opportunities is observed on CEX.IO-related exchange pairs, when CEX.IO serves as a sell-side exchange. As shown by Tables 23 and 25 in the Appendix, the time-weighted average of the positive first level price differences ranges from 0.4% to 0.8%, the maximum value - from 5% to 6%, which is at least two time higher than on other exchanges. Although CEX.IO has a relative high BTC and USD withdrawal fee and a relative small average quoted depth (see Table 2 in Section 3.2), sufficiently high price differences allow the bid and ask curves to intersect at a greater depth, as shown by the optimal trading quantities in Table 25. As a result, more than 95% of the taker-fee adjusted first level price differences cover the withdrawal fees of CEX.IO and yield positive arbitrage profits.

In the following sections, platforms with the most pronounced arbitrage profits are discussed in more detail. CEX.IO, bitFlyer and Bitfinex exhibit arbitrage profits, especially when they serve as sell-side markets.

4.7.1 Arbitrage Profits on CEX.IO

Table 25 in the Appendix confirms that CEX.IO-related exchange pairs are also the most profitable exchange pairs during the observation period, when CEX.IO serves a sell-side market. It yields at least 10 times higher time-weighted average of arbitrage profits than other exchanges on the sell-side, ranging from $30 to $400. The highest profits on CEX.IO can reach $4,000-$6,000, if a few spikes of more than $10,000 are not considered. Thus, the highest time-weighted average of arbitrage profits is associated with the highest average first level price difference and the highest average optimal trading quantity. However, CEX.IO as a sell-side market does not yield the highest average arbitrage profit *per arbitrage opportunity*, although its average price difference per opportunity (from 0.6% to 0.8%) is the highest. The highest average of arbitrage profits (around $750-$3,000) can be reached on Bitstamp due to several spikes at $20,000 (see Figure 22. Such a high profit is achieved on this exchange due to very high maximum optimal quantity, that is, very large order sizes available at the relevant prices. However, only 2% of the minutes of the sample period are profitable for arbitrage on Bitstamp, which constitutes on average 28 profitable minute-level arbitrage opportunities per day (see Table 4). This suggests that the price differences on Bitstamp are mainly driven by price volatility. In contrast, CEX.IO as a sell-side market is profitable for arbitrage on average 7 hours per day with average arbitrage profit from $700 to $1000 per opportunity. It corresponds to $294,000 per day (at 7 hours or 420 minutes per day of arbitrage opportunities, on average). Thus, CEX.IO is a profitable sell-side market since it consistently trades at higher bid prices relative to other exchanges, and the price differences with this exchange are the highest.

When CEX.IO is used as a buy-side market, arbitrage profits occur with a frequency of about 3%-10% of the sample period, which is again considerably higher compared to other exchange pairs. CEX.IO exhibits the lowest ask prices in the first half of July and in the last half of August. In this period CEX.IO serves as a profitable buy-side market and, after adjusting for the taker fees, exchange-pairs exhibit arbitrage opportunities in approximately 10% of minutes in the observation period. However, when CEX.Io is used as a buy-side market, the first level price differences are not large enough to reach the same magnitude of arbitrage profits as profits when CEX.IO serves as a sell-side market.

4.7.2 Arbitrage Profits on bitFlyer

Arbitrage opportunities associated with positive profits also frequently arise on exchange pairs involving bitFlyer. According to Figure 2 in Section 3.3, bitFlyer exhibits higher bid prices that deviate quite systematically in May and June, but in July the price difference reverts the direction, and bitFlyer quotes lower bid prices until the end of the sample period. In this distinct period from May to July, both CEX.IO- and bitFlyer-related exchange-pairs (with CEX.IO and bitFlyer on the sell-side) exhibited the highest and most persistent arbitrage profits. The timeline of these profits closely follows the cumulated price differences in Figure 2 in Section 3.3. The higher positive slope of the cumulated price difference curves of CEX.IO and bitFlyer indicates that these exchanges have consistently traded bitcoin at a premium in this period. As shown by Table 25 in the Appendix, the time-weighted average of the first level cross-market price differences on bitFlyer ranges from 0.07% to 0.11%, whereas the maximum value from 2% to 4%. As a result, up to 6% of the arbitrage profits are positive for bitFlyer-related exchanges with bitFlyer on the sell-side. This percentage corresponds to 1.5 hour per day on which the arbitrage profits are positive. However, the average profit per arbitrage opportunity is the smallest among all exchange pairs, ranging from $13 to $110.

4.7.3 Arbitrage Profits on Bitfinex

Besides CEX.IO and bitFlyer, the first level cross-market price differences are quite often positive when Bitfinex is used as a sell-side market: the frequency of its arbitrage opportunities is even comparable to CEX.IO-related exchanges – between 30% and 50% (see Figure 7). However, as reported by Table 25 in the Appendix, the price differences are not high enough to cover the proportional taker fees, yielding at most 10% positive arbitrage profits with CEX.IO on the buy-side. On average, there is one per day on which arbitrage profits are positive, when Bitfinex serves as a sell-side market. Per arbitrage opportunity, the profit ranges from $100 to $400.

4.7.4 Dynamics and Persistence of Arbitrage Profits

The dynamics and persistence of the arbitrage profits are also of particular interest. The highest and most frequent arbitrage profits are particularly noticeable in the first half of the sample period, starting on March 29 until the end of July. As expected, the frequency and the magnitude of arbitrage profits clearly drops over the sample since the prices across exchanges have become more integrated in the second half of the sample period. The low frequency and low duration of arbitrage opportunities on all exchanges in the sample, except CEX.IO and bitFlyer, suggest that the arbitrage opportunities are mainly driven by bitcoin price volatility. On average, the likelihood of occurrence of the arbitrage opportunity ranges from 10 minutes on Gemini to 54 minutes on Bitfinex in a given day (excluding CEX.IO and bitFlyer), even though the maximum deviations can reach 1031 minutes on Bitfinex, which corresponds to more than half of a day. As Table 4 reports, the average duration of the taker fee-adjusted first level price difference for all markets ranges from 7 to 15 minutes, but CEX.IO stands apart. Price deviations that correspond to arbitrage opportunities on CEX.IO persist, on average, 53 minutes. In the Bitcoin market, price deviations are thus more persistent than in the foreign exchange (FX) market with 1.55 seconds per day of the average duration of arbitrage opportunities (Foucault et al., 2017), but comparable to the ADR market with the average duration of price deviations of 12 minutes (Rösch, 2021).

4.7.5 Summary of Results

In summary, CEX.IO is the most profitable sell-side exchange across all exchanges in the sample, because arbitrage opportunities occur most frequently and last the longest when CEX.IO used as a sell-side market. Although the price differences on CEX.IO are the highest, the average arbitrage profit per opportunity is comparable to other exchanges. As shown by Figure 24, some exchange pairs are heterogeneous with regard to the magnitude of the profit per arbitrage opportunity. It ranges from $10 on Gemini to $2,000 on Bitstamp, which on average constitutes from $10 to $60,000 of arbitrage profits per day. Makarov and Schoar (2020) calculated arbitrage profits for the cryptocurrency exchanges from different geographical regions for the period between

November 2017 and February 2018. They show that average daily arbitrage profits between Europe and the US cryptocurrency exchanges - Coinbase, Kraken, Bitstamp, and Gemini - are around $500,000, which is significantly higher than the profits on the most profitable exchange in the sample - CEX.IO (average daily arbitrage profit is around $300,000). This means that three months ago the cryptocurrency markets were even more profitable for arbitrage, but the prices have become more integrated, as can be also seen on Figure 2. In the sample period from April 2018 to September 2018, average positive first level prices difference ranges between 0.01% and 0.8% (see Figure 23), whereas Makarov and Schoar (2020), in their sample period between November 2017 and February 2018, documented the average price discrepancies of 2% between exchanges that operate in the US and Europe. This shows a general trend of decreasing price differences and increasing efficiency of cryptocurrency markets.

In comparison to other financial markets, arbitrage profits on cryptocurrency market are higher, at least for some exchanges. Foucault et al. (2017) report that the average arbitrage profit from triangular[4] arbitrage opportunities within Reuters D-3000 is $7,291,2 per day (at 112 opportunities per day, on average). Brogaard et al. (2014) documented that, after incorporating NASDAQ trading fees, high-frequency traders in their sample earn, on average, $4,209.15 per stock-day on their market orders in large-cap stocks (see Table 5 in Brogaard et al. (2014)).

5 Discussion on the Sources of Bitcoin Price Differences

In the previous section it is shown that for all exchange pairs, except CEX.IO- and bitFlyer-related exchanges, more than 98% of the positive first level cross-market price differences do not constitute arbitrage opportunities after adjusting for the taker fees. However, CEX.IO-related exchange pairs exhibit 10%-35% of the minutes on which arbitrage profits are positive, bitFlyer-related exchange pairs - up to 6%. Average taker fee-adjusted price difference with CEX.IO is 0.2% and persists, on average, for 53.13 minutes. The taker fee-adjusted price difference on bitFlyer is on average 0.02% and it's average duration is 13.24 minutes. Whereas the taker fees reduce the frequency of the profitable arbitrage opportunities by 50%-99%, the USD and BTC withdrawal fees - by up to 50%, after optimally accounting for the liquidity at the relevant price levels. It means that at most in 50% of cases, the order book depth is large enough to cover the fees. When a profit-maximizing volume is traded, these arbitrage opportunities produce sizable

[4]Foucault et al. (2017) use data from Reuters D-3000 from January 2, 2003 to December 30, 2004 for three currency pairs: US dollar/euro (dollars per euro; hereafter USD/EUR), US dollar/pound sterling (USD/GBP), and pound sterling/euro (EUR/GBP).

profits for the trader: \$700-\$1000 per opportunity is achievable on CEX.IO, on average. This suggests the existence of additional market frictions as liquidity, measured in depth of the limit order book, and transaction costs can not fully explain all arbitrage opportunities on CEX.IO and bitFlyer. If the restriction of the same price of bitcoin, as a homogeneous asset with the same payoffs, is violated, then at least two conditions must be satisfied. First, there must be arbitrage constraints that create impediments that prevent arbitrageurs from fully taking advantage of the market inefficiency (Shleifer, 2000; Barberis and Thaler, 2003), in other words, limits to arbitrage potentially explain why price differences persist (see Gromb and Vayanos (2010)). Second, following Gromb and Vayanos (2010), there must be reasons why bitcoin has diverging prices in the first place, that is, reasons for bitcoin prices to deviate from fundamental value. The goal of this section is to analyze the second condition: Section 5.1 and 5.2 present examples of nonfundamental demand shocks, emphasizing that they arise from institutional frictions. The first condition of limits to arbitrage will be discussed in Section 5.3.

5.1 Risk Premia

Gromb and Vayanos (2010) states that price discrepancies are "commonly interpreted as arising because demand shocks, experienced by investors other than arbitrageurs, push prices away from fundamental values, and arbitrageurs are unable to absorb such shocks and correct price discrepancies". Also according to Gromb and Vayanos (2010), such nonfundamental demand shocks are attributed not only to investor irrationality[5], but can also arise because of institutional frictions. One example of the institutional friction on the Bitcoin market is uncertainty of investors in the quality of managers and the security of a cryptocurrency trading platform. The likelihood of compliance failure, bankruptcy or security breach of a cryptocurrency exchange is a *counterparty risk* or institutional risk that an investor will be exposed to when she trades on a centralized cryptocurrency exchange. The evaluation of the magnitude of the counterparty risk is often determined by the perception of the safety and the confidence of investors in the quality of managers of the exchange, which in turn depends on whether the exchange has already been attacked by hackers, whether it attracts the illegal activity and the ability to quickly withdraw unlimited amount of funds from the exchange. Each exchange has it's own security measures to limit the attack surface and illegality-motivated trading. In this way, the bitcoin exchanges are heterogeneous with respect to their institutional risks. Concerns about the counterparty risk may affect the volume of the Bitcoin exchange. Moore and Christin (2013) conducted a

[5]See the survey by Barberis and Thaler (2003) for behavioral explanations for the anomalies.

survival analysis on 40 Bitcoin exchanges and found that an exchange's average transaction volume is negatively correlated with the probability it will close prematurely. In other words, less popular exchanges are more likely to be shut than popular ones, which implies negative correlation between counterparty risk and an exchange's trading volume or market share.

The counterparty risk arises in all centralized cryptocurrency exchanges since the customers of the centralized exchanges do not have access to the private keys of their wallets and their funds are held by the single company running the exchange. On centralized exchanges, the vulnerability of a central network comes into play: the exchanges could be faced with external hacks (thieves need only hack into the data centres of these exchanges) and exchange failures caused by incompetence of the exchange operators and management or by technical glitches. For example, according to Russell (2015) and (Kroeger and Sarkar, 2017), "hackers successfully transferred funds out of Bitstamp accounts, although Bitstamp was able to cap its losses at $5.1 million and continue to operate and honor all customer claims". In August 2016, the Hong-Kong based exchange Bitfinex, currently one of the largest BTC/USD exchanges, lost around 70 million US dollars of its customers' bitcoins. Bitfinex continued operating, but reduced all client account balances by 36% (Tepper, 2016). In February 2014, after a security breach, approximately 850,000 bitcoins belonging to customers were stolen (450 millions US dollars at the time) from Mt.Gox, the largest cryptocurrency exchange by transaction volume at that time (Records, 2014). Mt.Gox declared bankruptcy and, the hack of Mt.Gox remains the most extreme example of a counterparty risk. More examples of temporary and permanent exchange shut-downs are documented by Moore and Christin (2013), the authors find that by early 2013, 45% of Bitcoin exchanges had closed, often with customer balances wiped out, and many of the remaining markets were subject to frequent outages and security breaches. There is an evidence that bitcoin traded at a discount following hacks. For example, immediately after the security breach on Bitfinex, the price of bitcoin on Bitfinex has dropped 10% (BBC, 2016). After the Mt.Gox hack, the prices dropped to below 20% of the prices on other exchanges, reflecting the market's perception of the small chance that Mt.Gox retrieve its customers' bitcoins (Kannenberg, 2014). The risk of losing account balances differs across exchanges, and this may lead to differences in premiums compensating traders for the risk of exchange failure.

In addition to the problems associated with security breaches, some cryptocurrency trading platforms attract illegality-motivated trading by lacking KYC (know-your-customer) and AML (anti-money laundering) compliance which reduce the anonymity of customers by design. Illegality-motivated trading can entail the government interference to prevent capital outflow and money laundering, leading to withdrawal delays, frozen accounts and, eventually, to the closure of the exchange and significant losses of exchange's customers. For example, on 12 January 2018, the news that South Korea could be preparing to ban trading in digital coins were followed by the depreciation of bitcoin by 6% (Kharpal, 2018). Another example is cryptocurrency exchange BTC-e, where bitcoin consistently traded at a discount relative to prices on other exchanges, attracted trades by lower levels of compliance with with KYC/AML regulations. Risk premium arose to compensate traders for a counterparty risk, and also the trades, associated with illegal activities, generated selling pressure, pushing prices on BTC-e below the prices on other exchanges (Kroeger and Sarkar, 2017). On 25 July 2017, then the founder of BTC-e Alexander Vinnik was arrested on charges of money laundering and the server equipment at one of their data centres was seized. These events led to the shut-down of the BTC-e exchange (U.S. Attorney's Office, 2017).

The anonymity of cryptocurrency attracts illegality-motivated traders, which deters banks to get tied up with cryptocurrency exchanges and to process bitcoin transactions. This can lead to the lack of *fiat liquidity*, which in turn may lead to price differences between exchanges.

5.2 Fiat Liquidity

Another example of the demand shock on the Bitcoin market also arises from the institutional friction - *fiat liquidity*. The exchange sample consists of the relatively few trading platforms that allows customers to deposit and withdraw fiat currency: US dollars, euros, yen, and others. Withdrawal policy of a fiat-to-crypto exchange implies the existence of several methods for withdrawing funds from the exchange's account. Typically, the fiat payment methods include Bank transfer, VISA, MasterCard, SEPA (for EUR) and other options. In this respect, fiat-to-crypto exchanges have to find a bank that would let them open an account or use other credit and payment facilities to withdraw and deposit funds of their customers. However, banks wary of cryptocurrency exchanges due to heavy compliance requirements and the anonymity of digital currencies which attracts illegality-motivated traders. For example, according to (Casey, 2019), Hong Kong-based exchange Bitfinex was challenged to find a bank that would let them open an account. The US banks were demanding that their counterparts in Hong Kong apply especially high KYC standard for cryptocurrency exchanges. The easiest way for the banks in Hong Kong was, therefore, to refuse to deal with Bitfinex (Casey, 2019). Also Bitfinex was not fully KYC compliant, attracting customers to anonymity of cryptocurrencies, but deterring banks from establishing relationships and offering financial services (Vigna, 2019). The lack of access to the bank accounts for fiat liquidity forces cryptocurrency exchanges to use less reliable payment facilities. For example, in the observation period of March 29th, 2018 to September 5th, 2018, two exchanges in the sample - Bitfinex and CEX.IO used banking services of a Panamanian payment processing firm Crypto Capital. By 2018, Bitfinex placed over one billion dollars of customers and corporate funds with Crypto Capital (New York County Clerk, 2019). Bitfinex reported that its partnership with Crypto Capital operated generally well until April 2018, when news emerged that funds at Crypto Capital had been seized by authorities in Poland due to money-laundering investigation (Zhao, 2019). As a consequence, Crypto Capital could not process withdrawals, leading to Bitfinex and CEX.IO having severe difficulties with customers' request to withdraw the fiat from their exchange's accounts. Delays for several weeks and months have been reported across multiple forums and exchange's websites (Morris, 2019). According to New York County Clerk (2019), in August 2018 Crypto Capital informed Bitfinex that in total, $851 million of Bitfinex's funds could not be accessed, because they were seized by governments of Portugal, Poland, and the United States.[6] In order to fill the gap, Bitfinex used reserves tied to Tether, a stablecoin pegged to the US dollar, the New York Attorney General said (James, 2019).

Delays with withdrawals of fiat currency could lead to price discrepancies on Bitfinex and CEX.IO, because customers, who wished to come out of these exchanges, were willing to pay a premium for bitcoins, which were easier to withdraw than US dollars. Bitfinex has already experienced a similar situation in March 2017 - after Bitfinex and Tether lost their banking partnership with Wells Fargo, an American multinational financial services company, the bitcoin premium soared above 8% (CryptoCompare, 2019). Price differences of the same nature were observed for a lengthy period shortly before Mt.Gox announced bankruptcy in 2014. According to Nasdaq (2015), the Department of Homeland Security seized Mt.Gox funds that were held by Dwolla, a Des Moines-based money transfer service, leading to slow processing of requests from customers, who wanted to withdraw money from their accounts on Mt.Gox. The prices on the exchange were pushed upwards by 12% due to high demand for bitcoins – the customers converted their fiat to bitcoin to withdraw via more reliable Bitcoin blockchain.

In the sample period, Bitfinex features the highest bid and ask quotes among all exchanges in the first week of May, and also tends to exhibit higher bid prices in July and August. Just a few days later, CEX.IO started to trade bitcoin at a premium and continued quoting higher prices until July. As in the previous examples above, a possible explanation for the premiums

[6] The scheme allegedly involved running a "shadow bank that processed hundreds of millions of dollars of unregulated transactions on behalf of numerous cryptocurrency exchanges", according to an announcement by the U.S. Attorney's Office (Cheng, 2020). Two managers of Crypto Capital were charged with bank fraud.

on Bitfinex and CEX.IO could be the news about the seizure of Crypto Capital funds in April and August 2018, and the related fiat withdrawal delays. Whereas the maximum premium on Bitfinex could reach 1% and persisted for a few days, the premium on CEX.IO reached 5% and persisted at least for two months. The bitcoin prices on CEX.IO could react more strongly on fiat withdrawal delays and related buying pressure than on Bitfinex, as CEX.IO is less liquid than Bitfinex (see Table 2), and CEX.IO could not use the Tether reserves as Bitfinex to restore fiat liquidity.

In contrast to Bitfinex and CEX.IO, Gemini and Coinbase Pro are known for much deeper, well-regulated banking relationships (Casey, 2019) and for the highest level of KYC compliance in the cryptocurrency market (Golstein, 2018). US dollar fiat funds submitted at Coinbase Pro and Gemini are held in anonymous FDIC[7]-insured banks located in the US, or in US treasuries (Zach, 2018). Also Kraken established a partnership with Fidor Bank, which is a German online bank, regulated by the German Financial Supervisory Agency BaFin (KRAKENFX, 2013).

Lack of access to the financial system forces cryptocurrency exchanges to use less reliable sources of fiat liquidity, leading to higher counterparty risk of their customers and frictions such as fiat withdrawal delays and restrictions. The evidences above show that price differences between cryptocurrency exchanges can arise due to fiat illiquidity. As a constraint to arbitrage, fiat illiquidity may lead to persistent price differences, since it deters arbitrageurs from exploiting arbitrage opportunities between exchanges. In summary, two examples in this section 5.1 and 5.2 describe demand shocks, stemming from institutional frictions, that had significant and long-term price effects. A natural question is why arbitrageurs can not get this inefficiency priced out of the market. The next section discusses the constraints to arbitrage, which may explain the persistence of arbitrage opportunities on the Bitcoin market.

5.3 Limits to Arbitrage

Bitcoin price deviations across exchanges suggest constraints to arbitrage that limit its effectiveness in achieving market efficiency (Shleifer, 2000; Barberis and Thaler, 2003). According to the law of one price, the price of bitcoin should be identical regardless of where it is purchased. Bitcoin is a homogeneous and mutually interchangeable asset – one bitcoin can be replaced by another one. The ownership of bitcoin can be transferred at a cost (miners' and exchange fees) and in some period of time. If price differences occur, at least theoretically they can be exploited

[7]The Federal Deposit Insurance Corporation (FDIC) is an independent agency created by the United States Congress to maintain stability and public confidence in the nation's financial system (FDIC, 2020).

by arbitrageurs to a profit. The Law of One Price is also enforced by the fact that the arbitrage possibilities are known. The bitcoin price differences can be directly compared based on exchanges' order books and market depth.

Referring to the literature survey by Gromb and Vayanos (2010), the limits to arbitrage may arise as transaction (Roll et al., 2007) and holding costs (Pontiff, 2006), fundamental and nonfundamental risks, capital constraints (Shleifer and Vishny, 1997) and short sale constraints (Ofek et al., 2004). Hautsch et al. (2020) shows that a new type of market friction - the blockchain-related settlement latency imposes limits to arbitrage by exposing cross-market arbitrageurs to the risk of adverse price movements, in a situation where inventory holdings and margins are exhausted, and the arbitrageur is forced to physically transfer the asset between markets. In distributed ledger technologies, such as Bitcoin blockchain, it is the time until the transfer of bitcoins across exchanges is verified by the Bitcoin network. The transfer of bitcoins is not accomplished unless the transaction is not verified. Three exchange characteristics might help arbitrageurs to avoid the exposure to the blockchain-related settlement latency - the ability to open a business account, to initiate short sales and to trade at margin. In this way, some arbitrage constraints depend on the arbitrage strategies, which were already discussed in detail in Section 4.1.

5.3.1 Transaction Costs: Fees and Bid-Ask Spread

Transactions costs limit arbitrage regardless of the strategy. They are incurred when transaction occurs, and include all trading fees, commissions, bid-ask spread and market impact (the impact a market participant's own trading has on the price) (Pontiff, 2006). On cryptocurrency exchanges, the magnitude of the trading fees to prevent arbitrageurs from implementing their strategies would depend on the type of a client, the transaction size and the exchange's fee schedule, it is therefore not clear a priori to what extent they contribute to limiting arbitrage in the Bitcoin market. Large players are less sensitive to trading fees since developed exchanges provide them with business accounts with preferential trading conditions including fees that are far below the fees for retail investors.

The fees arising from the Bitcoin network may in turn contribute to the bitcoin price deviations. When a trader withdraws bitcoins from an exchange, the exchange will need to make an on-chain transaction (i.e. the transaction has to be recorded on the blockchain) and pay transaction fees to the miners. The miner's fees are expressed and paid in bitcoins. Exchanges set a bitcoin withdrawal fee themselves, looking at the average bitcoin transaction fees and the network load. When many people are transacting, the miners give priority to the transactions with higher fees, and exchanges may increase the transaction fees to get faster the confirmation from the blockchain network. In this respect, the arbitrage constraints can be associated not only with cryptocurrency exchange characteristics, but the characteristics of bitcoin and bitcoin's blockchain.

As in traditional markets, price differences in the Bitcoin market could be related to the market frictions, associated with transacting (Roll et al., 2007). An endogenous market friction, liquidity, is pertinent to the arbitrageur. If the markets are illiquid, and the price difference is sufficiently wide, the illiquidity may have a lasting effect on price differences as arbitrageurs may choose to spread their trades out over time to decrease price impact (Kyle, 1985). Therefore, the large volume of arbitrage trades, needed to close the price difference, may not happen immediately (Roll et al., 2007). Also, thin order books are not sustainable, in particular, for institutional and large professional traders, leading to frictions that delay trades and impede traders from exploiting price differences.

An indicator of (il)liquidity – *bid-ask spread* is a market microstructure characteristic and a transaction cost along with trading fees, but determined by market participants. Specifically, bid-ask spread measures the cost dimension of liquidity, since it covers the cost of waiting (persons willing to wait before selling), carrying inventory costs, and adverse information costs (Demsetz, 1968; Stoll, 1978b; Glosten and Milgrom, 1985). In this respect, bid-ask spread is a proxy for liquidity risk, which is defined as a risk of failing to liquidate a position at a reasonable price without a time delay. An alternative measure of liquidity is a market depth, calculated as the average quoted depth, defined in equation 9. In the context of arbitrage, a trader needs a sufficient amount of orders near the inside quotes to ensure that the price difference can cover the trading fees. Also, sufficient order book depth makes it worthwhile for the arbitrageur to profit without incurring a large and adverse price impact and to make bigger trades for the right price without a trading time delay. Torre (1997) and Gatheral (2010) have shown that market impact increases proportionally with the square root or power 3/5 of the executed size (expressed as a proportion of the available liquidity).

Pontiff (2006) suggested that besides the transaction costs there are holding costs that can impede arbitrage. Holding costs are incurred every period that a position remains open. Holding costs include the opportunity costs of capital, the opportunity costs resulting from one's inability to collect full interest on proceeds from a short-sale (short-selling costs, see Ofek et al. (2004)) and idiosyncratic risk exposure. Bitcoin's idiosyncratic risk is related to price uncertainty of bitcoin and unrelated to the price risk of other cryptocurrencies. Shleifer and Vishny (1997) claim that idiosyncratic volatility matters to specialized arbitrageurs, since it cannot be hedged and arbitrageurs are not diversified. One of the proxy measures of the price uncertainty is the bitcoin price volatility.

5.3.2 Fiat Illiquidity as a Limit to Arbitrage

In the Bitcoin market, fiat illiquidity limits arbitrage along with bitcoin illiquidity. First, according to Gromb and Vayanos (2010), a nonfundamental risk arises from demand shocks when these affect prices. In the previous section, fiat illiquidity was discussed as a source of price differences, when the value of bitcoin increases artificially on the cryptocurrency exchange due to high demand. Second, the arbitrageur incurs holding costs, since the position remains open during the whole period of the withdrawal delay (Pontiff, 2006). Third, fiat illiquidity is associated with counterparty risk – the company is unable to process withdrawals in fiat currencies, and the funds may never be returned to its customers. Fiat illiquidity may force the arbitrageur to withdraw funds in bitcoin via the more reliable bitcoin blockchain, but then the arbitrageur is exposed to the price uncertainty, and the expected arbitrage profit also remains uncertain.

Another limit to arbitrage associated with fiat-to-crypto exchanges is a bank transfer used to deposit and withdraw fiat currency. In this study, although some cryptocurrency exchanges are located in different countries, all of them trade BTC/USD currency pair, and deposits and withdrawals are possible in USD. Some exchanges offer trading BTC/USD currency pair, withdrawal and deposit in USD only for the US residents. For example, until July 2, 2021, Coinbase Pro offered wallets in USD only for the US customers residing in a state, where Coinbase Pro offers service. Now institutional traders can make withdrawals and deposits in USD, EUR or GBR, and access to the related trading pairs regardless of the residency of the customer (Coinbase, 2021). BTC/USD is traded separately from BTC/EUR on Bitstamp, Coinbase Pro and Kraken, and from BTC/JPY on bitFlyer. Also, this study is focused only on the BTC/USD currency pair, other fiat currencies are not considered for arbitrage. There is therefore no need to exchange fiat currencies and the international bank transfer can be skipped. However, the arbitrageur still has to use bank transfers to deposit and withdraw USD to re-balance or to complete a round trip. This applies to both institutional and retail traders, beside that they both use trading at margin, which is offered by few exchanges. Bank transfers may take 1-2 days, meaning that the arbitrageur can re-balance or complete a round trip at most once a day.

Arbitrage constraints could create impediments that prevent arbitrageurs from fully taking advantage of the market inefficiency. This enables price deviations across exchanges to continue since arbitrageurs are deterred from exerting price pressure that eliminates price differences. In the next sections, we examine how the price differences across exchanges relate to such arbitrage frictions as illiquidity, transaction costs and price uncertainty.

6 Bitcoin Price Differences and Liquidity

In Section 5 the conditions, under which bitcoin may diverge and persist on cryptocurrency exchanges, are discussed in detail. In this section, it is attempted to more formally identify the dynamic relationship between bitcoin price differences and liquidity on market level for each exchange. Whereas, the paper of Hautsch et al. (2020) focused on the blockchain-related settlement latency as an important market friction in the Bitcoin market and showed that bitcoin price differences are particularly large during times of high latency implied price risk, this study focuses on another market friction - liquidity. The motivation for this study derives in part from the earlier work of Roll et al. (2007), who investigated the relation of price deviations from no-arbitrage values to the frictions associated with transacting. Specifically, Roll et al. (2007) examined the dynamic relation between the futures-cash basis and stock market liquidity, measured by quoted spreads. The authors provide the evidence of two-way Granger causality between the short-term absolute basis and liquidity. The link between price differences and liquidity could be contemporaneous as well as long term. Liquid market should facilitate arbitrage trades, making relative pricing in the market more efficient, and, when the market is illiquid, arbitrageurs may have difficulty closing the gap between two exchanges, especially when a price difference is sufficiently large (Roll et al., 2007). According to Kyle (1985), arbitrageurs may be not able to immediately provide the large volume of trades, required to close the price difference, as they may choose to spread their arbitrage trades out over time to avoid a large price impact. In the reverse direction, Roll et al. (2007) argue that wide price differences may trigger arbitrage trades, which in turn could create order imbalances, therefore, price differences could forecast future liquidity. However, Roll et al. (2007) do not specifically show that arbitrage trades induce order imbalances and hence, explain the positive effect of price differences on illiquidity (see also Foucault et al. (2017)). Regardless of that, if price differences indicate price pressures (transient demand or supply shocks), an increase in both price differences and spread should be driven by the inventory imbalances of liquidity providers (see also Rösch (2021)). Empirically, Comerton-Forde et al. (2010) show that spreads widen when inventory imbalances are high. Theoretically, the foundational work in microstructure by Stoll (1978a) and O'Hara and Oldfield (1986) suggests that price pressures induced by order imbalances can have a lasting impact on liquidity. Price pressures occur since market makers, faced by the inventory problems caused by order imbalances, are expected to respond by revising price quotations and changing bid-ask spread (Chordia et al., 2002). Inventories affect spread if market makers face capital constraints (Amihud and Mendelson, 1980), holding costs (Shen and Starr, 2002) or averse to risk (O'Hara and Oldfield, 1986). Financing constraints increase marginal costs of providing liquidity in the direction of imbalances, causing a widening of bid-ask spread (Rösch, 2021). Also price pressures caused by imbalances have a direct effect on returns, increased fluctuations of which cause a widening of bid-ask spread due to the increased inventory risk (Chordia et al., 2002). In this way, if market orders tend to occur, for example, on the sell side of the market, the liquidity providers will revise their price quotations (bids) on the buy side of the limit order book, since they are faced with inventory problem. To elicit trading on the buy side of the market, the liquidity providers on the sell side of the limit order book are expected to adjust their quotes (ask), i.e. worsen their offered terms of trade. If the quotes are not adjusted quickly enough, the bid-ask spread will widen. Hence, given that price differences indicate price pressures, the contemporaneous and long term positive association between price differences and bid-ask spread, is due to persistent market makers' inventory problems, caused by order imbalances, and also their inability to adjust price quotations quickly enough during periods of large imbalances (Chordia et al., 2002). Chordia et al. (2002) also shows that the arguments above are empirically upheld by market-wide order imbalances, returns and bid-ask spread for NYSE stocks.

Roll et al. (2007) showed that index futures basis have a lasting and significantly positive effect on stock market illiquidity, and argued that arbitrage trades, in response to wide basis, induce order imbalances which in turn strains liquidity. However, as already stated above, Roll et al. (2007), do not specifically show that arbitrage activities explain this relationship. Rösch (2021) proxies arbitrage activity by the daily maximum price deviations as an inverse measure, and finds that arbitrage activity is positively associated to liquidity, given that the author also shows that price deviations in the ADR market from 2001 to 2016 mainly arise because of price pressure. According to Gromb and Vayanos (2010) and Copeland (2019), whether arbitrageurs provide liquidity by exploiting mispricings depends on the cause of arbitrage opportunities. When arbitrage opportunities arise due to price pressure (demand or supply shocks), arbitrageurs improve liquidity by trading against price pressure and decreasing inventory holding costs (Holden, 1995; Gromb and Vayanos, 2010). If arbitrage opportunities are due to asymmetric private valuation shocks ("toxic" arbitrage opportunities, see Foucault et al. (2017)), then market makers, exposed to the risk of trading with arbitrageurs at stale quotes, charge wider bid-ask spreads to cover adverse selection risk (Copeland and Galai, 1983). Foucault et al. (2017) theoretically and empirically showed that arbitrageurs' response to the "toxic" arbitrage opportunities impairs liquidity by raising adverse selection costs for liquidity providers.

Price differences in the Bitcoin market may also occur due to both: price pressures (non-fundamental demand and supply shocks, see Sections 5.1 and 5.2) and due to asynchronous information revelation (Kotz et al., 2012; Voigt, 2020) and asymmetric private valuation shocks (Foucault et al., 2017). The investigation of the ratio of the toxic bitcoin price differences to the price differences that reflect price pressure is beyond the scope of this thesis. Previous literature[8] shows a strong positive relation between order flow and prices in traditional markets: typically order flow imbalances explain 15%-30% of the daily variations of stock returns. Makarov and Schoar (2020) find that this relation is significantly higher for the Bitcoin market: up to 85%, although in cryptocurrency market it is less obvious whether there are any traders who have more information than others. By the preceding arguments, it is expected that the relation between bitcoin price differences and illiquidity is positive in the Bitcoin market, regardless of the reason for the arbitrage opportunity to arise. If price deviations reflect price pressure, then they are themselves a measure of illiquidity and an inversely measure of arbitrage activity (Rösch, 2021). Both price differences and spreads are driven then by inventory imbalances of market makers. If arbitrage opportunities are "toxic", liquidity providers charge larger spreads to cover the risk of trading at stale quotes (Copeland and Galai, 1983; Foucault et al., 2017).

Following these ideas, the goal of this section is to explore them for the cryptocurrency markets, i.e. to examine the joint dynamic interactions between bitcoin price differences and liquidity and how they vary across exchanges. In the next section time series properties of the bitcoin price differences and bid-ask spreads will be investigated for each exchange.

6.1 Time series properties of Price Differences and Bid-Ask Spreads

Prior to conduction of vector autoregressions, following Roll et al. (2007), the common regularities and trends from the time series will be investigated and removed to avoid spurious results. The minute-by-minute quote data is used to construct an hour-by-hour and a day-by-day measure of the bitcoin price differences for each exchange i. One concern is that the price differences can be mechanically linked to spread. To mitigate this effect, the price differences were computed as

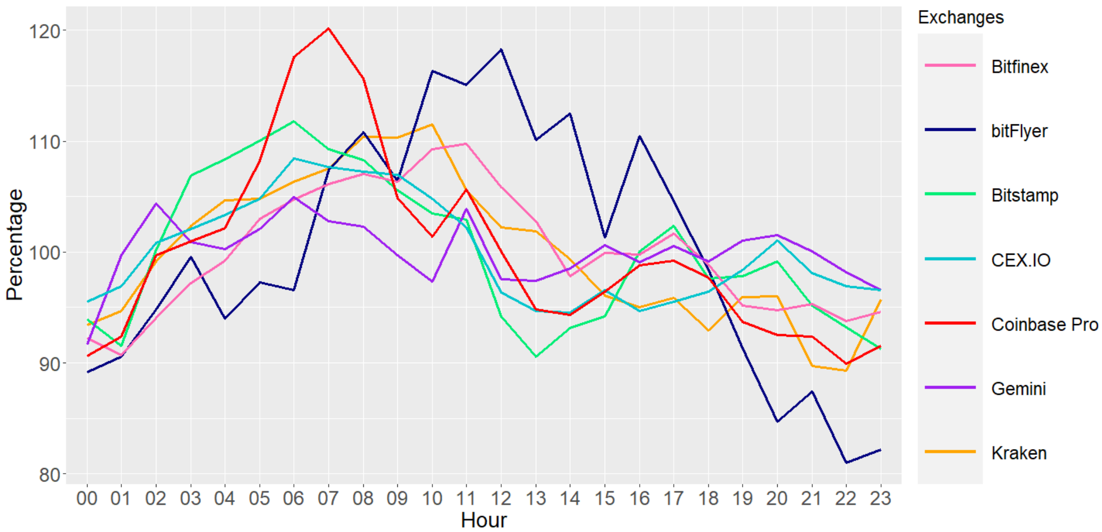
[8] For example, Evans and Lyons (2002), Berger et al. (2008), Fourel et al. (2015) for foreign exchange markets, Brandt and Kavajecz (2004) for the US Treasury markets, Deuskar and Johnson (2011) for S&P 500 futures market, Chordia et al. (2002), Goyenko et al. (2009), Hendershott and Menkveld (2014) for NYSE stocks

deviations from midquote prices. Specifically, the hourly price difference of exchange i at time t is defined as the cross-sectional hourly average of differences between the price level, calculated as a midquote (see Equation 6), on each exchange and the midquote prices on all other exchanges. The hourly price difference is denoted by PD_{it} and is defined in Equation 14 in Section 3.3:

$$PD_{it}^{M} := \frac{|m_{it} - \overline{m}_t|}{\overline{m}_t}, \tag{49}$$

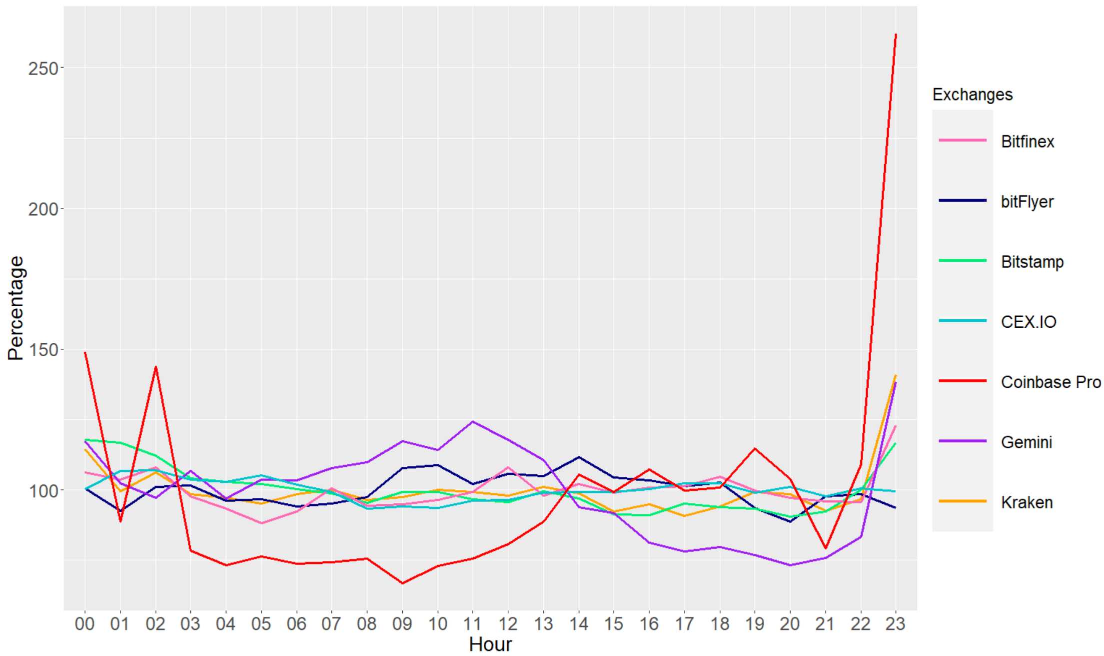
In this way, PD_{it}^{M} measures the deviation from the average midquote price across exchanges. The daily price differences are constructed analogously. Since deviations can arise in either direction, the absolute value of price differences is taken as a more sensible measure to analyze. The liquidity is measured by a quoted bid-ask spread, defined in Equation 7 in Section 3.2. Since the price differences are scaled by the bitcoin price, bid-ask spreads are not scaled by the price to avoid attributing movements in bitcoin prices to movements in liquidity. To ensure that the results are not driven by outliers, price deviation and illiquidity measures are winsorized at the 1% and 99% level on each day.

Figure 10. Intraday Dynamics of Price Differences



Notes: This figure displays intraday variation of the bitcoin price differences for each exchange in the sample. First, a time-series of minute-by-minute percentages of absolute price differences is created. Then, the percentage price differences are averaged within each hour to create a time-series of hour-by-hour price differences. The percentage price differences are averaged on a hourly basis, then scaled by the variable daily average.

Figure 11. Intraday Dynamics of Bid-Ask Spread



Notes: This figure displays intraday variation of the quoted bid-ask spread for each exchange in the sample. First, a time-series of minute-by-minute quoted bid-ask spreads is created. Then, the bid-ask spread is averaged within each hour to create a time-series of hour-by-hour bid-ask spreads. The bid-ask spreads are averaged on a hourly basis, then scaled by the variable daily average.

The intraday variations of the price differences and spreads presented in Figures 10 and 11 provide information about trading behavior and regularities throughout the day. As cryptocurrency exchanges do not have opening and closing times, the typical U-shape of spread, commonly found in equity markets (Mcinish and R.A., 1992; Madhavan et al., 1997), is not as pronounced in the Bitcoin market and the whole pattern is not clear. This observation could be explained by an international market participation, which leads to an overlay of intraday patterns from different time zones and the existence of automated trading. Although this observation is generally in line with Dyhrberg et al. (2018), a small regularity can be still observed. The spreads are slightly wider at the start and end of the trading day, whilst the spreads are tighter in the middle of the day for the most of the exchanges in the sample. This might reflect lower market activity at the beginning and end of the day. In contrast to the spread, the bitcoin price differences' daily pattern is more clear: it increases overnight to peak early in the morning, between 6:00

and 10:00 CET, and then declines throughout the day. The observed pattern indicates that a large number of trades originate in Europe during normal trading hours, even on the US cryptocurrency exchanges such as Coinbase Pro, Gemini and Kraken. The Japanese cryptocurrency exchange bitFlyer exhibits a slightly different pattern: the price differences begin to decline after 13:00 CET, which corresponds to 6:00 GMT, indicating that trades for the BTC/USD pair on bitFlyer originate mostly in the United States.

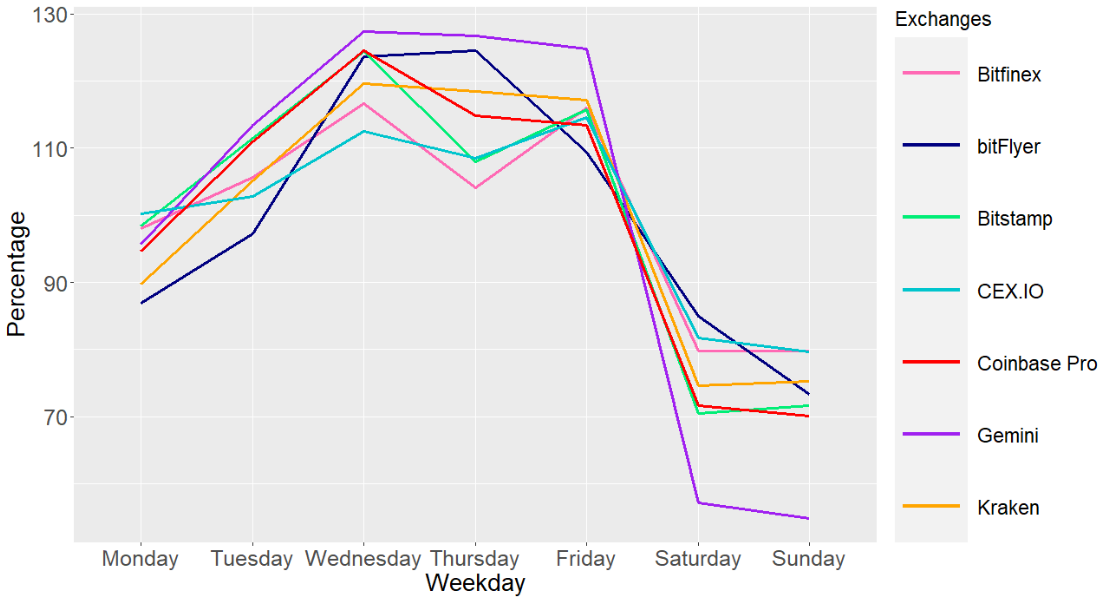
By the preceding arguments, the following variables were used to adjust the hourly price differences and spreads: (i) hour of the day to account for the intensity of trading activity, (ii) a linear and squared time-trend to remove any long-term trends, since over the time period, the price differences have become more integrated, especially in the second half of the sample period, i.e. in July and August, (iii) day of the week to account for decreasing trading activity in the end of the week, (iv) 7 calendar month dummies for March through September. In Section 6.2, innovations (residuals) from the adjustment regressions are used in vector autoregressions. For spreads explanatory variables also include a dummy to account for the shift in minimum tick size from 1 to 0.1 for Bitfinex (June 15, 2018), and from 0.001 to 0.01 for Coinbase Pro (May 23, 2018). A reduction in spreads is expected on Bitfinex after the minimum tick size was reduced, an increase - on Coinbase Pro.

Table 9 in the Appendix presents coefficients from the adjustment regressions for the hourly price differences. The reference group includes 24:00 as an hour of the day, Tuesday as a day of the week and April as a calendar month. Table 9 reveals that bitcoin price differences are higher between 04:00 and 11:00 CET (excluding the Japanese exchange bitFlyer), which is in line with Figure 10 and reflect the impact of the lower market activity on price differences at the nighttime. The bitcoin price differences are higher on Sunday throughout Wednesday, but lower on Thursday, Friday and Saturday relative to Tuesday. Figure 12 illustrates that the trading volume for BTC/USD currency pair starts to rise on Monday, achieves the highest values on Wednesday, Thursday and Friday, and finally decreases on Saturday till Sunday. Although cryptocurrency exchanges are opened 7 days a week, the trading activity is decreasing over the weekend and this pattern is pronounced for each exchange. The trading activity is also reduced due to the risks holding positions over the weekends. In line with Roll et al. (2007), higher trading volume on Wednesday could result in higher price pressures, and consequently higher bitcoin price differences. This phenomenon was also observed by Roll et al. (2007) in a case of the futures-cash basis. After Wednesday the bitcoin price differences begin to decrease and achieve the lowest values on Saturday. On Sunday the price differences start to increase and are at their highest, but the trading volume is the lowest among all weekdays, implying that the absence of large volume required to close the price gap may result in wider bitcoin price differences. It can be also seen that the price differences are higher in March, May and June, but lower in July, August and September relative to April. In addition to the previous observation, the bitcoin price differences exhibit a downward trend, confirming the findings in Section 3.3, where Figure 2 illustrates a slight trend of decreasing price differences.

Table 10 in the Appendix presents coefficients from the adjustment regressions for the hourly quoted spread. The table reveals that the quoted spreads are higher on Friday through Sunday relative to Tuesday (the omitted weekday) and are also lower in all months except March relative to April (the omitted month). The effect of month is increasing with time, indicating increasing liquidity of the exchanges. Moreover, spreads exhibits a downward trend. There are no specific intraday patterns of spread, which is consistent with Figure 11. Although the signs of the tick size effect are as expected, there is no significant reduction in spreads after the minimum tick size was reduced from 1 to 0.1 on Bitfinex, and no significant increase after the minimum tick size was increased from 0.001 to 0.01 on Coinbase Pro.

Since the daily price differences and spread were computed by averaging the hourly time series, it is expected that the daily-horizon price differences and spread at a daily horizon have the same time-trends and calendar regularities as the hourly-horizon series. Therefore, the daily bitcoin price differences and spread were adjusted using the following variables: (i) a linear and squared time-trend to remove any long-term trends, since over the time period, the price differences have become more integrated, especially in the second half of the sample period, i.e. in July and August, (ii) day of the week to account for decreasing trading activity in the end of the week, (iii) 7 calendar month dummies for March through September. Residuals from the adjustment regressions are used in a vector regressions in the next section.

Figure 12. Weekday Dynamics of Trading Volume



Notes: This figure displays weekday variation of the daily volume (BTC) for BTC/USD currency pair for each exchange in the sample. The volume is averaged on a daily basis, then scaled by the variable daily average.

To investigate the time series characteristics of the bitcoin price differences the method of Box and Jenkins (1970) is used. The inspection of Figures 26 27 for the hourly series, and Figures 26, 27 for the daily series, reveals that the bitcoin price differences are persistent. The ARIMA model is applied to find the best fit of a time-series model to past values of the time series. The ARIMA model expresses time series using appropriate differencing and an ARMA model, which includes moving an average (MA) process and autoregressive (AR) model. The ARIMA (p, d, q) is expressed as

$$\Phi(B)\Delta^d y_t = \Theta(B)\epsilon_t \qquad (50)$$

Where B is the backward shift operator

$\Delta^d = (1 - B)^d$

$\Theta(B) = 1 - \theta_1 B - \ldots - \theta_q B^q$ is a moving average polynomial of order q,

$\Phi(B) = 1 - \phi_1 B - \ldots - \phi_p B^p$ is an autoregressive polynomial of order p.

Following the Box and Jenkins methodology, the price differences need to be stationary to use equation 50 for modelling the series. The augmented Dickey-Fuller test strongly rejects the existence of a unit root for the adjusted and not adjusted price differences, quoted spreads, quoted depth and volatility. Tables 7 and 8 in Appendix shows that the Augmented Dickey-Fuller and Phillips-Perron tests rejected null hypothesis of the unit root at the 1% significant level for all exchanges. The lag length is chosen by Akaike's Information Criteria (AIC) individually for each exchange and each variable. In order to ensure that sufficient lags are included in the testing regressions such that the error term becomes white noise, the autocorrelation of the residual series is studied by inspecting autocorrelation and partial autocorrelation functions. The results are not presented for brevity, but show that no serial correlation in the residuals of the testing regressions.

In order to identify the ARIMA structure of the price differences, information criteria and Box and Jenkins' methodology of pattern recognition have been applied. The lag order p and q is determined by comparing a set of estimated models by information criteria. Tables 11 and 12 in Appendix shows the best fitted ARIMA models based on the AIC measure for the hourly and daily series respectively. By visually inspecting correlograms and partial correlograms in Figures 26, 27 and 28, 29 in Appendix, the chosen lag orders have been ascertained. For the hourly series, according to AIC, the best ARIMA models are AR(2) for Coinbase Pro, AR(4) - for bitFlyer and Kraken, AR(5) - for Bitfinex, and AR(6) - for Bitstamp, CEX.IO and Gemini. For the daily bitcoin price differences, the best models according to AIC are AR(1) and AR(2), and only one exchange (Bitstamp) exhibits price differences that follow ARMA(1,1) process, meaning that the price differences are influenced not only by the past values, but also by recent shocks or random errors to the series. The autoregressive models explain a large fraction of the time-series variability of bitcoin price differences for most exchanges. Specifically, for the hourly series, the R^2 ranges from 57% to 94%, and the first autoregressive coefficient - from 0.59 to 1.05. For the daily price differences, the R^2 is between 0.25% and 0.61%, and the first autoregressive coefficient - between 0.30 and 0.81. These results indicate that bitcoin price differences are persistent: for the hourly series, 4-6 previous hours contribute to the output, for the daily series - 1-2 days.

According to daily ACF and PACF 28, 29 in Appendix, all exchanges except bitFlyer exhibit similar statistical behavior. Compared to other Bitcoin platforms, bitFlyer has weaker positive autocorrelation and stronger negative autocorrelation in price differences. The weak positive autocorrelation indicates that price differences on bitFlyer have more fluctuations, whereas the strong negative autocorrelation, according to Shi et al. (2019), "is brought about by the market prevention that protects Bitcoin price from deviating actual value, implying that there are more violent fluctuations in bitFlyer." By analyzing the properties of price returns in five leading Bitcoin platforms, Shi et al. (2019) detect that the kurtosis of the price return distribution, the power-law exponent of the price return distribution, and the autocorrelation of price return in bitFlyer all highly deviate from other exchanges. After considering the best prices that traders quote, Shi et al. (2019) found that that "buyers and sellers quote price far beyond current best price (higher than best bid or lower than best ask) at the same time" (Shi et al., 2019). The authors argue that some traders tried to "manipulate the price by creating a false impression of an active market. If the abnormal ask orders and the abnormal bid orders are placed concurrently again and again during a certain time period, it may not be coincidental but deliberate. Thus, it may be potentially linked to either price manipulation or money laundering. Furthermore, bitFlyer was punished by the Japan Financial Services Agency due to the Know Your Customer (KYC) policy vulnerability on June 22, 2018" (Shi et al., 2019). For these reasons, to avoid potential misinterpretation of the estimation results, bitFlyer is excluded from the VAR analysis in this thesis. Moreover, bitFlyer is the smallest fiat-to-crypto exchange by BTC/USD trading volume with the smallest average quoted depth and the largest time-weighted bid-ask spread, suggesting less trading activity in BTC/USD currency pair relative to other fiat-to-crypto markets.

6.2 Vector Autoregressions and Robustness Checks

By the preceding arguments, discussed in the beginning of Section 6, liquidity and price differences are likely determined jointly, and both variables may have a contemporaneous as well as a persistent impact on each other. For each exchange, the liquidity and price differences dynamics is modeled by the method of vector autoregression, using data that spans more than 3,000 hours or 159 days. Considering the heterogeneity of the sample exchanges in terms of liquidity and the magnitude and persistence of price differences, it is expected that the impact of liquidity on price difference and vice versa may differ between the markets. Controlling for fluctuations in liquidity driven by volatility dynamics is performed by including exchange-specific spot volatility in the VAR system. Hourly and daily spot volatility is computed by converting minute-level spot volatility estimates to the hourly and daily level respectively. The minute-level spot volatility is estimated following Kristensen (2010) (see Equation 12), and converted to the hourly and daily level by multiplying it by $\sqrt{60}$ and $\sqrt{1440}$ respectively, which is the square root of the number of minutes in an hour and in a day. According to Stoll (1978a), increased volatility tends to decrease liquidity by increasing inventory risk. In the reverse direction, low liquidity could increase price volatility (see, Longin (1997), or Subrahmanyam (1994)) or liquidity can affect volatility by attracting more trading (see Amihud and Mendelson (1986)). As for price differences, volatility creates uncertainty about the duration of arbitrage opportunity, suggesting that higher volatility is associated with higher price differences. In the reverse direction, bitcoin price differences should positively affect volatility, since price differences can be associated with higher order imbalances, which in turn increase volatility (Chordia et al., 2002).

Following Roll et al. (2007), to study robustness to alternative measures of liquidity, average depth in used instead of the quoted bid-ask spread in the VAR system. Average depth is positively associated with liquidity, and is defined as an average quantity that the trader can trade at the best prices (see Equation 9 in Section 3.2). Then, the depth measure is averaged over the hour and over the day to build hourly and daily series, respectively.

The next robustness check involves the effects of the Blockchain network activity on the bitcoin price differences. In order to capture the effects of the Bitcoin network activity, number of transactions waiting for verification and the network fee per byte is included in the model as exogenous variables. The number of waiting transactions is relevant for cross-market arbitrageurs, for which inventory holdings and margins are exhausted, and they are forced to use the Bitcoin blockchain to physically transfer funds between exchanges. On average every 10 minutes, the Bitcoin blockchain bundles together up to 1 MB (on average 1000 transactions) volume of transaction data that wait for verification into a new block. An increasing number of transactions waiting for verification implies a decreasing probability of being included in the next block. Also Easley et al. (2019) and Biais et al. (2019) showed that the number of waiting transactions increases the latency of all transactions waiting for confirmation from the Blockchain network. According to Hautsch et al. (2020), settlement latency implies arbitrage costs for cross-market arbitrageurs, because they are exposed to non-hedgeable price risk. In this respect, more waiting transactions are associated with higher network activity, which increases settlement latency. Therefore, the number of waiting transactions is expected to increase the price risks of cross-market arbitrageurs, potentially leading to higher price differences between exchanges. The number of waiting transactions is included as a exogenous variable in the VAR system since price differences and bid-ask spread are not expected to affect network activity and subsequently settlement latency. Voigt (2020) examined a novel dataset that allows to identify potential arbitrage transactions and found that cross-market trades with the purpose of exploiting price differences represent a minor fraction of all transaction waiting for verification. In this way, Voigt (2020) illustrated that the presence of price differences do not induce inflated network activity. Another reason to control for the network activity in the VAR system is that an increase in number of transactions waiting for verification might narrow spreads (Voigt, 2020).

Whereas an increase in network activity is expected to increase settlement latency, paying higher fees is expected to reduce it. Traders have the option to set higher fees to make it attractive for miners to include their transactions in the next block (Easley et al., 2019). There are time series of fees per byte available for this study that was gathered by running a full node (computer that connects to the Bitcoin network is referred to as a node) and collecting transaction-specific information in real time. Fee per byte (in BTC) is the total fee per transaction divided by the size of the transaction in bytes. Each transaction contains the relevant information for this study: a unique identifier of the transaction, a timestamp of the initial announcement to the network and the fee per byte the initiator of the transaction offers validators to confirm the transaction. The fees per byte are aggregated to the minute, hourly and daily level by taking the average of the fees of all transactions entered memory pool in a given minute (Hautsch et al., 2020).

Using an hourly and daily interval to analyze the joint dynamics of liquidity and the bitcoin price differences has several justifications: one of them is that trading and arbitrage allow high frequency trading which enables fast reaction to price changes and to an occurrence of arbitrage opportunities. Also, as stated by Roll et al. (2007), arbitrageur's and market makers inventory considerations are important over rather short horizons, hourly and daily as opposed to weekly or monthly. The second justification for the hourly and daily horizon is also given by Roll et al. (2007), and involves the impact of asynchronous trading on the price differences. This justification could also be applied to the case of bitcoin price differences, since different cryptocurrency exchanges trade bitcoin at different frequencies due to the market segmentation. Some exchanges are more advanced and host more institutional traders, some primarily serve retail users of bitcoin, as inferred from differential trade sizes in Table 2. Table 2 shows that the average quoted depth (in BTC) is relatively large on Coinbase Pro and Bitfinex, suggesting a substantial amount of traders and the type of traders that generate large limit orders. On the other hand, CEX.IO and Kraken are small-capitalization exchanges according to the average quoted depth and the trading volume for BTC/USD currency pair. Also, different cryptocurrency exchanges are located in different countries and different time zones (see Table 10), and fiat currency can not flow seamlessly across them, which causes asynchronous price information arrival. Market fragmentation is also created in restrictions for some traders and exchange governance rules, for example, in August 15, 2018 Bitfinex has prohibited all US citizens and US corporations from using it's services (Bitfinex, 2018). Due to markets fragmentation, information about bitcoin may arrive asynchronously on cryptocurrency exchanges, leading to cross-market price differences. A potential concern is that larger bitcoin price differences may be accompanied by high asynchronous trading and high spreads (Roll et al., 2007). If a substantial part of arbitrage opportunities (which are first calculated at the minutely level) is "toxic", then the hourly and daily price differences will miss most of the pattern, since, in the sample period, bitcoin price differences persist between 7-15 minutes. Also Foucault et al. (2017) shows that most of the "toxic" arbitrage opportunities are short-lived, because arbitrageurs react faster to such opportunities and often equipped with advanced trading software systems. If the majority of price differences are due to price pressures, an hourly, and especially a daily horizon should reduce the impact of asynchronous trading (Chordia et al., 2002; Roll et al., 2007).

6.2.1 Model Specification

As it was previously discussed, bitcoin price differences, liquidity and bitcoin price volatility are likely determined jointly, giving rise to a potential endogeneity concern of simultaneity. Thus, the empirical analysis is motivated to employ a three equations vector autoregressive model to uncover the impact of liquidity on the bitcoin price differences. Define $Y_t = (PD_{it}, QS_{it}, \sigma_{it})$. The system of the following form is considered for each exchange i:

$$Y_{it} = \gamma_i + \Phi(L)Y_{it} + \Gamma\beta + \varepsilon_{it}. \tag{51}$$

Here, $\Phi(L)$ is a lag-function that controls for the effects of the lagged endogenous variables, and Γ contains a host of exogenous variables to explain price differences and quoted spreads. In the first step the lag order in the VAR is chosen on the basis of the Akaike Information Criterion (AIC) and the Schwarz Information Criterion (SIC) individually for each exchange. The VAR models of several exchanges, built by the lag order p_i chosen due to the minimum AIC value, exhibit serially correlated residuals according to the Portmanteau test for the hourly series and Breusch-Godfrey LM test for the daily series[9]. Thus, additional lags were added if the test indicated the presence of autocorrelation. The null hypothesis of the test is that there is no serial correlation of any order up to h, and h was chosen according to the seasonal pattern of the data and was set to twice its value. Although the daily pattern was not strongly pronounced for all exchanges,

[9] Since the approximate $\mathcal{X}^2-$ distribution is obtained under the assumption that h (the lag order for the test) goes to infinity with the sample size, the Portmanteau test should be used to test for high-order residual autocorrelation, whereas the LM test is suitable for testing for low order residual autocorrelation, since it will be not possible to estimate the parameters in the auxiliary regression model for large h (see Lütkepohl and Kilian (2017) and Lütkepohl (2005))

h was set to 48 hours for the hourly data. For the daily data, h was set to 14, corresponding to the weakly pattern of the bitcoin price differences. The lag order is eventually selected as p_i for each exchange i due to relatively minimized value of AIC and no autocorrelation in the model residual[10]. If the models with longer lags produced impulse responses that feature wiggles and jagged oscillations that are symptomatic of overfitting (Belongia and Ireland, 2021), large lag order was avoided. Then, the strategy, suggested by Lütkepohl (2005), is performed: different lag orders, obtained with minimizing AIC and reducing autocorrelation in the model residuals, are compared. The results are largely unaltered by the use of different lag orders. The choice of smaller lag lengths is also justified because slopes of the information criterion (as a function of lags) are typically flat for larger lag lengths (Roll et al., 2007). The chosen lag orders p_i are shown in the corresponding tables for Granger-causality tests and figures with impulse response functions in Appendix.

The model is estimated equation-by-equation using ordinary least squares. A sufficient condition for the consistency of the LS estimator would be that innovations ε_{it} is continuous i.i.d. with four finite moments random vector (see Lütkepohl (2005), Chapter 3). The white noise assumption does not allow serial correlation in the errors, but may be relaxed to allow for conditional and unconditional heteroskedasticity (see Lütkepohl and Kilian (2017), Chapter 2). Estimating univariate GARCH(1,1) models for the VAR residuals and inspecting autocorrelation functions of the squared residuals show that both processes (price differences and spreads) exhibit conditional heteroskedasticity[11]. Following Hafner and Herwatz (2009), in large samples as in this study, it is sufficient[12] to use heteroskedasticity consistent Granger causality test based on the White (1980) correction of standard OLS inference.

The models were estimated using both unadjusted and adjusted for seasonality and trends time series. The adjustment of price differences for calendar regularities and trends might or might not remove possible causation. For example, Figure 2 shows that, over the time period, the bitcoin price differences have become more integrated, especially in the second half of the sample period, i.e. in July and August. This change could be attributed to a long-term increase in liquidity, while other events, for example, fiat (il)liquidity (as discussed in Section 5.2, fiat (il)liquidity was a major problem for fiat-to-crypto exchanges in the sample period), or techno-

[10] Results of the autocorrelation test for the residuals of the VAR models for each exchange are available upon request

[11] The estimates from this exercise are not reported for brevity, but the results are available upon request.

[12] Hafner and Herwatz (2009) shows that a modified least squares statistic that is heteroskedasticity consistent behaves better in large samples (T = 1000, 2000 according to the Monte Carlo study) than the suggested bootstrap versions.

logical innovations such as new trading and arbitrage bots substantially speeding up the trade execution, could also have played a role. Therefore, in order to assess the impact of the adjustment procedure, the vector autoregressions are recomputed using the unadjusted series and reported in the tables below. The results will show that the conclusions about Granger causality tests and impulse response functions turn out to be insensitive to the adjustment procedure for most exchanges.

6.2.2 Estimation Results 1 - Granger Causality Tests

Granger-causality tests were performed to assess the effect of spreads on price differences and vice versa for each exchange. In particular, it is tested whether the coefficients on all lags of the causing variable are jointly equal to zero. Table 5 provides the results of the Granger-causality tests for the hourly time series. The rows in Panel B show $\mathcal{X}^2$-statistic and p-value. The table consists of two columns, where the test results for both adjusted and unadjusted series are presented. Table 13 in Appendix reports pairwise correlations of innovations from each VAR. Apart from Bitstamp and Gemini, all correlations between spread and price differences are positive and range from 0.04 to 0.11.

In line with Roll et al. (2007), the results suggest that for most exchanges, liquidity, measured by quoted bid-ask spread, is positively related to price differences. The movements in price differences and spreads are positively correlated due to either contemporaneous causal relationship or because of the common influence of other variables (order imbalances, for example). The price difference Granger-cause spread only on two out of six exchanges: Coinbase Pro and CEX.IO. Reverse causality running from spreads is not found on any exchange, except Bitstamp, however the causality turns to be insignificant after controlling for calendar regularities and trends. On Bitfinex, Gemini and Kraken volatility alone Granger-causes quoted spread, on Kraken volatility alone Granger-causes price differences.

Table 17 in the Appendix reports Granger-causality tests for the daily series. First, as at the hourly level, there is no causality running from spreads to price differences on any exchange at the daily frequency. As for the reverse direction, the results are consistent with the Granger tests at an hourly frequency for Coinbase Pro: price differences Granger-cause spreads, suggesting that the effect of price differences on spreads on Coinbase Pro remains significant at a daily horizon, at which the impact of asynchronous trading should be fairly small. The results also reveal that price difference innovations on Coinbase Pro have a stronger effect on liquidity. As for other exchanges, the Granger-causality tests at the daily level are in line with the hourly Granger

Table 5. Granger Causality Tests with Quoted Spread for the hourly series

Bitfinex	Not Adjusted (26 lags)			Adjusted (18 lags)		
	PD	QS	σ_t	PD	QS	σ_t
PD	-	19.85 (0.40)	21.8 (0.29)	-	9.25 (0.16)	39.98*** (< 0.01)
QS	16.34 (0.63)	-	38.95*** (0.004)	3.73 (0.71)	-	15.17** (0.01)
σ_t	18.25 (0.50)	59.93*** (< 0.01)	-	6.09 (0.42)	50.92*** (< 0.01)	-
Bitstamp	Not Adjusted (26 lags)			Adjusted (10 lags)		
PD	-	29.39 (0.21)	18.91 (0.76)	-	12.04 (0.28)	11.78 (0.32)
QS	21.97 (0.58)	-	23 (0.52)	8.09 (0.62)	-	7.94 (0.63)
σ_t	13.39 (0.95)	34.52* (0.07)	-	1.19 (0.98)	4.46 (0.61)	-
CEX.IO	Not Adjusted (29 lags)			Adjusted (28 lags)		
PD	-	34.07** (0.03)	31.95** (0.04)	-	28.76 (0.42)	47.84*** (< 0.01)
QS	14.94 (0.78)	-	53.64*** (< 0.01)	20.08 (0.86)	-	62.03*** (< 0.01)
σ_t	25.60 (0.18)	115.03*** (< 0.01)	-	25.95 (0.57)	105.95*** (< 0.01)	-
Coinbase Pro	Not Adjusted (24 lags)			Adjusted (9 lags)		
PD	-	27.43 (0.28)	35.67* (0.06)	-	18.11** (0.03)	37.32*** (< 0.01)
QS	32.03 (0.13)	-	28.37 (0.24)	15.02* (0.09)	-	10.81 (0.29)
σ_t	29.94 (0.17)	36.56** (0.05)	-	15.47* (0.08)	44.22*** (< 0.01)	-
Gemini	Not Adjusted (25 lags)			Adjusted (9 lags)		
PD	-	20.47 (0.43)	28.15 (0.11)	-	4.92 (0.42)	45.91*** (< 0.01)
QS	13.59 (0.85)	-	30.21* (0.07)	2.61 (0.76)	-	683.63*** (< 0.01)
σ_t	32.14 (0.15)	36.56* (0.06)	-	9.49* (0.09)	15.50*** (0.008)	-
Kraken	Not Adjusted (20 lags)			Adjusted (20 lags)		
PD	-	31.90 (0.13)	39.35** (0.02)	-	9.92 (0.13)	35.24*** (< 0.01)
QS	21.03 (0.64)	-	22.27 (0.56)	2.00 (0.92)	-	27.27*** (< 0.01)
σ_t	24.13 (0.45)	72.77*** (< 0.01)	-	13.77** (< 0.01)	59.32*** (0.001)	-

Notes: The table reports Granger-causality tests. Null hypothesis: row variable does not Granger-cause column variable, more precisely, it is tested whether the coefficients on all lags of the causing variable i are jointly zero when j is dependent in the VAR. Both chi-square statistics and p-values (in parentheses) of pairwise Granger-causality tests between the endogenous variables are presented. p-values are based on heteroscedasticity-consistent (HC) standard errors (White (1980) corrected standard errors). The lags in parentheses represent the lag order chosen by AIC and SC information criteria.

results: no causality running from price differences to spread and vice versa. On Gemini volatility alone Granger-causes price differences and spread, on Coinbase Pro - price differences.

6.2.3 Estimation Results 2 - Impulse Response Functions

Whereas, Granger causality test is based on a single equation, impulse response functions (IRFs) examine the joint dynamics of liquidity and price differences implied by the full VAR system by tracing the impact of one standard deviation innovation to a specific variable on the current and future values of the chosen endogenous variables. Tables 13 and 16 in the Appendix show that innovations are correlated, so the impulses need to be orthogonalized by using the inverse of the Cholesky decomposition of the residual covariance matrix. In order to identify the effects of shocks to price differences or spreads on the subsequent time paths of both, an assumption about whether the correlation is due to current price difference affecting current spread or due to spread affecting price difference has to be made. Given the evidence for lagged effects from Granger causality tests (price differences Granger-cause spread but not vice versa), price difference is placed before spread in the ordering. Since volatility alone could capture the common dynamics of price differences and liquidity, it is placed first in the ordering. Although the impulse response functions are generally sensitive to the ordering of the endogenous variables, reestimating the VAR model using all of the other 5 possible permutations of the order of the input variables shows that the qualitative results have not changed[13].

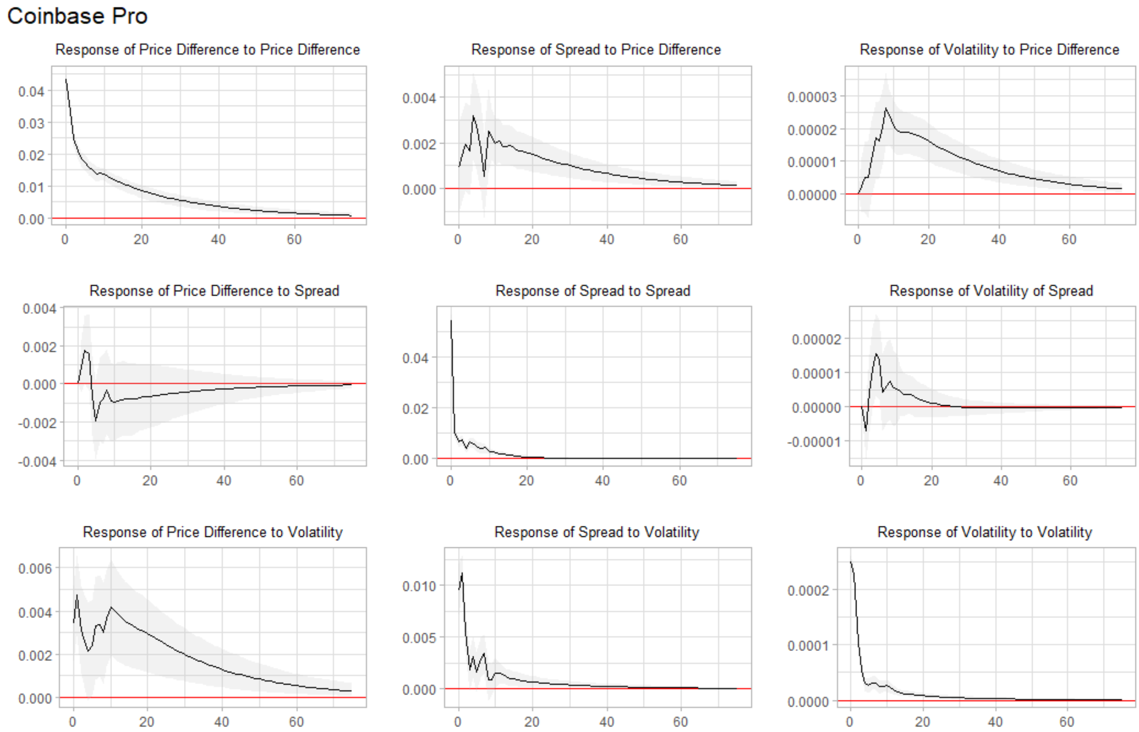
The impulse responses in Figs. 30 to 32 in the Appendix are estimated by exchange using the hourly time series, and illustrate the response of an endogenous variable to a unit standard deviation orthogonalized shock in the endogenous variable for a period, chosen by AIC and SC information criteria as reported by Table 5 in the Appendix for corresponding exchanges. To evaluate statistical significance of the responses, bootstrapped confidence intervals (0.95%) based on 1000 replications are provided along with the estimated responses. All impulse responses decay over time and the shock dies out quickly, supporting the results from the unit roots tests that indicate stationarity in Table 7 in Appendix.

The persistency of bitcoin price differences and bid-ask spreads is confirmed by the shocks which are informative in forecasting future values of the same variable. The diagonal panels in Figs. 30 to 32 in the Appendix show the effects of shocks to each variable on its own future values. The off-diagonal panels (bottom-left and top-right) show the effects of a spread shock on the path of price difference and vice versa. The IRFs of three out of six cryptocurrency exchanges

[13]The graphs of the reverse ordering are not reported for brevity, but are available upon request.

- CEX.IO, Coinbase Pro, Kraken reveal that shocks to the price differences are significantly informative in predicting future shift in spreads, but not in other way around. The results also indicate the robustness of the impulse responses to the inclusion of volatility and suggest that volatility alone does not capture the dynamics of bid-ask spread. The IRFs for Bitfinex, Bitstamp, Gemini and Coinbase Pro are consistent with Granger causality tests: innovations in the price differences have a significantly positive effect on spread for Coinbase Pro, for Bitfinex, Bitstamp and Gemini - cross-effects between price differences and spread are not significant in any period and in any direction, and volatility alone explains price differences and spread. As Figure 13 shows, on Coinbase Pro the price difference is informative in predicting shocks to spread from the 6th hour to 50th hour, i.e. almost two days, for CEX.IO and Kraken the effect is less persistent - from the current period (contemporaneous effect) to the 10th period for CEX.IO and to the third period for Kraken. For CEX.IO and Kraken, for which the Granger-causality tests do not reveal any relationships between price differences and spreads, the results of the IRFs show that after accounting for the joint dynamics, including the persistence of price differences and spreads, shocks to price differences are statistically informative in forecasting spreads at least in the first three hours, although the effect seems to be rather fleeting.

Figure 13. Impulse response functions of Coinbase Pro for the VAR with the hourly price differences, quoted bid-ask spread and volatility

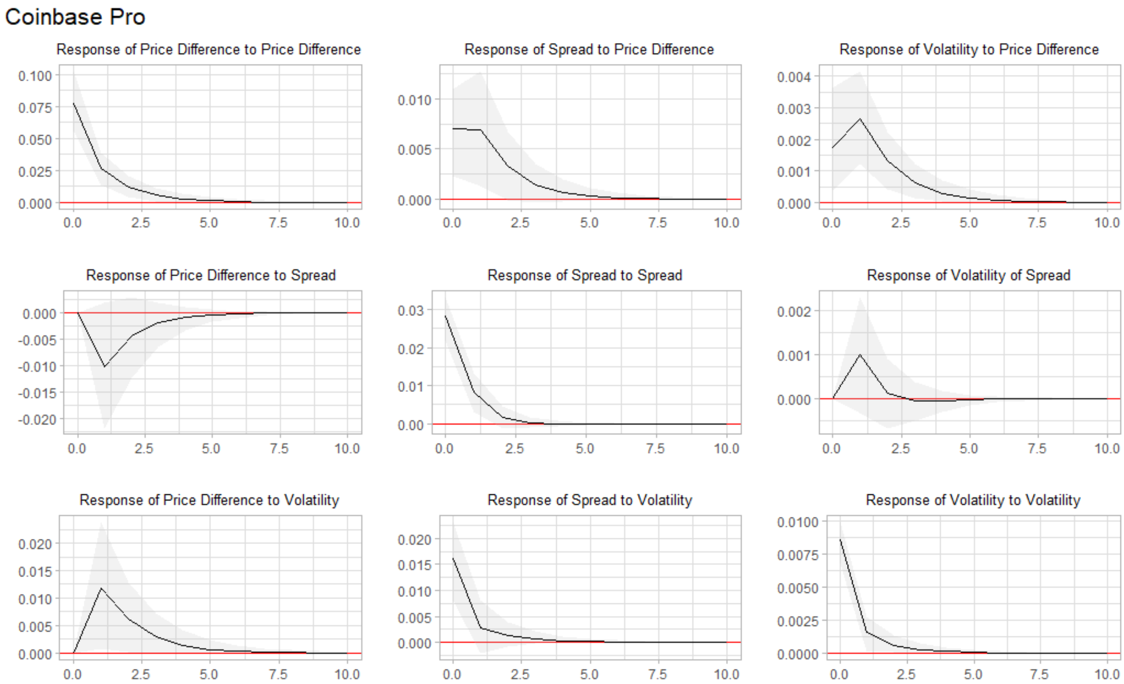


Notes: Impulse responses are presented for the vector autoregression that includes three endogenous variables: daily price differences, quoted bid-ask spread and volatility on Coinbase Pro over the time period from March 29, 2018 to September 5, 2018 (all variables are adjusted for time trends and other regularities, see Section 6.1). Bootstrapped confidence intervals (0.95%) are provided along with the estimated responses.

The bottom-left panel of the IRFs for the hourly time series shows that there is no statistically significant response of price difference to spread for any exchange and for any period. The effect is also always zero in the current period (at zero lag). This could be a result of the identification assumption: the condition that spread has no immediate effect on price difference was imposed in order to identify the shocks. The identification assumption imposes the condition that any "common shocks" that affect both time series are assumed to be price difference and volatility shocks, with spread shocks being the part of the spread VAR innovation that is not explained by the common shock. In this way, contemporaneous correlation in shocks runs only from the price differences and volatility to spread. This might cause the spread shocks to have smaller variance and might also overestimate the effect of the price difference and shocks on spread. To assess the sensitivity of the conclusions to the ordering assumption, the IRFs making the opposite assumption: that contemporaneous correlation in the innovations is due to spread shocks affecting bitcoin price differences, are examined. The graphs of the reverse ordering are not reported for brevity, but are available upon request. As expected, the effect of the price difference on spread now begins at zero. However, as with the previous ordering, there is still a significant response of spread to price differences at later lags for all exchanges except Bitstamp. In this way, at an hourly frequency, the impulse response functions are not sensitive to the ordering of the endogenous variables the VAR model at the hourly level. These results suggest that the price differences have a significant effect on spread but not vice versa, because interpreting the VAR results in favor of spreads's effect (by putting them first in the order) does not give spread an effect on the price differences, and the price difference shocks are clearly important for both variables.

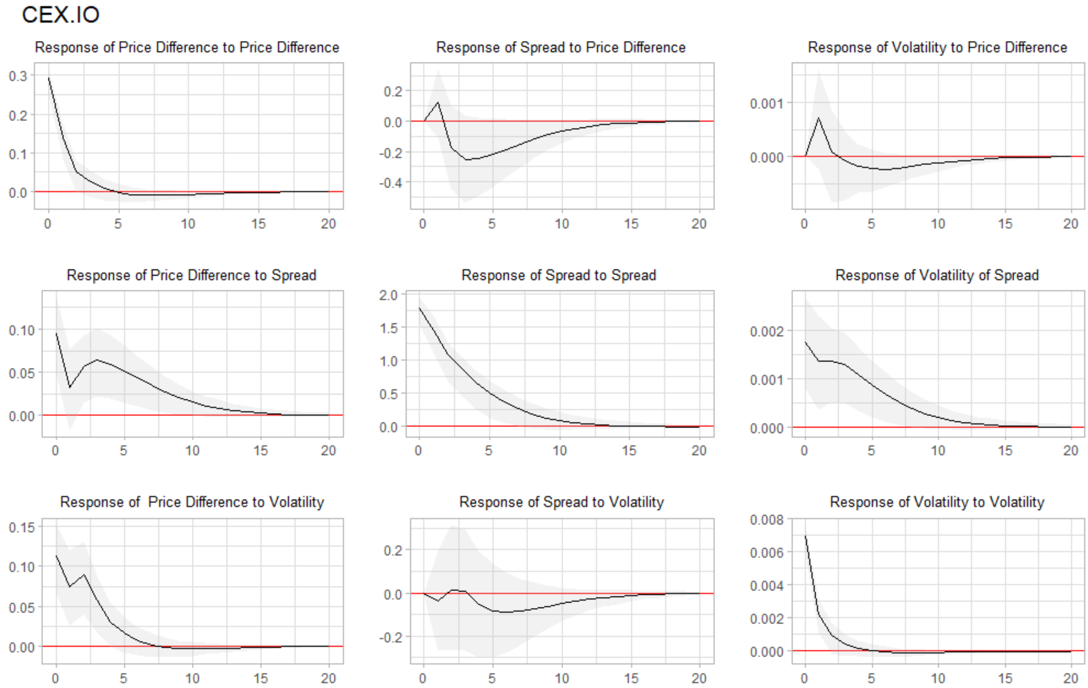
The IRFs functions of the daily time series are shown in Figs. 36 to 38 in the Appendix. The IRFs are largely consistent with the Granger-causality results at a daily frequency: the price difference is informative in predicting shocks to the spread only on Coinbase Pro. Figure 15 shows that impulse responses of spreads to price differences are significant at CEX.IO too, however, this result is sensitive to the ordering of the variables in the VAR system. Specifically, the IRFs results on Coinbase Pro and CEX.IO show that shocks to price differences are statistically informative in forecasting spreads at the first and second day. Whereas on Coinbase Pro this effect is insensitive to the VAR ordering, on CEX.IO the results show insignificant impulse responses of spread to price differences in VAR orderings where spread is placed before price differences. Since there is no confidence which identification assumption is better, there is a considerable uncertainty left in interpreting the effect of price differences to spread on CEX.IO. However, in contrast to other exchanges, where there is no significant impulse responses of price differences to spreads at any frequency, there is a statistically significant positive response from price differences to spread on CEX.IO at the daily frequency. Moreover, this effect is insensitive to the VAR ordering. With the orderings, where spreads come before price differences, the significant positive effect of spread to price difference starts in the first period, continues in the third, and lasts up to the tenth period. These results suggest that on CEX.IO, at the daily level, the spread has a stronger effect on price difference than vice versa. Interpreting the VAR results in favor of the spread effect virtually wipes out the effect of price difference on spread, but the effect of spread is still significant in all VAR orderings. These results were expected, since on CEX.IO, the effect of price difference on spread is not very persistent at the hourly frequency, it lasts 10 hours, but does not continue on the next day. In contrast, the results of Coinbase Pro in Figure 14 shows that the price differences have a significant positive effect on spread but not vice versa, and the effect is insensitive to the VAR ordering and to the frequency of the time series, and consistent with Granger-causality tests at each frequency. These results again suggest that price difference innovations on Coinbase Pro have a stronger and longer-lasting effect on liquidity. Possible reasons for that will be discussed in the next section.

Figure 14. Impulse response functions of Coinbase Pro for the VAR with the daily price differences, quoted bid-ask spread and volatility



Notes: Impulse responses are presented for the vector autoregression that includes three endogenous variables: daily price differences, quoted bid-ask spread and volatility on Bitfinex over the time period from March 29, 2018 to September 5, 2018 (all variables are adjusted for time trends and other regularities, see Section 6.1). Bootstrapped confidence intervals (0.95%) are provided along with the estimated responses.

Figure 15. Impulse response functions of CEX.IO for the VAR with the daily price differences, quoted bid-ask spread and volatility



Notes: See the notes for Figure 14

The IRF's results of Bitstamp and Gemini at the daily level are also consistent with Granger-causality tests at each frequency and with the IRFs estimation results at an hourly frequency: there is no significant relations between price differences and spread. As for Bitfinex and Kraken, the IRFs results are sensitive to the order in which variables appear in the model. This sensitivity reflects high contemporaneous correlation of the residuals of the estimated model. The high residual correlations are the result of a strong common component in the shocks, when volatility is omitted from the model. Cholesky decomposition effectively loads the common component onto the first variable in the ordering. It seems reasonable to put volatility always in the first place, and assess sensitivity only to permutations of the order of price differences and spread. With this approach, there is little sensitivity to the ordering of price differences and spreads. The IRFs results for Bitfinex and Kraken are consistent with Granger-causality tests at each frequency: there is no any significant relations between price differences and spread. This result was expected, since, at the hourly frequency, the IRF of Kraken shows that the effect of price difference to spread lasts up to 10 hours, but does not exceed a single day.

It is noteworthy to explain, why hourly Granger Tests are not significant at Kraken and CEX.IO, but their impulse responses shows significant impact of price differences on spread. First, hourly frequency due to high autocorrelation forces to include many lags in the Granger test, leading to some insignificant lags. Since it is tested whether the coefficients on all lags of the causing variable are jointly equal to zero, many insignificant lags may lead to accepting the null hypothesis. Second, Dufour and Tessier (1992) show that if variable X in multivariate systems (with more than two variables), does not cause Y in the sense of Granger, the innovations of X may still account for a sizeable proportion of the variance of Y. In particular, they give general and sufficient condition that shows that the duality between AR and MA characterizations of non-causality, holding in bivariate systems, does not generally extend to multivariate systems which include variables other than X and Y (which is the case in this thesis) (Dufour and Tessier, 1992).

6.2.4 Estimation Results 3 - Discussion on the Order Imbalance

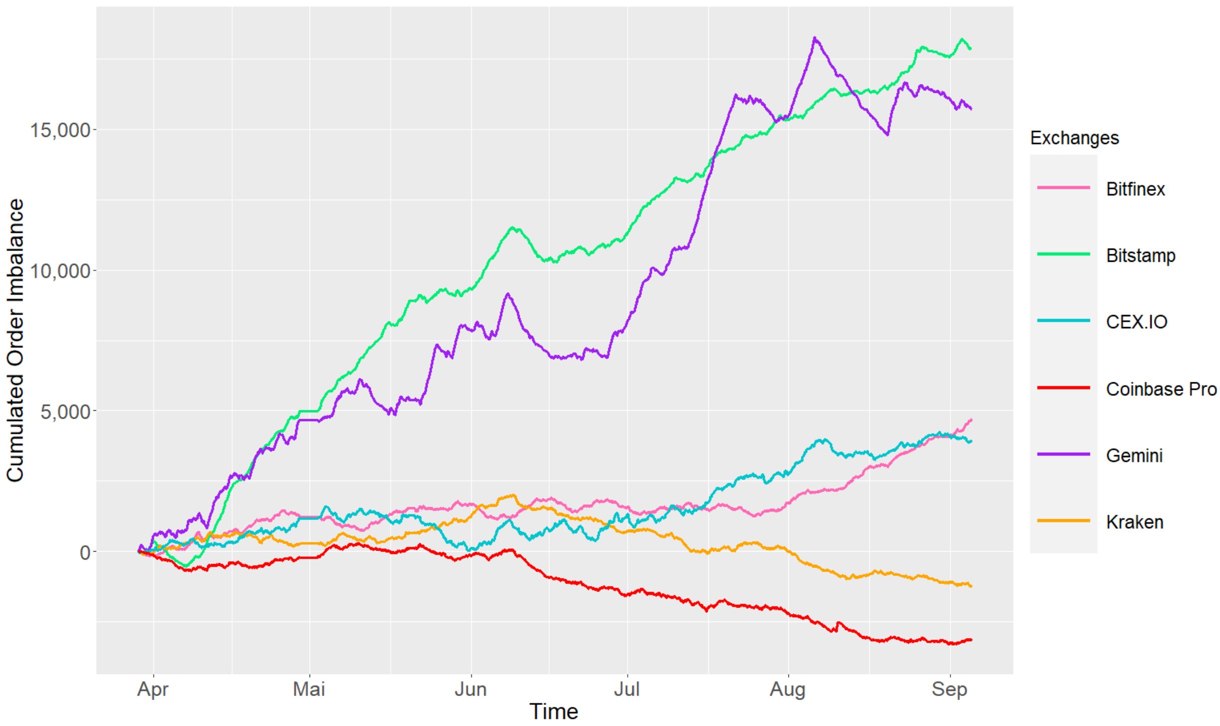
Coinbase Pro is only one exchange, where the effect of price differences on spread is longer-lasting - the effect persists beyond a single day. The exchange exhibits a very high order book imbalance during the sample period, and the imbalance is the highest among all the exchanges in the sample. Table 6 shows the order book imbalance, which is split into positive and negative parts. Order book imbalance is calculated as in Equation 10, and is positive (negative), when there is an excess of buy (sell) limit orders. Therefore, these measures reflect inventory imbalances faced by liquidity providers. Both excess sell and buy orders are quite substantial on Coinbase Pro - 0.68 on average, whereas on other exchanges, order imbalances are medium in favor of both sides. If price differences on Coinbase Pro are mostly associated with price pressures, the high order imbalance could be a reason for the longer-lasting effect of price difference on bid-ask spread, which is consistent with the inventory paradigm, which suggests that inventory problems, exacerbated by order imbalances, create price pressures and reduce liquidity (Chordia et al., 2002). Figure 16 visualises cumulative sum of the minutely order imbalances over time and across exchanges. By summing the values up over time, the time series have been smoothed, making it easier to distinguish and compare multiple overlapping plots in one graph. Chordia et al. (2002) finds that the effect of order imbalance on S&P500 returns is asymmetric: excess sell orders have an impact four times that of excess buy orders. Compared to other exchanges, excess sell orders on Coinbase Pro is higher and occur more often than excess buy orders. Hence, the estimation results on Coinbase Pro are consistent with the block trading literature, which suggests that price pressures caused by the excess of sell orders are greater than those for buy orders (Chordia et al., 2002). In this way, such a high order book imbalance on Coinbase Pro indicates severe inventory problems that have an extended effect on liquidity, which persists beyond a single day as shown by the results of the impulse response functions and Granger-causality tests.

Table 6. Summary statistics of the exchange-specific measures for excess buy and sell orders

Exchange	Excess sell			Excess buy		
	Mean	Median	p75	Mean	Median	p75
Bitfinex	0.49	0.49	0.73	0.50	0.51	0.74
Bitstamp	0.41	0.38	0.62	0.46	0.44	0.69
Coinbase Pro	0.68	0.77	0.93	0.66	0.75	0.92
Gemini	0.42	0.39	0.64	0.53	0.49	0.89
CEX.IO	0.37	0.33	0.55	0.38	0.34	0.56
Kraken	0.40	0.37	0.60	0.40	0.36	0.60

Notes: Descriptive statistics is given for minutely order imbalance measures for the sample exchanges over the time period from 02:00:00, March 29, 2018 to 01:00:00, September 5, 2018. Excess buy orders are calculated as Max$[0, OI_t]$, Excess sell order -|Min$[0, OI_t]$|, where $OI_t \in [-1, 1]$ and calculated as in Equation 10 in Section 3.2.

Figure 16. Cumulated Order Imbalance across Exchanges over Time



Notes: This figure shows the cumulative sum of the minute-level order imbalances over the time period from 02:00:00, March 29, 2018 to 01:00:00, September 5, 2018. Order imbalances are calculated as in Equation 10 in Section 3.2.

In contrast to Coinbase Pro, other exchanges either do not show any significant impact of price differences on spread or show the impact, which lasts at maximum 10 hours. If their price differences arise due to price pressure, the estimation results could be explained by their medium order imbalances (see Table 6), which are almost two times smaller than on Coinbase Pro. It implies that, during the periods of imbalances when market orders tend to occur on one side of the market, to elicit trading on the other side of the market, liquidity providers manage to adjust the quotes within an hour for Bitfinex, Bitstamp and Gemini, and within 10 hours at most for Kraken and CEX.IO. If most of the deviations arise as a result of differences in information, which is possible at the exchange level, then price differences should positively affect spread because liquidity providers charge larger spreads to cover the risk of trading at stale quotes (Copeland and Galai, 1983; Foucault et al., 2017). In this case, however, the hourly frequency could be too low to detect the effect of price differences on spread and will miss most of the pattern. First of all, according to Foucault et al. (2017), "toxic" price differences are mostly short-lived, because arbitrageurs react faster to such opportunities, and also technological possibilities should reduce the duration of arbitrage opportunities. For example, Foucault et al. (2017) observed arbitrage opportunities in Reuters D-3000 from January 2003 to December 2004 and found that 25% of all arbitrage opportunities last less than half a second. As for the Bitcoin market, taker fee-adjusted price differences last from 8 to 15 minutes on all exchanges in the sample, except CEX.IO, where the average duration is 53 minutes (see Table 4). It less clear whether there are more high frequency arbitrageurs in the Bitcoin market, who buy and sell on two markets simultaneously or more arbitrageurs, who transfer the assets across markets during arbitrage and are exposed to the blockchain related settlement latency (see Section 4.1). The frequency at which trading patterns and bitcoin price deviations evolve is also unclear. Moreover, it is not obvious whether there are traders who are more informed than others, and to what extent hourly and daily price differences reflect asynchronous adjustments following information arrival. Nevertheless, hourly frequency may be too low to capture the effect of asynchronicities on liquidity and miss most of the pattern, given that bitcoin price differences persist on average 10 minutes and the prices may already converge within an hour. On the other side, the hourly and daily intervals are justified when price differences mainly reflect price pressures, caused by imbalances in inventories, because inventory effects are most likely to be pronounced over rather short horizons as opposed to weekly or monthly intervals, while very high frequency data is plagued by inter-asset synchronicities (Chordia et al., 2002). To sum up, the effect of price differences on spread on Bitfinex, Bitstamp and Gemini is insignificant, either because their price differences are mostly driven by the differences in information and the hourly frequency is too low for the adverse selection effect to be manifested, or most of the price differences reflect price pressures, but inventory problems of liquidity providers do not persist beyond a single hour, since order book imbalances are not as severe on these exchanges as on Coinbase Pro (see Table 6). The last argument applies also for CEX.IO and Kraken, where the effect of price differences on spread lasts several hours, but does not persist beyond a single day. As a result, the hourly frequency might be too low to capture the effect of price differences regardless of the reason of their occurrence. Although the order book imbalances on Bitfinex, Bitstamp and Gemini are at the same level as on CEX.IO and Kraken, CEX.IO and Kraken still show significant positive effects of price differences on bid-ask spread. One possible explanation is the speed of the quotes adjustment as a response to inventory problems, which may be lower on CEX.IO and Kraken, since these exchanges are relative small and illiquid during the sample period.

In order to estimate the impact of order imbalance on price differences, one can include order imbalance's measure or excess sell and buy orders as endogenous variables in the VAR system with price differences, spread and volatility. However, as stated by Makarov and Schoar (2020), bitcoin is traded simultaneously on multiple exchanges, therefore, when the traders decide how much to buy or sell at the exchange, they will take into account prices on other exchanges, where bitcoin is traded. As a result, the effect of order imbalance on price differences, measured in a vector autoregression in each exchange separately, may be biased (Makarov and Schoar, 2020). One way to accommodate the case of multiple exchanges would be to follow Makarov and Schoar (2020) and combine factor analysis with price decomposition in Hasbrouck (1995). Order flow on each exchange can be decomposed into a common component and an idiosyncratic, exchange-specific component (Makarov and Schoar, 2020). This analysis is beyond the scope of this thesis, however, it could be it's possible empirical extension.

6.2.5 Summary of Results

Summing up the estimation results, Granger tests do not show two-way causality: price difference Granger-causes spread on Coinbase Pro, and there is no reverse causality running from spreads to price differences at any exchange, which is not quite in line with Roll et al. (2007), who shows two-way Granger causality between the 3-month absolute basis and liquidity. The IRF's results are nevertheless largely in line with Roll et al. (2007): for the hourly frequency impulse response functions reveal that shocks to the absolute price differences are significantly informative in predicting future shifts in spreads on CEX.IO, Coinbase Pro and Kraken. At the daily level, the effect remains significant only on Coinbase Pro (according to both: Granger causality tests and IRF). Regardless of the reason for the arbitrage opportunity to arise, a lasting and significantly positive relation between price deviations and illiquidity was expected. If arbitrage opportunities arise due to asymmetric private valuation shocks ("toxic" arbitrage opportunities, see Foucault et al. (2017)), then market makers, exposed to the risk of trading with arbitrageurs at stale quotes, charge wider bid-ask spreads to cover adverse selection risk (Copeland and Galai, 1983). If price differences reflect price pressure, then price differences will have a positive impact on spread, because the market maker's inventory problems, exacerbated by order imbalances, implicitly raise the costs of providing liquidity in direction of imbalances, and, hence, widen bid-ask spreads (Chordia et al., 2002).

Shocks to spreads are informative in forecasting price differences only on CEX.IO at the daily frequency, suggesting that liquidity affects arbitrageurs more in relatively illiquid markets. This result is also consistent with Roll et al. (2007), who found that liquidity affects arbitrageurs only in the relatively less actively traded longer-term futures contracts (according to IRFs). Furthermore, Section 3.3 shows that CEX.IO stands out from other cryptocurrency markets with the longer-lasting and larger price deviations. Price deviations that correspond to arbitrage opportunities on CEX.IO persist, on average, 53 minutes, whereas the average duration on all other exchanges in the sample ranges from 7 to 15 minutes. Such a long duration of arbitrage opportunities suggest that CEX.IO - an illiquid (see Table 2) market with fiat withdrawal problems (see Section 5.2), was unattractive for arbitrageurs, especially for institutional traders. The combination of relatively large price differences together with fiat and bitcoin illiquidity may have prevented arbitrageurs from bringing the bitcoin prices closer to fundamental values on CEX.IO. When liquidity was higher, it could facilitate arbitrage trades, making relative pricing more efficient on CEX.IO. This effect can be captured however only at a daily frequency, the hourly horizon is not long enough for this effect to be significant at CEX.IO.

According to Hautsch et al. (2020), 91% of price differences[14] adjusted for transaction costs are within arbitrage bounds, which increase with spot volatility, risk aversion, expected blockchain-related settlement latency, and uncertainty in this latency. Further, Hautsch et al. (2020) finds that risk-averse arbitrageurs explore price differences only if potential gains offset latency-implied price risk. Therefore, the finding that the effect of spread on price difference is insignificant for all exchanges except CEX.IO is consistent with the evidences provided by Hautsch et al. (2020). The majority of price differences on such relative liquid exchanges as Bitfinex, Bitstamp, Coinbase Pro, Gemini and Kraken, do not constitute arbitrage opportunities, and their arbitrage opportunities persist for 7-15 minutes. Hence, one can expect that the arbitrageurs were able to correct for the price discrepancies in a relative short period of time, and their trades were not sensitive to liquidity, at least during the sample period. Another reason might be a frequency at which the VAR is estimated. Although it is unclear at which frequency pattern of trading and bitcoin price deviations evolve, the hourly frequency may miss most of the pattern, given the duration of arbitrage opportunities on Bitfinex, Bitstamp, Coinbase Pro, Gemini and Kraken. In contrast to these exchanges, CEX.IO is a relatively illiquid exchange and exhibits higher price differences, 35% of which correspond to positive arbitrage profits, which persist almost one hour, indicating that arbitrageurs needed a longer time to eliminate the mispricing. The result of the

[14]Hautsch et al. (2020) retrieved all open buy and sell orders for the first 25 levels on a minute interval from January 1, 2018, to October 31, 2019. In this way, the sample period of this thesis coincides with the period considered by Hautsch et al. (2020)

impulse responses of price differences to spread suggest that illiquidity was one of the frictions that prevented arbitrageurs from exploiting price differences at CEX.IO.

6.2.6 Robustness Checks

The first robustness check involves an alternative measure of liquidity - average quoted depth, which is positively related to liquidity. The depth measure is calculated by averaging the order volumes at the best bid and ask quotes for each minute, as stated by Equation 9. Then, the resulting measure is aggregated to hourly and daily averages of average depth. Using the variables from Section 6.1 above, the depth is adjusted to avoid spurious covariation. in line with (Roll et al., 2007), the correlation between VAR innovations is negative for Coinbase Pro, Gemini and Kraken for the hourly series and for all exchanges except Kraken for daily series, suggesting that depth is negatively related to price differences, which is expected, since depth is positively related to liquidity. The Granger-causality tests and the IRFs also show (in Appendix see Table 15 and Figure 34 for the hourly frequency, Table 19 and Figure 40 for the daily frequency), that for Coinbase Pro, the effect of price differences on spread is not only longer-lasting, it is also robust to the alternative measure of liquidity - average quoted depth. The Table 15 in Appendix show that the price differences on Coinbase Pro Granger-cause quoted depth, and the results are unaltered by the adjustment procedure. Impulse response functions, demonstrated by Figure 40 are consistent with the Granger causality results. At the hourly level, as expected, a positive shock to price differences predicts a decrease in quoted depth. If the price differences on Coinbase Pro arise mainly due to price pressures, then this result is consistent with inventory paradigm, discussed above.

The previous inferences for CEX.IO and Kraken are altered by the use of depth instead of bid-ask spread, since impulse response functions do no reveal significant relations between depth and price differences in any direction. This result is not surprising, since CEX.IO and Kraken show a less persistent effect of price differences on spread: the effect lasts from the current period (contemporaneous effect) to the tenth period for CEX.IO and to the third period for Kraken. Also, this finding suggests that bid-ask spread may react faster and stronger to the increase in price differences and, hence, to the changes in order imbalance and arbitrage activity, than depth.

In the reverse direction, results for Coinbase Pro indicate that a one standard deviation positive shock to depth predicts an increase in price differences. This result is not in line with Roll et al. (2007), who shows that liquidity, measured in quoted depth, decreases price differences.

Liquid market should facilitate arbitrage trades, which move prices to an appropriate level, making Bitcoin market more efficient. Such a surprising result can be explained by the significant selling pressure, which Coinbase exhibited during the sample period (see Table 6 and Figure 16). The quoted depth were growing because of excess of sell orders, which pushed the bitcoin price on Coinbase Pro down. Although Coinbase Pro is a relative liquid exchange, and the data was winsorized to ensure that the results are not driven by outliers, cryptocurrency markets are still not as deep as traditional financial markets, therefore, there is still no sufficient volume of pending orders on both the bid and ask side, preventing a large order from significantly moving the price. Selling pressure on Coinbase Pro is explained by much deeper, well-regulated banking relationships (Casey, 2019) and for the highest level of KYC compliance in the cryptocurrency market (Golstein, 2018). US dollar fiat funds submitted at Coinbase Pro are held in anonymous FDIC[15]-insured banks located in the US, or in US treasuries (Zach, 2018). Furthermore, relative to other exchanges, Coinbase Pro is very liquid: according to Table 2, Coinbase Pro shows the highest median quoted depth of 16.21 BTC during the sample period, which is by far larger than the second and the third highest average quoted depth of 8.79 BTC and 2.55 BTC on Bitfinex and Gemini, respectively. Other exchanges, for example Bitfinex and CEX.IO are known for their serious problems with fiat withdrawals during the sample period (see Section 5.2), therefore, Coinbase Pro is one of the few fiat-to-crypto exchanges, attractive for fiat withdrawals, especially for the large institutional traders, who might posses large amounts of bitcoins in their wallets.

The final robustness check involves the effects of the Blockchain network activity on the bitcoin price differences. In order to capture the effects of the Bitcoin network activity, number of transactions waiting for verification and the network fee per byte is included in the model as exogenous variables (see the argumentation in the beginning of Section 6). The contemporaneous impact of two exogenous variables on both the bitcoin price differences and bid-ask spread is mostly not significant, and the impulse response functions also remain unchanged.

As stated by Makarov and Schoar (2020), one possible concern might be non-linearity in relationship between liquidity and price differences, which is not picked up in the VAR model. When the price differences on one exchange are very high, market makers might adjust which exchange they trade and how much liquidity they provide (see (Makarov and Schoar, 2020)), and as a result, the effect of price differences on liquidity might be stronger for larger price differences. Nevertheless, the results of vector regression analysis show that relation between bitcoin price

[15]The Federal Deposit Insurance Corporation (FDIC) is an independent agency created by the United States Congress to maintain stability and public confidence in the nation's financial system (FDIC, 2020).

differences and liquidity are consistent with the notion that price differences forecast and, in turn, are forecasted by liquidity. Bitcoin price deviations on a given hour (day) can help in predicting liquidity on subsequent hours (days) at least on some exchanges, which is consistent with both inventory paradigm and asymmetric information. Although it is less obvious whether there are any traders who have more information than others in the Bitcoin market, and the investigation of the ratio of the toxic bitcoin price differences to the price differences that reflect price pressure is beyond the scope of this thesis, previous literature, for example, Makarov and Schoar (2020) find that up to 85% of the day-to-day variation in bitcoin returns is explained by order flow imbalances. Also, hourly and daily are long enough to reduce the the influence of asynchronous trading (Chordia et al., 2002).

7 Conclusion

This thesis studied arbitrage and trading frictions in the Bitcoin market using a large dataset of minute-by-minute orderbook snapshots of the seven largest cryptocurrency exchanges that feature BTC versus USD trading. After describing the price differences between the cryptocurrency exchanges, a measure of the instantaneous arbitrage profit is constructed by optimally accounting for the prevailing orderbook depth and the fee schedule. More precisely, each minute, it is evaluated whether the trader, starting with a dollar position, could profit from purchasing bitcoins with dollars at one exchange and selling bitcoins for dollars at another exchange, all simultaneously at the relevant bid and ask prices. In that sense, the profit measure characterizes arbitrage profit in a scenario where the arbitrageur is not fully exposed to the settlement latency: inventory holdings on an exchange are not exhausted, or short-selling, trading at margin, borrowing from hodlers is accessible with negligible costs and risks. The full magnitude of arbitrage opportunity for each exchange pair is captured by identifying the trading quantity that maximizes the arbitrage profit given the bid-side and ask-side depth profiles and the fee schedule of corresponding markets at each minute. An arbitrage opportunity occurs if exchanges are crossed, e.g., the bid price at one exchange exceeds the ask price of the other exchange. It was shown that, by the construction of a limit order book, the quantity of bitcoins that maximizes arbitrage profit corresponds to the intersection of the bid and ask curves of the sell- and buy-side exchange, respectively. Also, the trading quantity corresponding to the intersection of the bid and ask curves is profit-maximizing after incorporating all fees, stemming from the exchanges and the blockchain network.

The arbitrage profits, documented in this thesis, provide systematic empirical evidence on the extent to which bitcoin price differences are related to the frictions associated with transaction costs: liquidity, measured by the order book depth, and exchange-specific fees arising from transacting, such as proportional taker fees, BTC and USD withdrawal fees. For all exchange pairs, except CEX.IO-related ones, 50%-80% of the minutely observed instantaneous first level cross-market price differences do not constitute arbitrage opportunities. In other words, for more than a half of the sample period, the markets are not crossed, i.e. the bid price at one market is lower than the ask price of the other market. Further, for all exchange pairs, excluding pairs involving CEX.IO and bitFlyer, more than 98% of the *positive* first level cross-market price differences do not correspond to arbitrage opportunities after adjusting the prices for the proportional taker fees. Thus, in the sample period, the instantaneous bitcoin price differences between the largest fiat-to-crypto exchanges Bitfinex, Bitstamp, Gemini, Coinbase Pro and Kraken can be almost fully explained by the exchange-specific taker fees. Indeed, the average magnitude of the positive first level cross-market price differences is 0.01%-0.1% (excluding CEX.IO and bitFlyer), which is below the taker fees ranging from 0.15% to 0.26%. Whereas, the bitcoin prices on these exchanges are more integrated, CEX.IO and bitFlyer stand out with longer-lasting and larger price deviations. In 60% of minutes in the sample period, CEX.IO as a sell-side market trades bitcoin at 0.7% premium on average. As a result, 10%-35% of minutes are profitable for arbitrage on CEX.IO with average profit ranging from $700 -$1000 per opportunity (if the optimal amount of bitcoins is traded). Arbitrage opportunities on CEX.IO persist on average 53.13 minutes, on the remaining exchanges - from 7.27 to 15.19 minutes.

After optimally accounting for the liquidity at the relevant price levels of the corresponding order books, up to 50% of the taker fee-adjusted first level price differences could not cover the USD and BTC withdrawal fees. Thus, the liquidity in the form of the order book depth plays a lesser role in explaining bitcoin price differences than the proportional taker fees, but still reduce the amount of profitable arbitrage opportunities.

In summary, price differences and arbitrage profits appear to constitute two groups. In one group, where the reference exchanges are CEX.IO and bitFlyer and they serve as sell-side markets, bitcoin consistently trades at higher bid prices on CEX.IO compared to the other exchanges, leading to higher and persistent arbitrage profits. For the remaining exchanges, which are also the largest fiat-to-crypto exchanges by trading volume for BTC/USD currency pair in the sample period, price differences and arbitrage profits are typically more episodic and smaller in magnitude. The magnitude of the taker fee-adjusted bitcoin price differences in this group is comparable to price deviations in other markets and assets. It ranges from 0.001% for Gemini to 0.02% on bitFlyer. To compare with "traditional" markets, Gagnon and Karolyi (2010) documented that the mean price difference for ADRs (American Deposit Receipts) is 0.049%, Wenxin et al. (2018) observed mean daily deviations for the covered interest parity that fluctuate from 0.06% to 0.19% basis points annualized. Nevertheless, the bitcoin price deviations in this group are more persistent than in the foreign exchange market with 1.554 seconds per day (Foucault et al., 2017), but comparable to ADR market with average duration of price deviations of 12.41 minutes (Rösch, 2021).

Overall, during the sample period, the highest and most frequent arbitrage profits are particularly noticeable in the first half, starting on March 29, 2018 until the end of July, 2018. The frequency and the magnitude of arbitrage profits clearly drops over the sample since the prices across exchanges have become more integrated in the second half of the sample period. In comparison to price differences in earlier periods, documented by Makarov and Schoar (2020) and Kroeger and Sarkar (2017), the bitcoin price differences have become smaller, suggesting a general trend of decreasing price differences and increasing efficiency of cryptocurrency markets.

Documented price differences and arbitrage profits suggest the existence of impediments that prevent arbitrageurs from taking advantage of the market inefficiency. To explore the idea that bitcoin price differences are related to the frictions associated with liquidity, the intraday (hourly) and daily horizon time-series interaction between price differences and bid-ask spread is modelled by vector autoregressions for each exchange. The empirical findings include the following: (1) In line with Roll et al. (2007), contemporaneous innovations to the bitcoin price differences and spreads are positively correlated. The correlations are negative if depth instead of spread is included in the VAR system; (2) fiat-to-crypto exchanges are heterogeneous with respect to the strength of the relation between price differences and liquidity; (3) in contrast to Roll et al. (2007), there is no Granger causality running from spreads to price difference; (4) one out of six cryptocurrency exchanges - Coinbase Pro - exhibits one-directional Granger causality running from bitcoin price difference to quoted spreads, but not in the other way around; (5) by accounting for the full dynamics of the VAR system, the impulse response functions of three out of six exchanges - Coinbase Pro, CEX.IO and Kraken reveal that the innovations in the price differences have a significant positive effect on the spreads, but not in the reverse direction, which is also consistent with Granger causality test results. These results suggest that the price differences have a significant effect on spread but not vice versa, because interpreting the VAR results in favor of spreads's effect (by putting them first in the order) does not give spread an effect on the price differences, and the price difference shocks are clearly important for both variables; (6) only Coinbase Pro shows a longer-lasting effect of price differences on spread, which persists beyond a single day. Moreover, at the hourly frequency, this effect is robust to the alternative measure of liquidity - average quoted depth; (7) CEX.IO and Kraken show a less persistent effect of price differences on spread: the effect lasts from the current period (contemporaneous effect) to the tenth period for CEX.IO and to the third period for Kraken; (8) for the remaining exchanges Bitfinex, Bitstamp and Gemini cross-effects between price differences and spread are not significant in any period and in any direction, and volatility alone explains price differences and spread; (9) shocks to spreads are not informative in forecasting shifts in price differences at any exchange and any frequency, except CEX.IO at the daily level. This finding is consistent with Roll et al. (2007), since CEX.IO is a relatively illiquid exchange with substantial and longer-lasting price differences, which may suggest that liquidity affects arbitrageurs more in the relatively less liquid markets.

The heterogeneity of cryptocurrency exchanges with regard to the strength of the effect of price differences on spread can have two explanations, which boil down to the fact that the hourly and daily frequency at which the VAR was estimated, might be too low to capture the effect of price differences on bid-ask spread. First, Coinbase Pro exhibits a very high order book imbalance, which suggest severe inventory problems that have an extended effect on liquidity, which persists beyond a single day. In contrast to Coinbase Pro, other exchanges exhibit medium order imbalances, which are almost two times lower that on Coinbase Pro. Therefore, in response to the order imbalances, market makers manage to adjust the quotes faster than on Coinbase Pro, and inventory problems of liquidity providers do not persist beyond a single hour at Bitfinex, Bitstamp and Gemini, and last at most ten hours at CEX.IO and Kraken. Second, if price differences on Bitfinex, Bitstamp and Gemini are mostly driven by the differences in information, the hourly frequency might be too low for the adverse selection effect to be manifested and might miss most of the pattern, given that the average duration of the taker-fee adjusted price differences between 7 and 15 minutes.

In another direction, shocks to spreads are informative in forecasting price differences only on CEX.IO at the daily frequency, suggesting that liquidity affects arbitrageurs more in relatively illiquid markets. 98% of price differences on such relative liquid exchanges as Bitfinex, Bitstamp, Coinbase Pro, Gemini and Kraken, do not constitute arbitrage opportunities, and their arbitrage opportunities persist for 7-15 minutes. Hence, one can expect that the arbitrageurs were able to correct for the price discrepancies in a relative short period of time, and their trades were not sensitive to liquidity. Another reason might be a frequency at which the VAR is estimated. Although it is unclear at which frequency pattern of trading and price deviations evolve, the hourly frequency may miss most of the pattern, given the duration of arbitrage opportunities on Bitfinex, Bitstamp, Coinbase Pro, Gemini and Kraken.

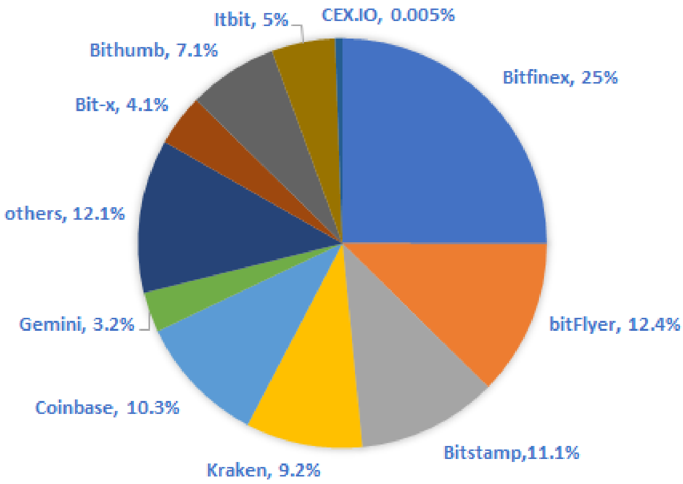
All results are preserved after controlling for bitcoin price volatility and for the effects of the Blockchain network activity. Overall, the estimation results of some cryptocurrency exchanges are consistent with the notion that price differences forecast and, in turn, are forecasted by liquidity at least on some exchanges. Bitcoin price deviations from the law of one price on a given hour (day) can help in predicting liquidity on subsequent hours (days), which is consistent with both inventory paradigm and asymmetric information. If arbitrage opportunities arise due to information, then liquidity providers, exposed to the risk of trading with arbitrageurs at stale quotes, charge wider bid-ask spreads to cover adverse selection risk (Copeland and Galai, 1983; Foucault et al., 2017). If price differences reflect price pressure, price differences will have a positive impact on spread, because the market maker's inventory problems, exacerbated by order imbalance, implicitly raise the costs of providing liquidity in direction of imbalances, and, hence, widen bid-ask spreads (Chordia et al., 2002; Rösch, 2021). In the reverse direction, liquidity plays an important role in moving the Bitcoin market toward more efficiency, if market's efficiency is measured by the degree to which the law of one price satisfied (Roll et al., 2007).

To empirically extent this analysis, it would be interesting to conduct the same analysis of price differences as in Rösch (2021) for the ADR market, i.e. to investigate to which extent bitcoin price differences are "toxic" as in Foucault et al. (2017) and to which extent bitcoin price differences arise due to price pressures and, hence, indicate inventory concerns. This analysis would help to understand the true impact of arbitrage trades on liquidity in the Bitcoin market, since if arbitrage opportunities are "toxic", arbitrageurs can harm market liquidity (Foucault et al., 2017). If price deviations arise as a result of price pressure, then arbitrageurs provide liquidity by trading against deviations. In the next step, to accommodate the case of multiple exchanges, one could use factor analysis and price decomposition as in Hasbrouck (1995) and dynamically relate the bitcoin price differences and order imbalances as in Makarov and Schoar (2020). For example, if there are problems to withdraw fiat funds from the exchange – withdrawals are delayed, rejected or even lost, then price changes can arise from order imbalances, resulting from an excess of buy or sell orders. Also, Easley et al. (2008) and Holden and Jacobsen (2014) suggest that order imbalances can be used as a measure of adverse selection, therefore, the market makers can consider orders imbalance as an indication of a "toxic" order flow, regarded as "toxic" when it originates from a better-informed counterparty, creating an adverse selection problem. In the ADR market, Rösch (2021) showed that less that 10% of all price deviations are "toxic", and found that price deviations indicate inventory concerns which are mitigated by arbitrageurs. How large is the fraction of "toxic" arbitrage opportunities and how arbitrage trades affects liquidity in the Bitcoin market would be an interesting area for further research.

8 Appendix

8.1 Figures

Figure 17. Total market share by exchange



Notes: The pie chart shows the market share by exchange for all currency pairs for the 1th of May, 2018 (bitcoinity.org, 2020). Coinbase trading volume is a trading volume of its two core products: Coinbase Pro (formerly GDAX) – professional trading platform, and a user-facing retailer broker of cryptocurrencies for fiat currency. One should note that the market shares by exchange are varying from source to source due to different sets and number of exchanges involved in the measurements.

Figure 18. Intersection in a point

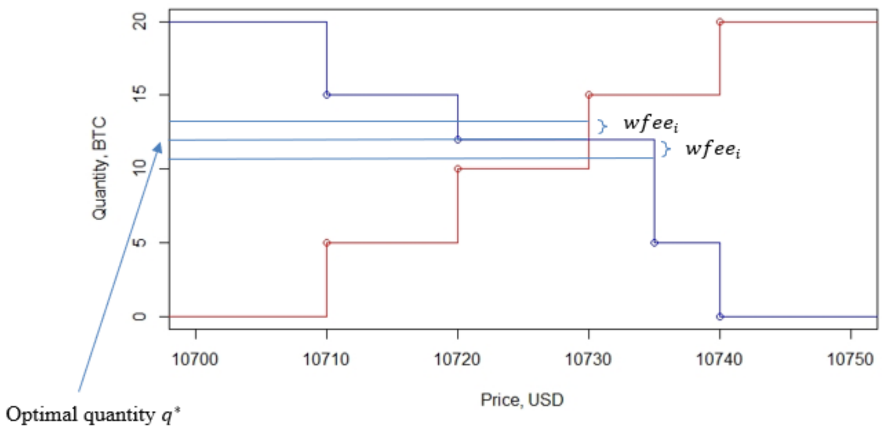


Figure 19. Vertical Tangency

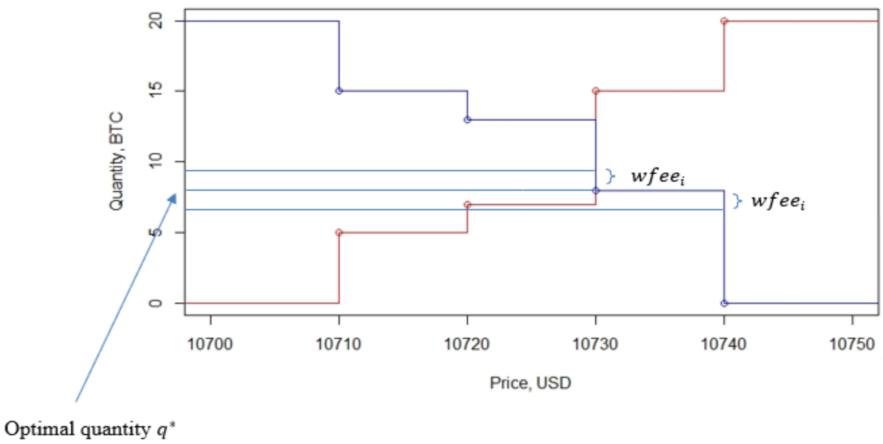


Figure 20. Horizontal Tangency

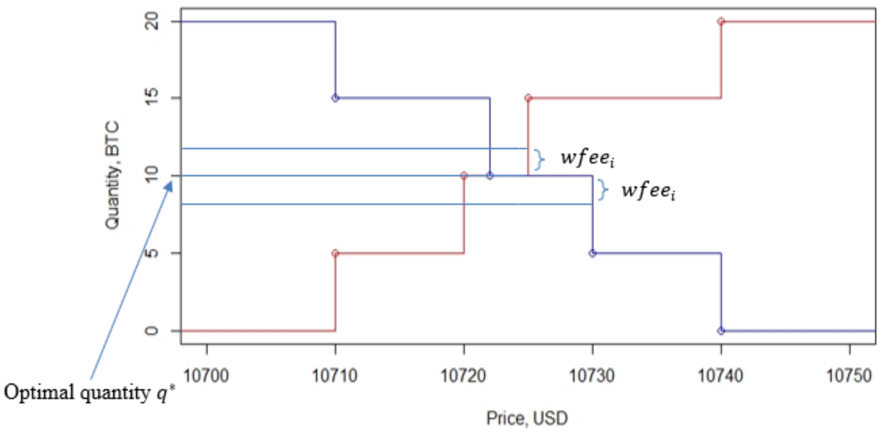


Figure 21. Arbitrage strategy

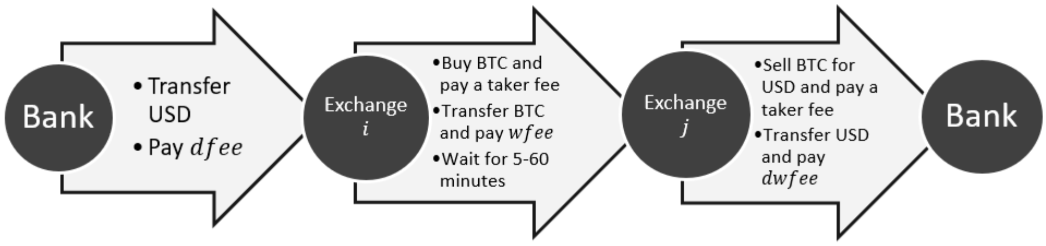
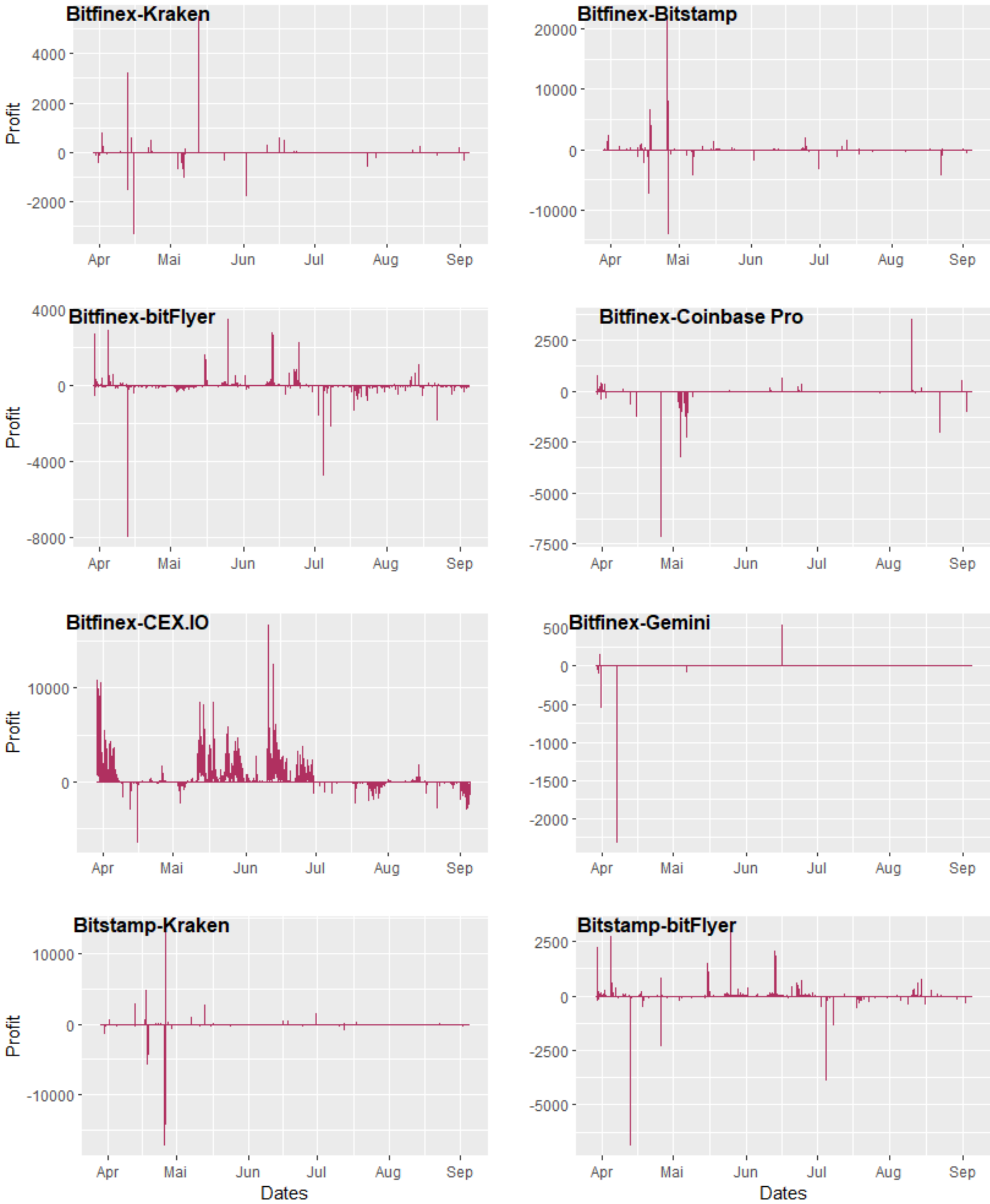
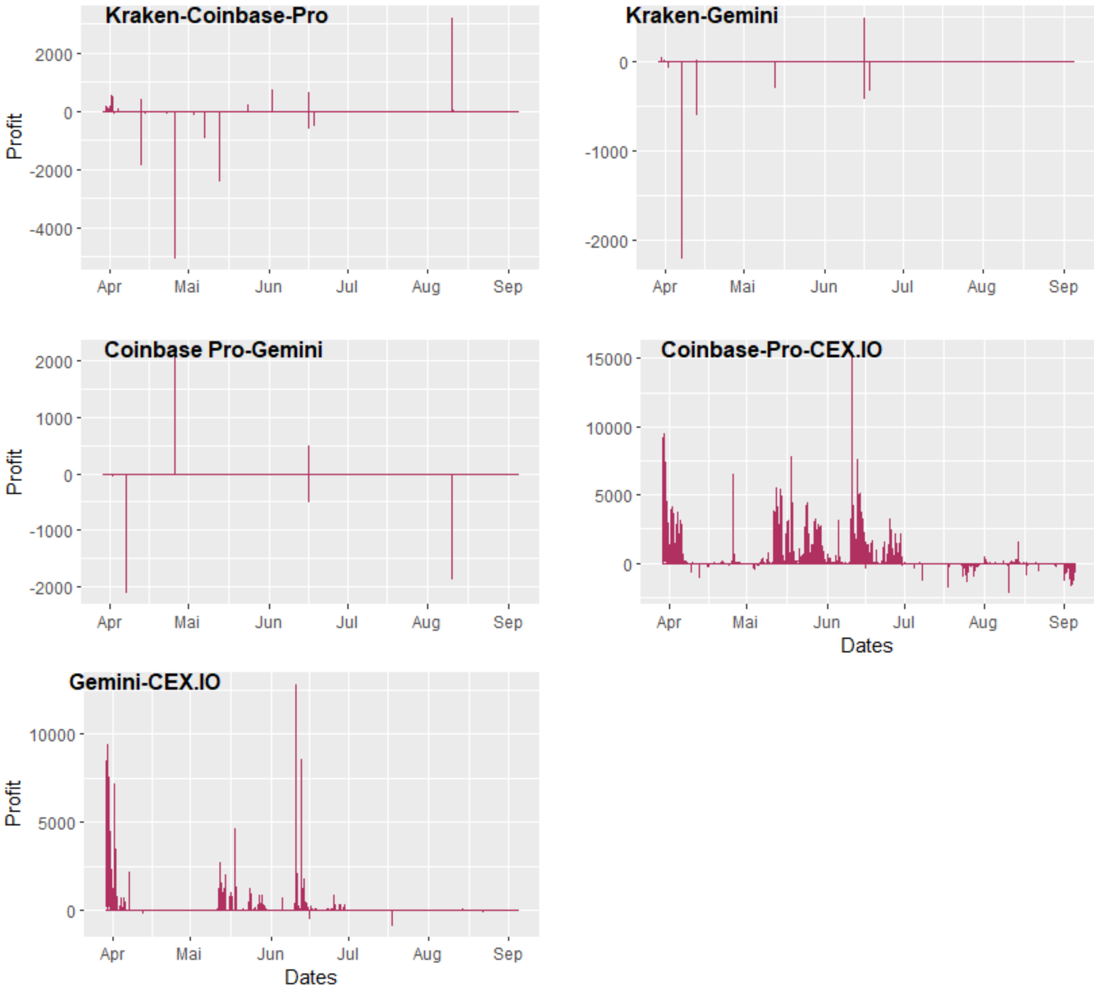


Figure 22. Arbitrage Profits over Time

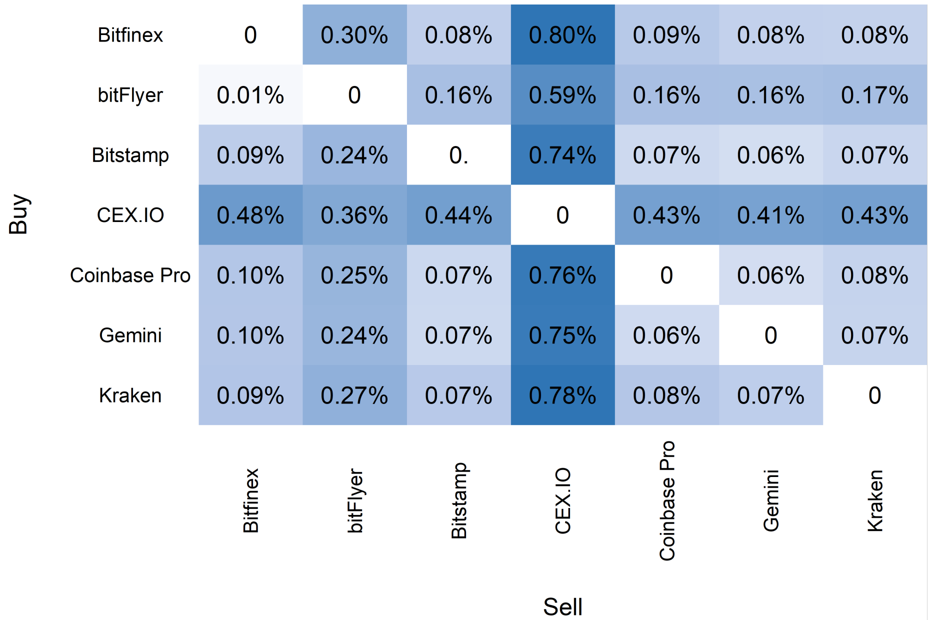



Bitstamp-Coinbase Pro
Bitstamp-Gemini
Bitstamp-CEX.IO
bitFlyer-Gemini
bitFlyer-Kraken
bitFlyer-CEX.IO
bitFlyer-Coinbase Pro
Kraken-CEX.IO
Profit
Dates
Apr
Mai
Jun
Jul
Aug
Sep



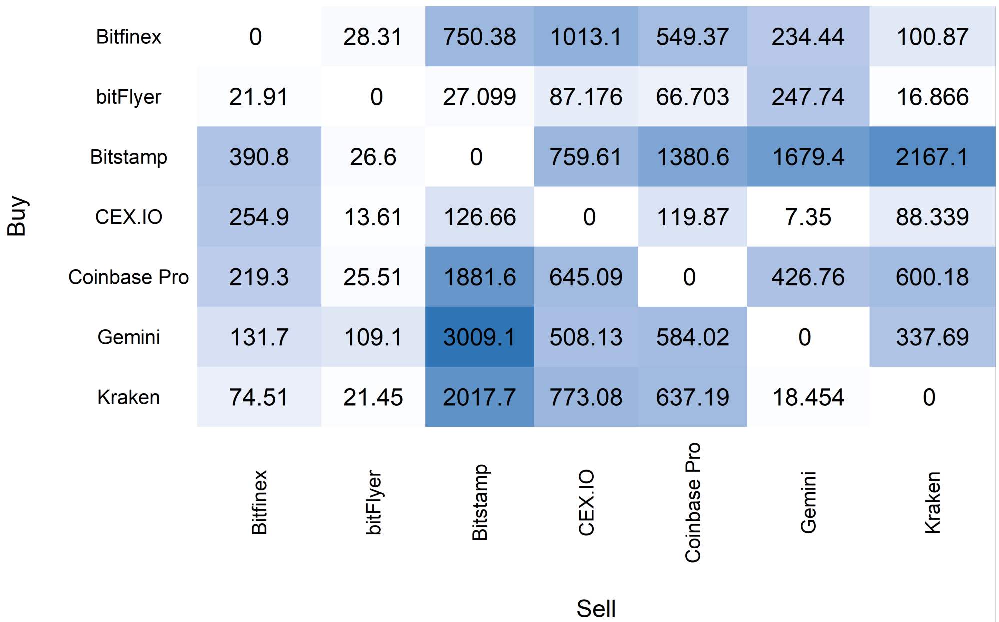
Notes: These figures show minute-level arbitrage profits $\pi_t^{ij}(q^*, wfee_i, dwfee_{ij})$, defined in Equation 47, over the time period from 02:00:00, March 29, 2018 to 01:00:00, September 5, 2018, by exchange pair ij. Each panel shows the arbitrage profits in both directions, where the upper part of the graph corresponds to the direction indicated in the title of the plot and the lower part of the graph corresponds to the opposite direction. The profits are reported in USD.

Figure 23. Time-series means of the exchange pair-specific positive first level price differences



Notes: The heatmap shows the time series means of the positive first level price differences $\delta_{t,1}^{ij}$ in %, defined in Equation 18 and normalized by the average of midquotes of corresponding exchanges, over the time period from 02:00:00, March 29, 2018 to 01:00:00, September 5, 2018, by exchange pair ij. The darker the colour, the higher the value.

Figure 24. Time-series means of the Profit per Arbitrage Opportunity



Notes: The heatmap shows the time series means of the profit per arbitrage opportunity (when $\delta_{t,1}^{ij} > 0$), $\pi_t^{ij}(q^*, wfee_i, dwfee_{ij})$, defined in Equation 47, over the time period from 02:00:00, March 29, 2018 to 01:00:00, September 5, 2018, by exchange pair ij. The darker the colour, the higher the value.

Figure 25. Summary statistics of Arbitrage Profits, Price Differences and Optimal Quantity

Pair	Arbitrage Profit		Price Difference			Optimal Quantity	
	Mean	Max	Mean	StDev	Max	Mean	Max
Bitfinex-Kraken	0.09	5530.53	0.03	0.06	3.63	0.0060	68.29
Bitfinex-CEX.IO	356.83	16759.23	0.48	0.64	5.74	9.2057	112.21
Bitfinex-Bitstamp	2.76	22120.96	0.03	0.08	3.66	0.0528	85.29
Bitfinex-bitFlyer	2.96	3500.22	0.11	0.22	3.51	0.1579	65.42
Bitfinex-Coinbase Pro	2.14	3528.46	0.04	0.09	4.06	0.0479	60.65
Bitfinex-Gemini	0.01	541.92	0.03	0.06	4.12	0.0001	9.37
Kraken-Bitfinex	0.62	3289.76	0.04	0.08	1.96	0.0614	63.94
Kraken-CEX.IO	257.90	14699.65	0.47	0.62	5.81	7.6645	94.58
Kraken-Bitstamp	2.34	17137.51	0.02	0.07	3.75	0.0240	94.44
Kraken-bitFlyer	1.55	2587.37	0.10	0.20	3.62	0.0807	65.42
Kraken-Coinbase Pro	1.88	3214.30	0.03	0.09	3.99	0.0351	34.55
Kraken-Gemini	0.01	491.84	0.03	0.05	3.94	0.0004	7.30
Bitstamp-Bitfinex	3.59	13841.31	0.04	0.09	3.13	0.1082	121.56
Bitstamp-CEX.IO	253.89	15909.73	0.45	0.59	5.65	7.8672	109.58
Bitstamp-Kraken	2.69	13656.69	0.02	0.07	3.44	0.0361	73.68
Bitstamp-bitFlyer	1.42	3114.65	0.08	0.18	3.54	0.0628	65.42
Bitstamp-Coinbase Pro	3.56	12208.07	0.02	0.08	3.92	0.0499	66.10
Bitstamp-Gemini	0.73	6491.29	0.01	0.06	4.10	0.0159	66.10
CEX.IO-Bitfinex	33.86	6362.56	0.14	0.30	2.91	1.7299	76.99
CEX.IO-Kraken	8.70	7375.46	0.12	0.26	2.96	0.6995	62.40
CEX.IO-Bitstamp	12.29	14184.35	0.12	0.26	3.42	0.8341	77.37
CEX.IO-bitFlyer	1.01	1067.19	0.07	0.19	2.07	0.0932	12.97
CEX.IO-Coinbase Pro	10.80	2146.36	0.12	0.27	2.93	0.7253	61.70
CEX.IO-Gemini	0.02	818.08	0.11	0.25	3.95	0.0014	19.85
Gemini-Bitfinex	0.02	2303.32	0.05	0.09	3.70	0.0010	34.69
Gemini-Bitstamp	1.36	12310.89	0.02	0.07	3.85	0.0222	80.18
Gemini-bitFlyer	0.40	2799.16	0.08	0.18	3.52	0.0116	25.79
Gemini-Coinbase Pro	0.50	2100.04	0.03	0.07	4.00	0.0162	26.57
Gemini-Kraken	0.02	2190.52	0.03	0.06	3.70	0.0005	25.02
Gemini-CEX.IO	70.58	12821.18	0.46	0.60	5.76	1.9803	121.59
Coinbase Pro-Bitfinex	2.72	7130.51	0.06	0.10	2.49	0.1646	86.84
Coinbase Pro-CEX.IO	70.58	12821.18	0.47	0.61	5.60	1.9803	121.59
Coinbase Pro-Kraken	0.41	5027.49	0.03	0.07	3.53	0.0091	50.99
Coinbase Pro-bitFlyer	1.40	3434.37	0.09	0.19	3.98	0.0625	60.39
Coinbase Pro-Bitstamp	1.87	21805.75	0.02	0.08	5.87	0.0197	80.18
Coinbase Pro-Gemini	0.06	2140.09	0.02	0.05	4.00	0.0024	51.39
bitFlyer-Bitfinex	1.48	7927.44	0.08	0.16	5.71	0.1175	58.37
bitFlyer-CEX.IO	26.53	8129.22	0.33	0.49	5.06	0.8664	53.51
bitFlyer-Kraken	0.44	3569.70	0.06	0.12	4.74	0.0431	45.02
bitFlyer-Bitstamp	0.59	6844.70	0.05	0.12	5.49	0.0364	68.78
bitFlyer-Coinbase Pro	1.15	6687.35	0.05	0.12	5.27	0.0487	48.26
bitFlyer-Gemini	0.09	5202.70	0.05	0.11	5.61	0.0015	34.31

Notes: The heatmap shows summary statistics of the minute-level arbitrage profits in USD, defined in Equation 47, with corresponding first level price differences in %, defined in Equation 18 and optimal quantities in BTC by exchange pair over the time period from 02:00:00, March 29, 2018 to 01:00:00, September 5, 2018. The columns present the mean, the standard deviation, and the maximum. The darker the colour, the higher the value.

Figure 26. Autocorrelation functions of the hourly price differences for each exchange (up to 100 lags)

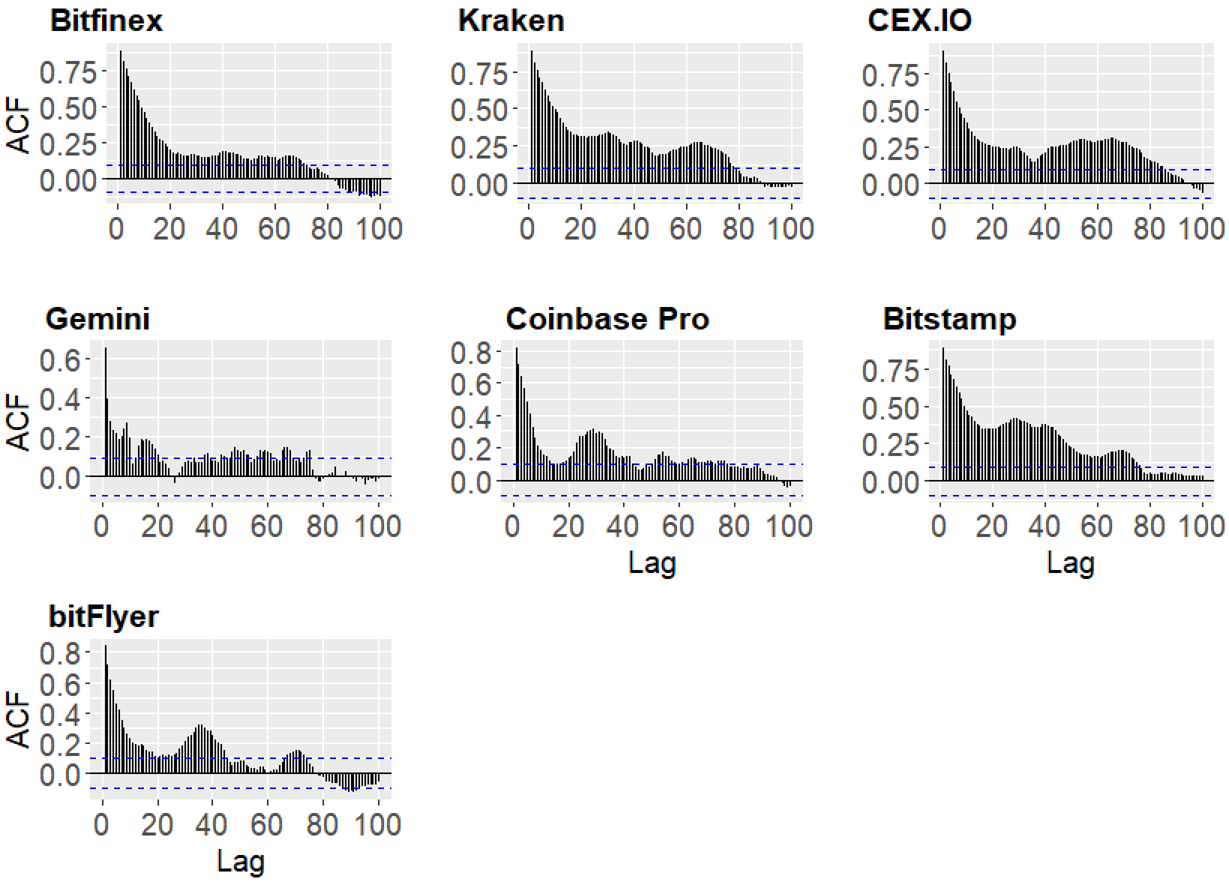


Figure 27. Partial autocorrelation functions of the hourly price differences for each exchange (up to 100 lags)

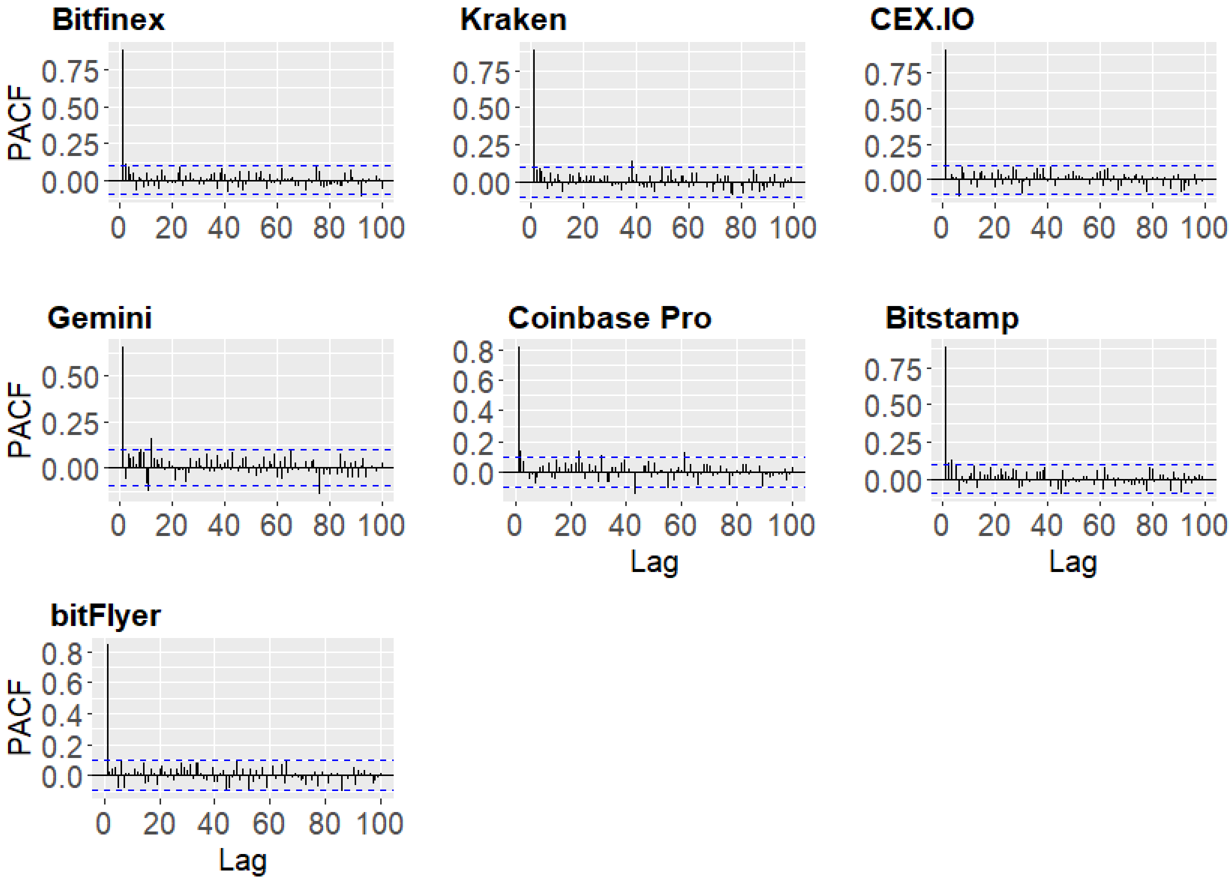


Figure 28. Autocorrelation functions of the daily price differences for each exchange (up to 20 lags)

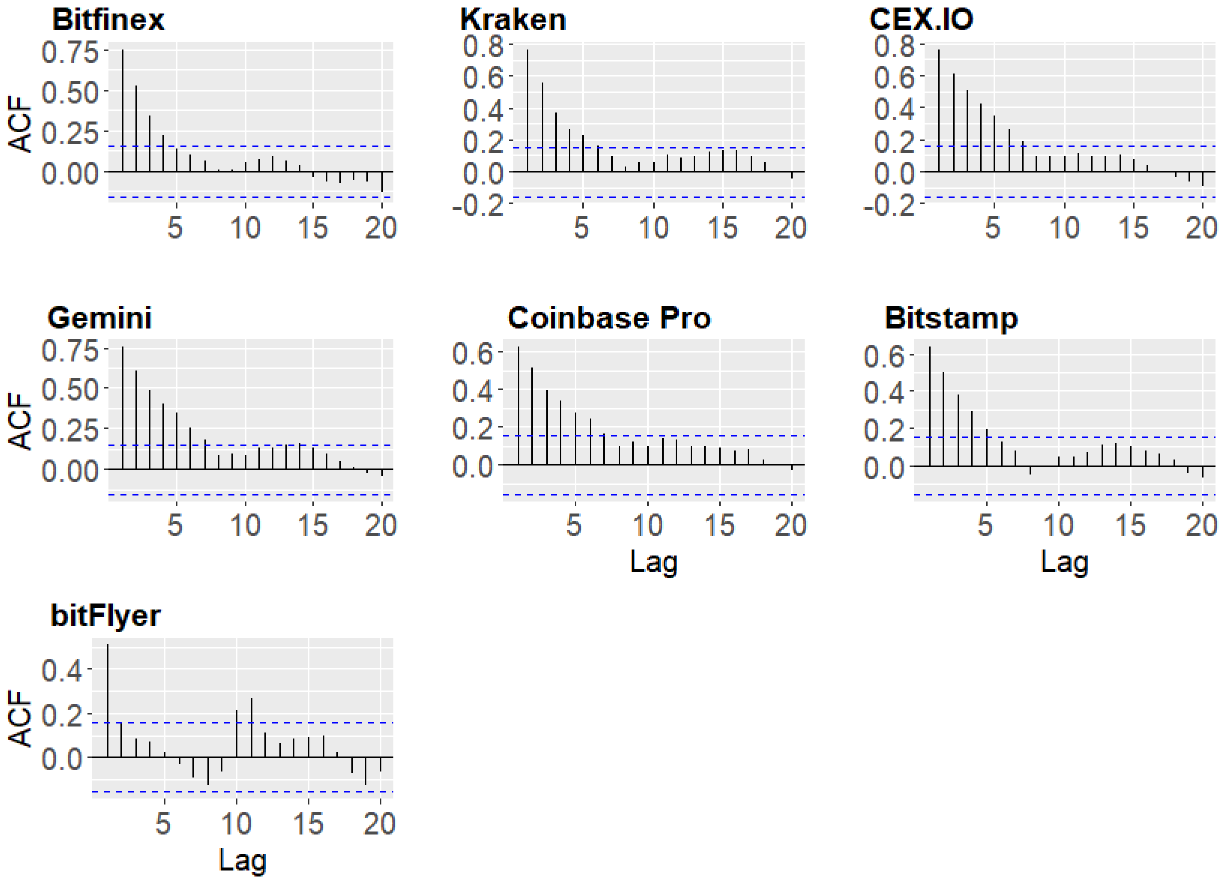


Figure 29. Partial autocorrelation functions of the daily price differences for each exchange (up to 20 lags)

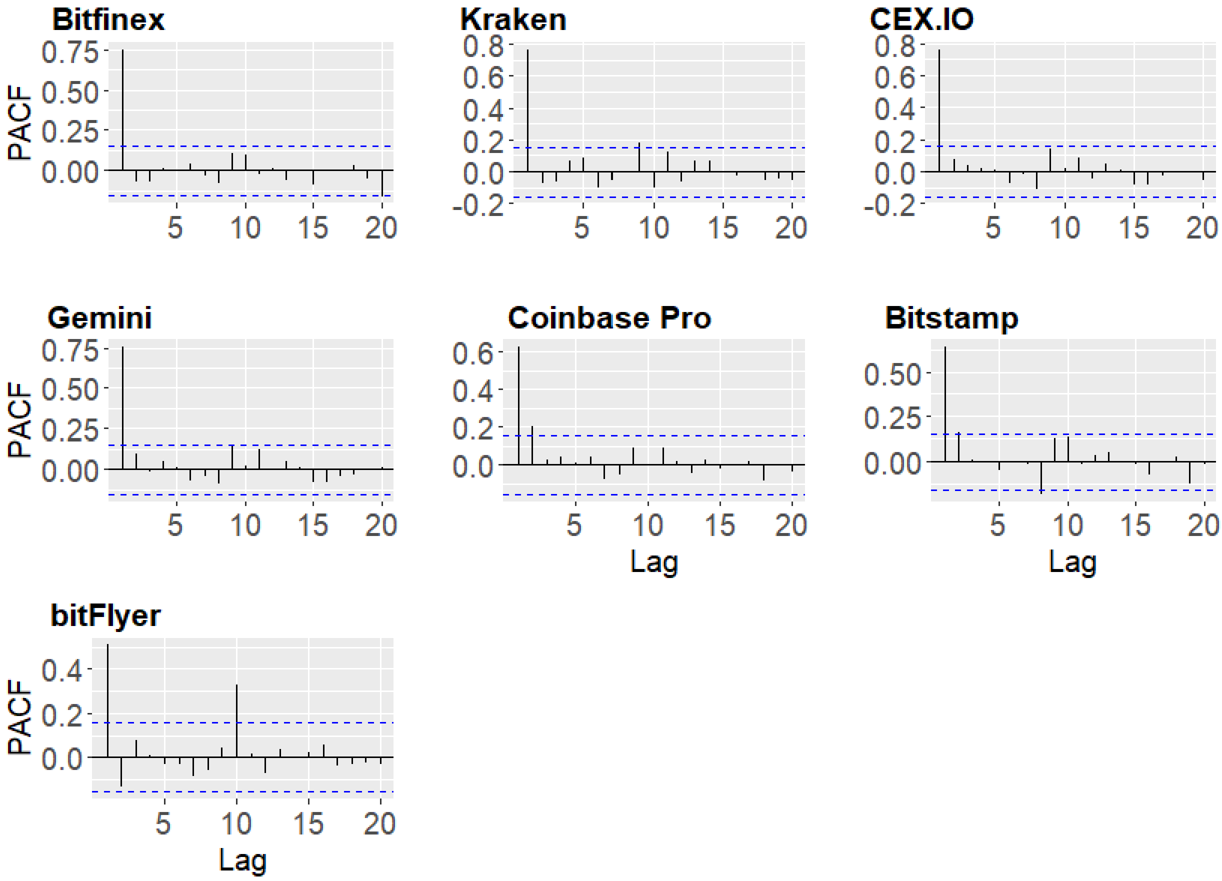
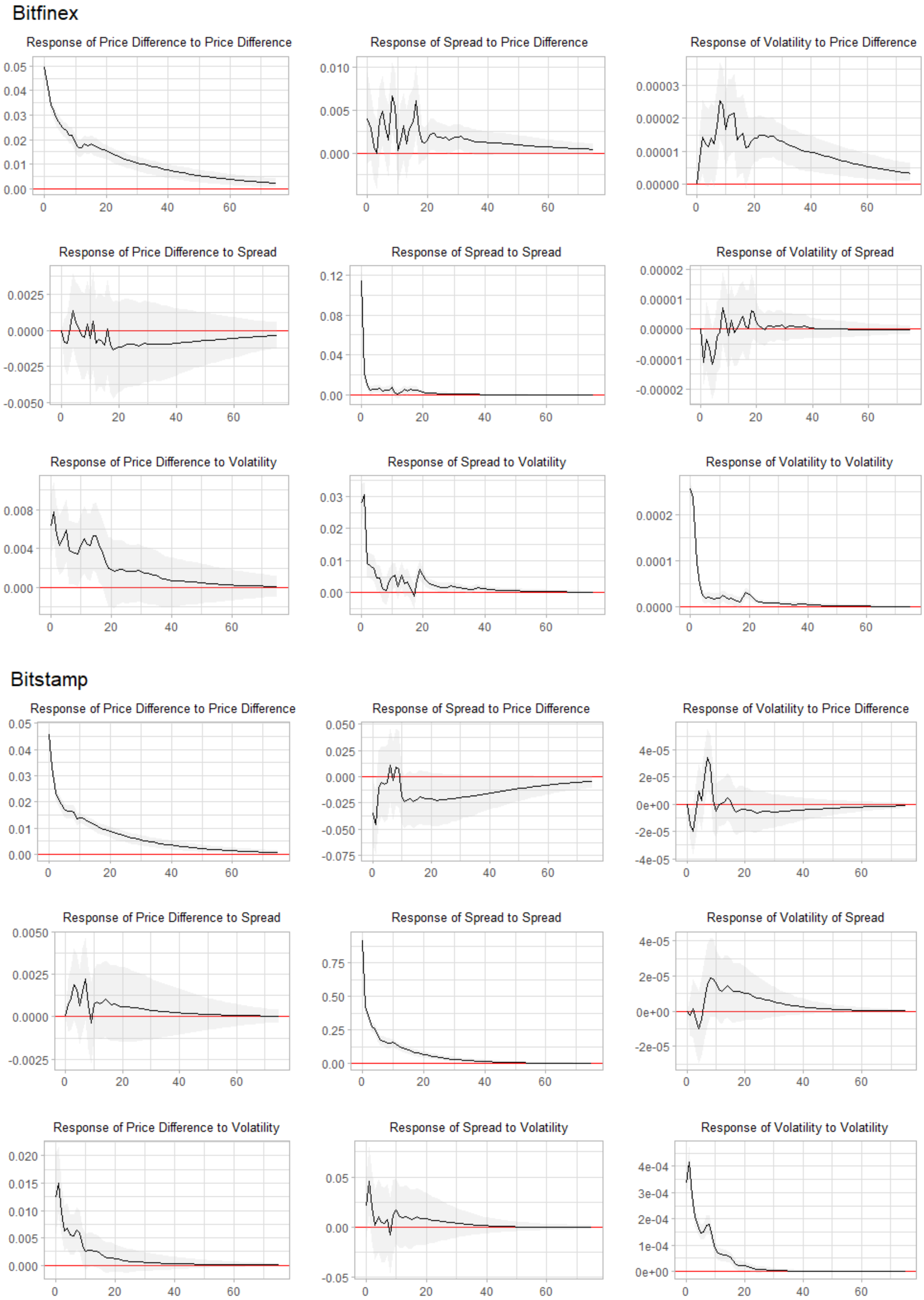
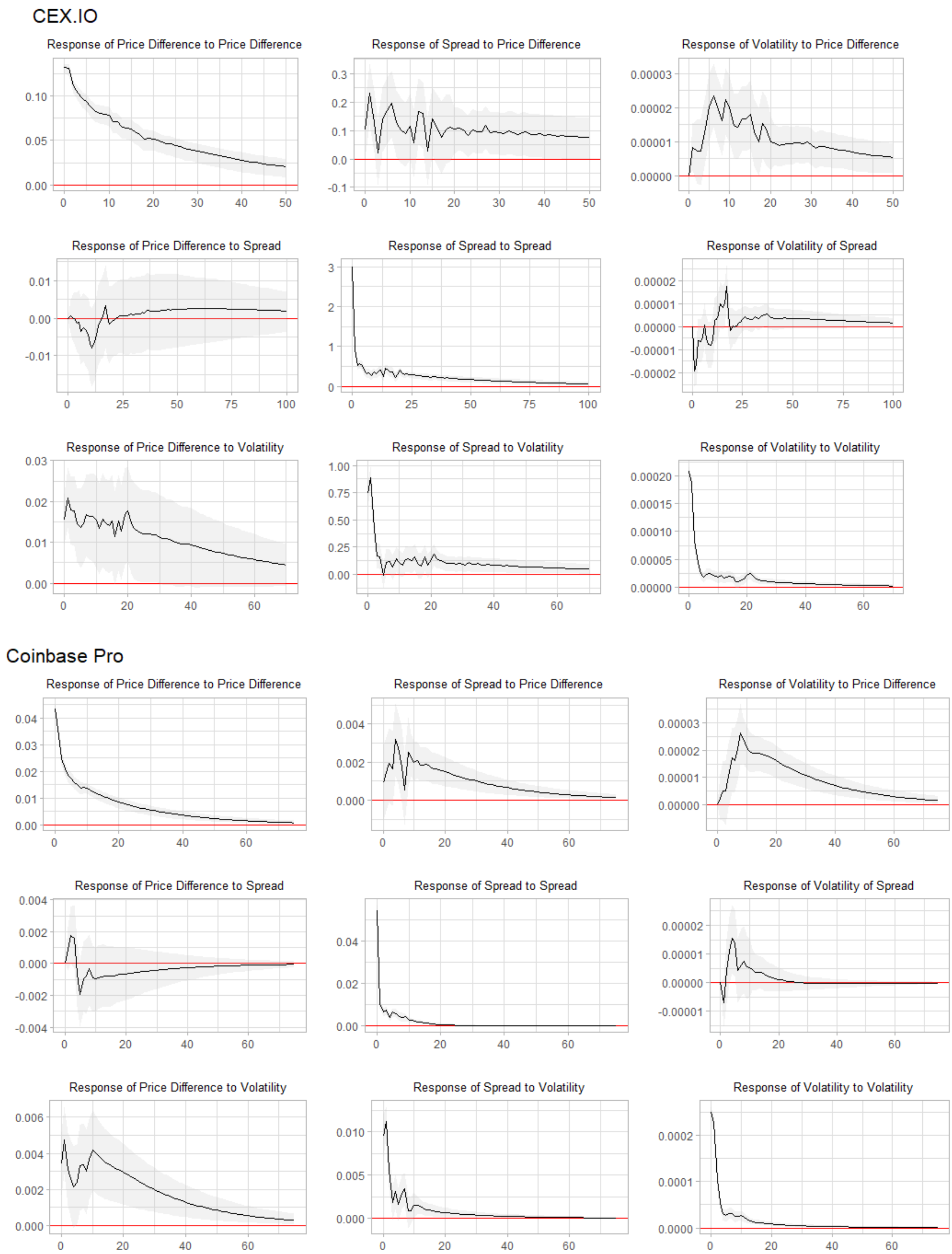


Figure 30. Impulse response functions of Bitfinex and Bitstamp for the VAR with the hourly price differences, quoted bid-ask spread and volatility



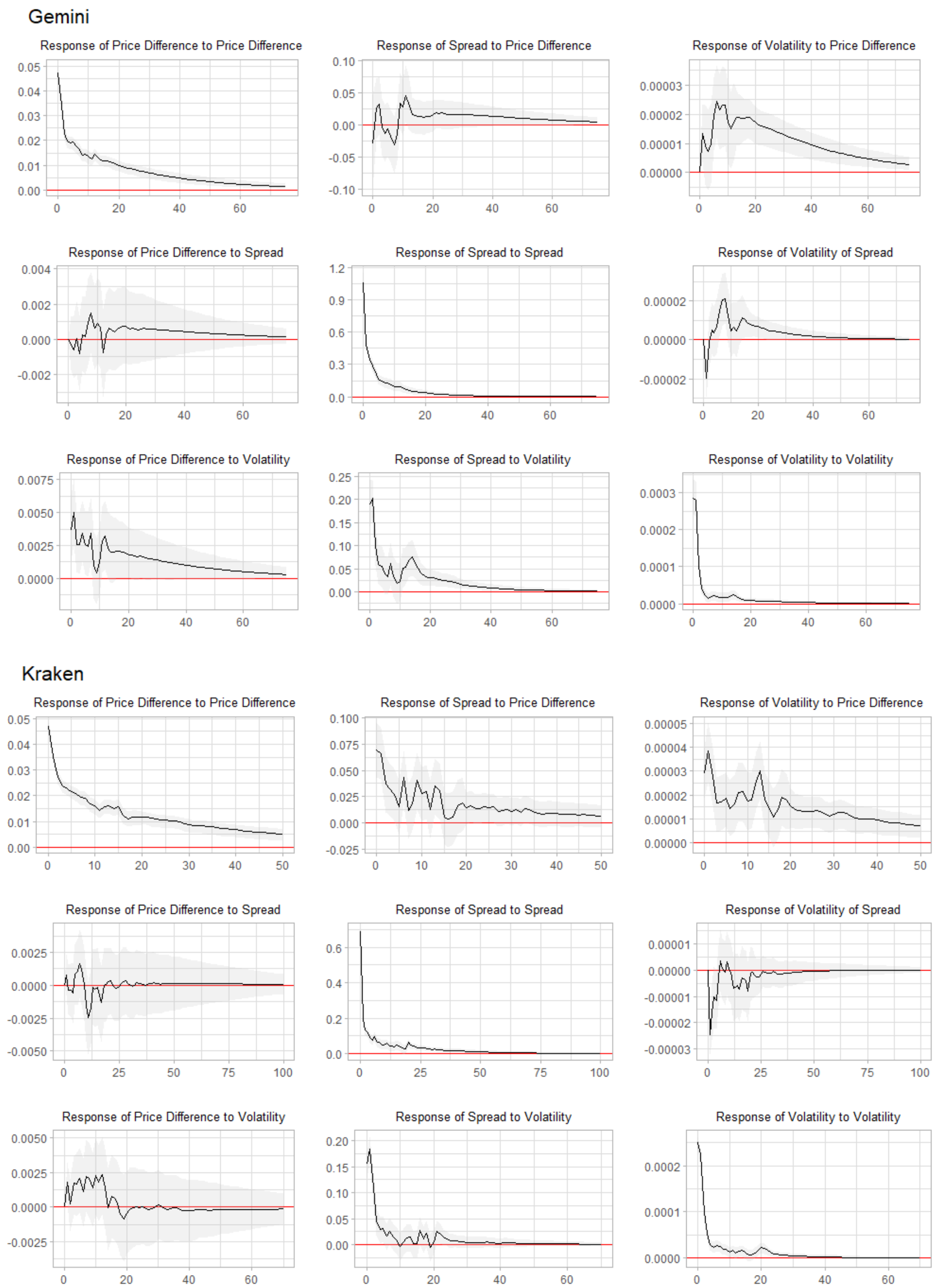
Notes: Impulse responses are presented for the vector autoregression that includes three endogenous variables: hourly price differences, quoted bid-ask spread and volatility on Bitfinex over the time period from March 29, 2018 to September 5, 2018 (all variables are adjusted for time trends and other regularities, see Section 6.1). Bootstrapped confidence intervals (0.95%) are provided along with the estimated responses.

Figure 31. Impulse response functions of CEX.IO and Coinbase Pro for the VAR with the hourly price differences, quoted bid-ask spread and volatility



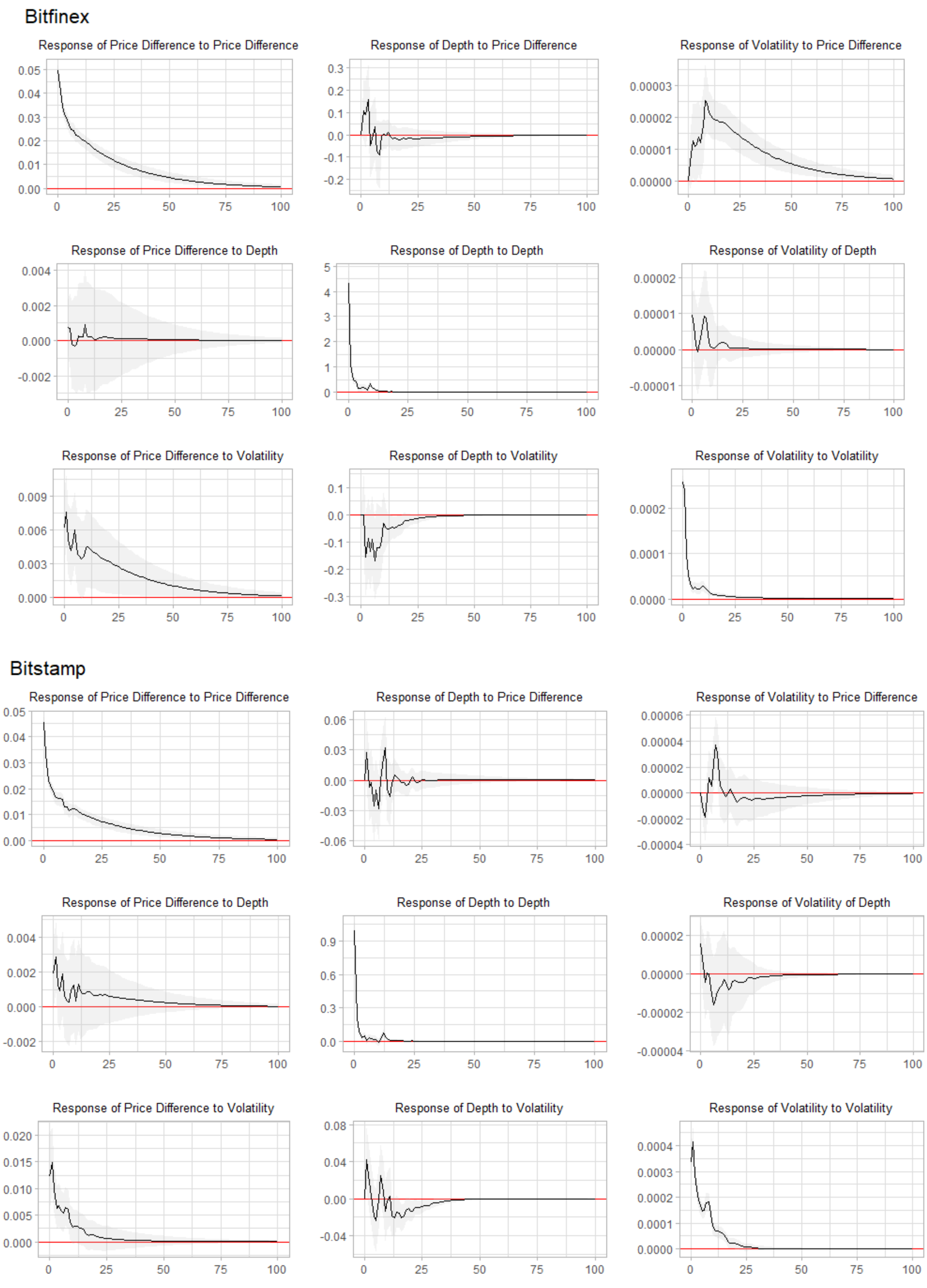
Notes: See the notes for Figure 30

Figure 32. Impulse response functions of Gemini and Kraken for the VAR with the hourly price differences, quoted bid-ask spread and volatility



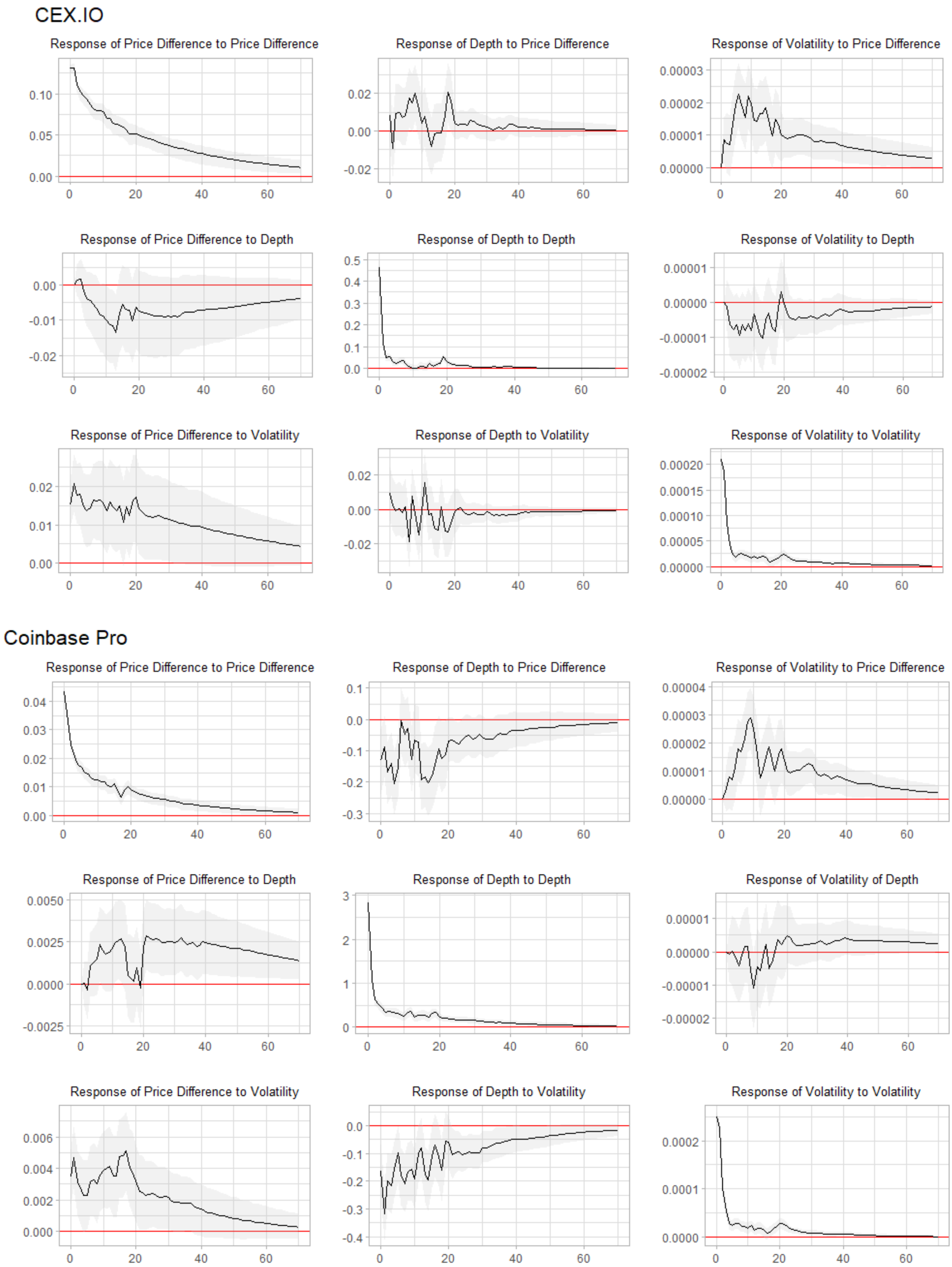
Notes: See the notes for Figure 30

Figure 33. Impulse response functions of Bitfinex and Bitstamp for the VAR with the hourly price differences, quoted depth and volatility



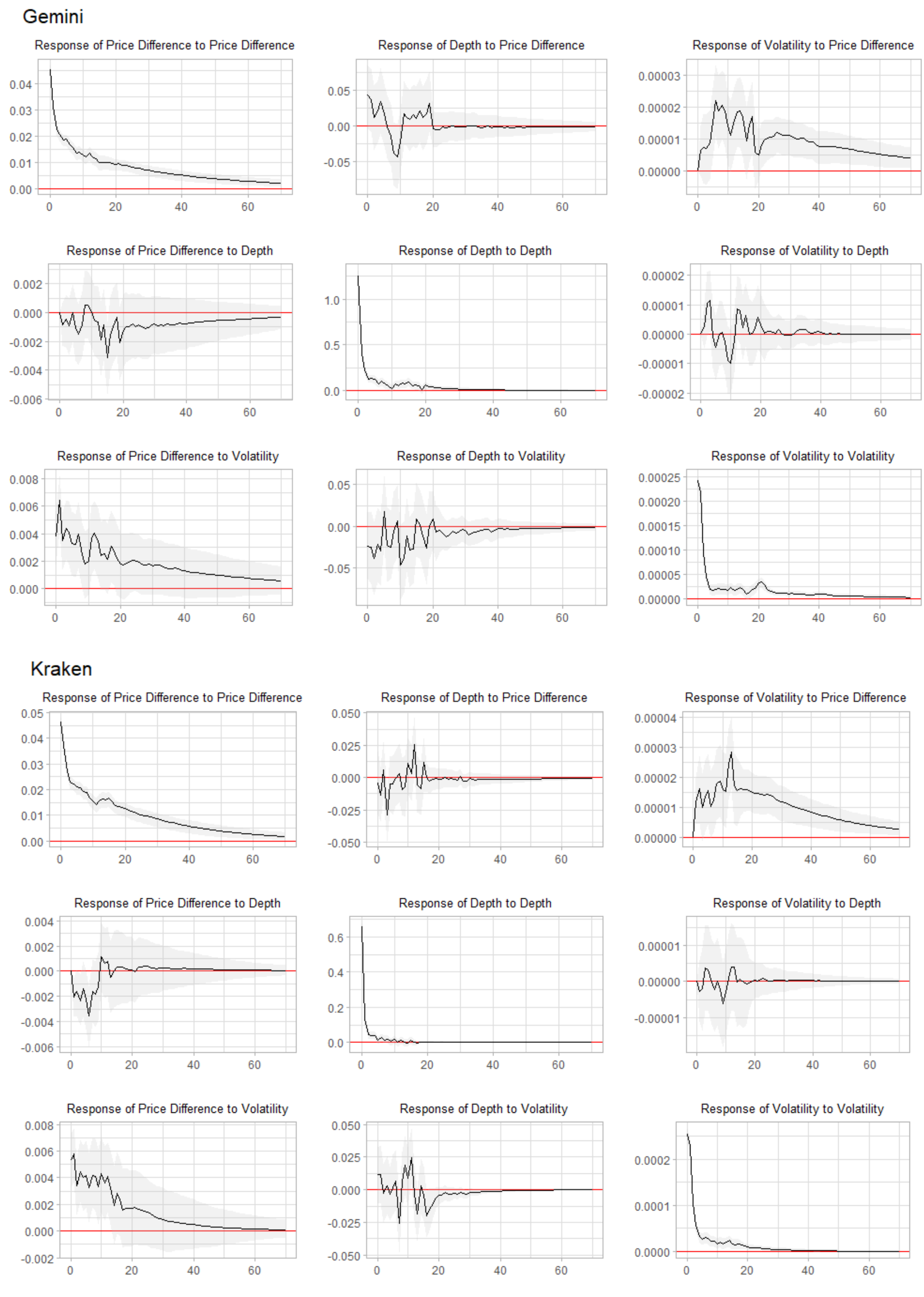
Notes: Impulse responses are presented for the vector autoregression that includes three endogenous variables: hourly price differences, quoted depth and volatility on Bitfinex over the time period from March 29, 2018 to September 5, 2018 (all variables are adjusted for time trends and other regularities, see Section 6.1). Bootstrapped confidence intervals (0.95%) are provided along with the estimated responses.

Figure 34. Impulse response functions of CEX.IO and Coinbase Pro for the VAR with the hourly price differences, quoted depth and volatility



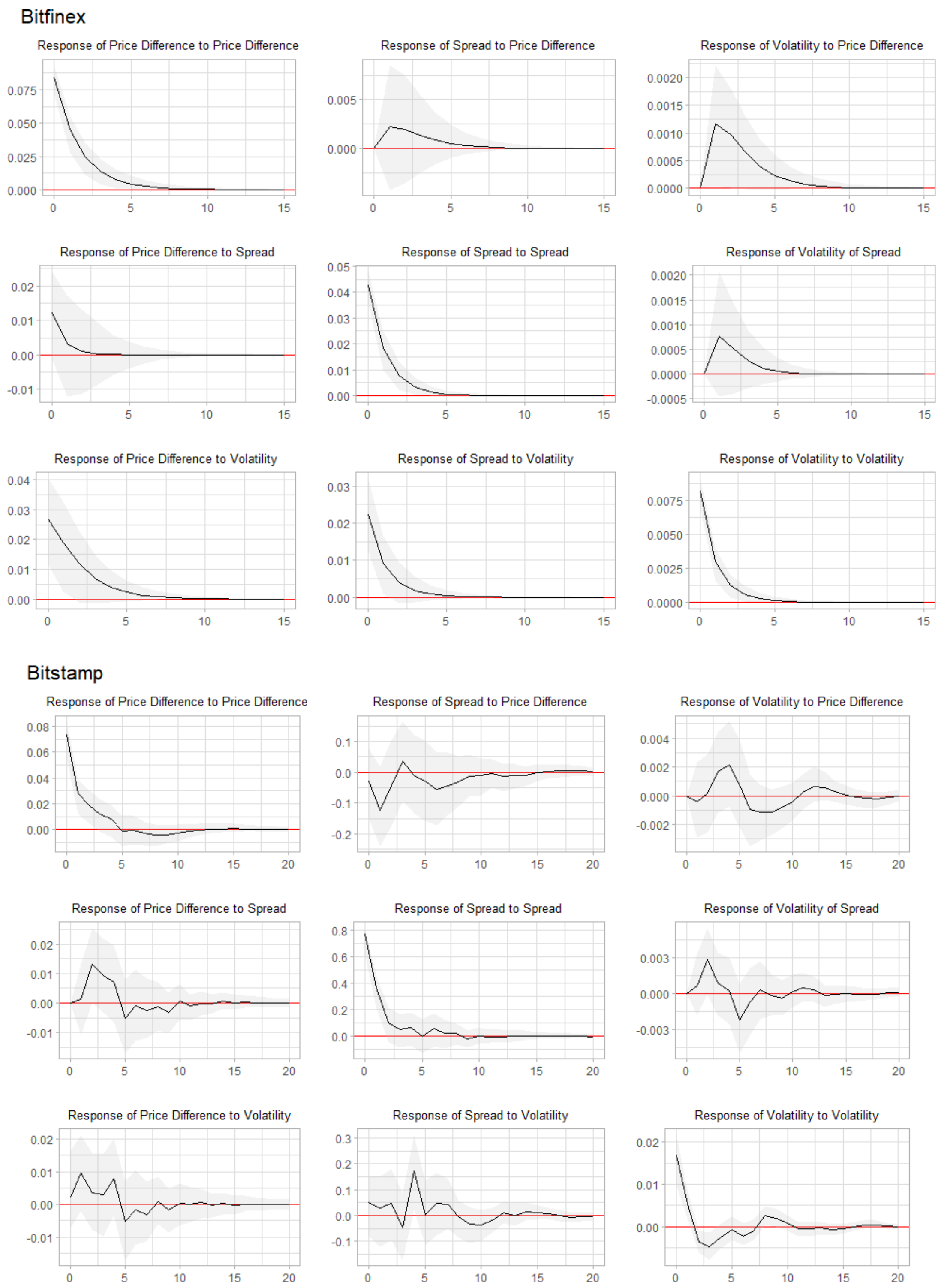
Notes: See the notes for Figure 33

Figure 35. Impulse response functions of Gemini and Kraken for the VAR with the hourly price differences, quoted depth and volatility



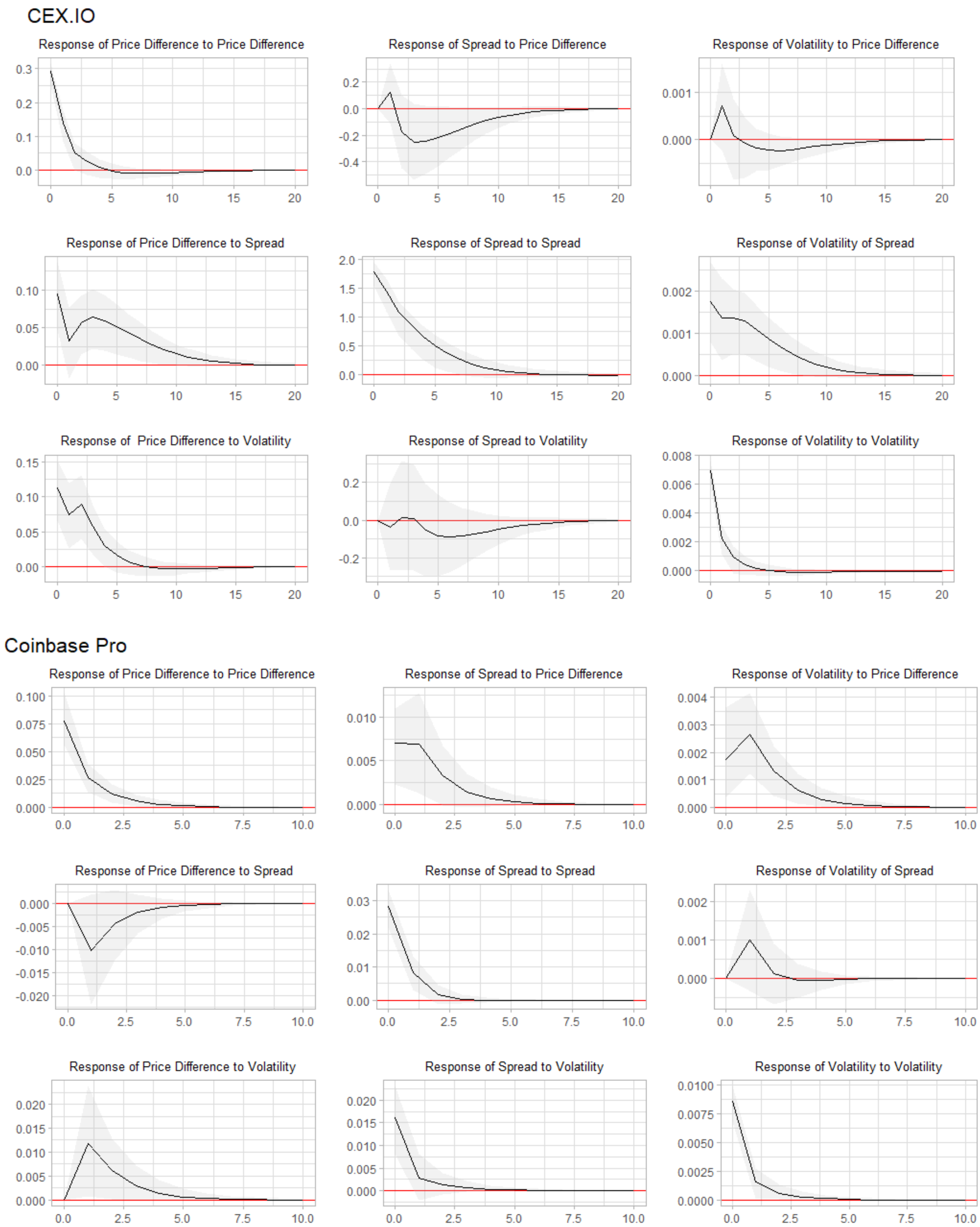
Notes: See the notes for Figure 33

Figure 36. Impulse response functions of Bitfinex and Bitstamp for the VAR with the daily price differences, quoted bid-ask spread and volatility



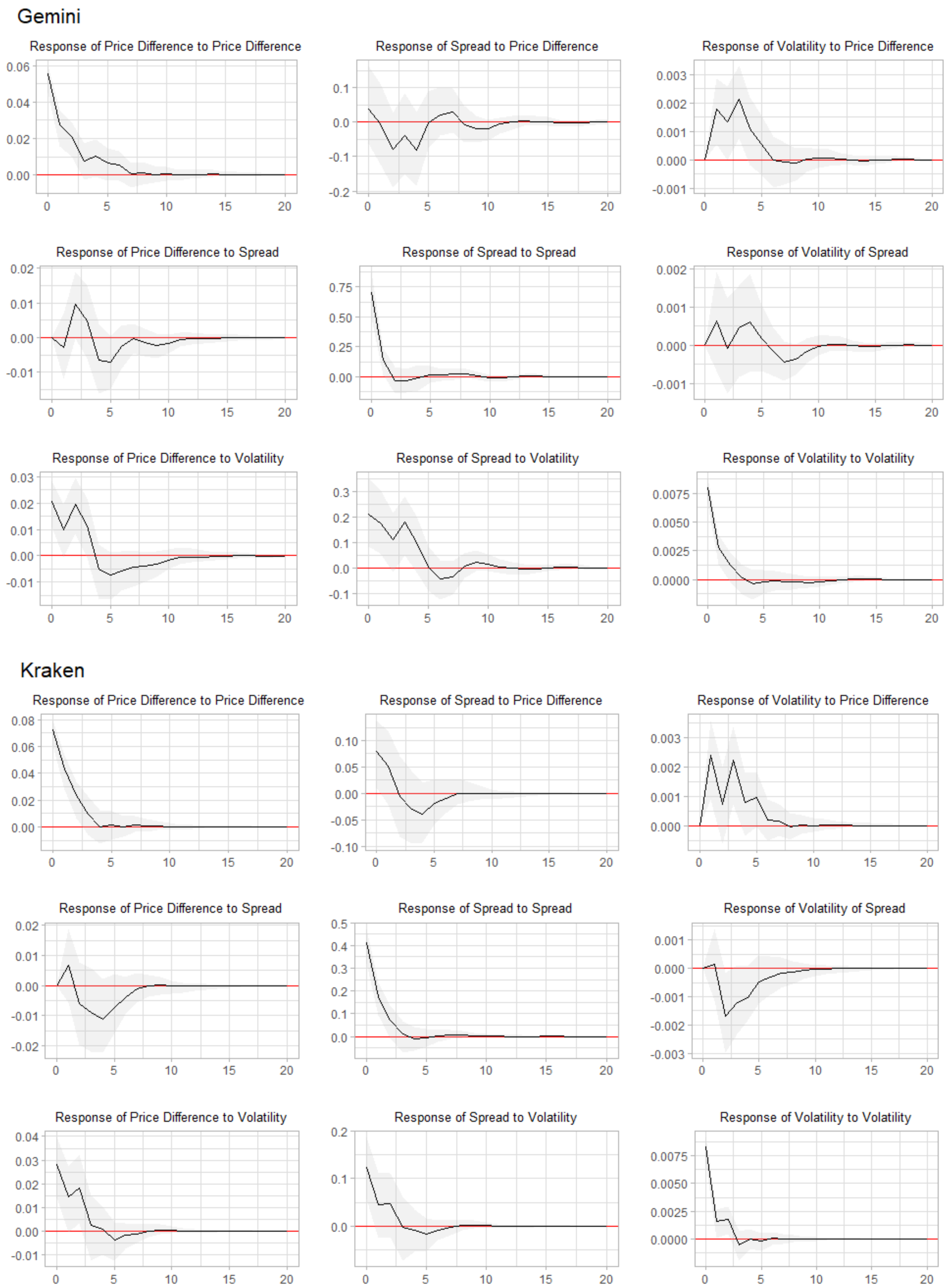
Notes: Impulse responses are presented for the vector autoregression that includes three endogenous variables: daily price differences, quoted bid-ask spread and volatility on Bitfinex over the time period from March 29, 2018 to September 5, 2018 (all variables are adjusted for time trends and other regularities, see Section 6.1). Bootstrapped confidence intervals (0.95%) are provided along with the estimated responses.

Figure 37. Impulse response functions of CEX.IO and Coinbase Pro for the VAR with the daily price differences, quoted bid-ask spread and volatility



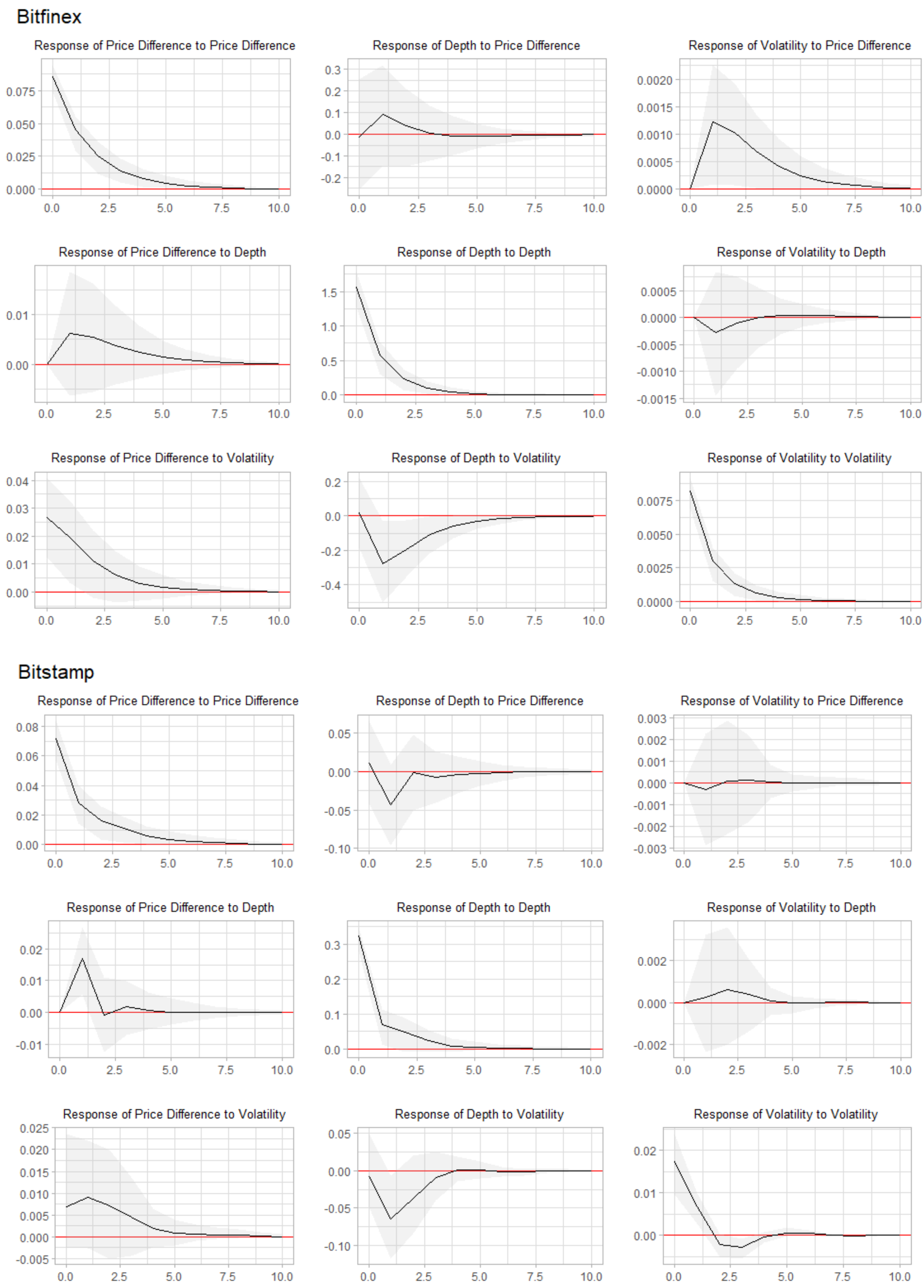
Notes: See the notes for Figure 36

Figure 38. Impulse response functions of Gemini and Kraken for the VAR with the daily price differences, quoted bid-ask spread and volatility



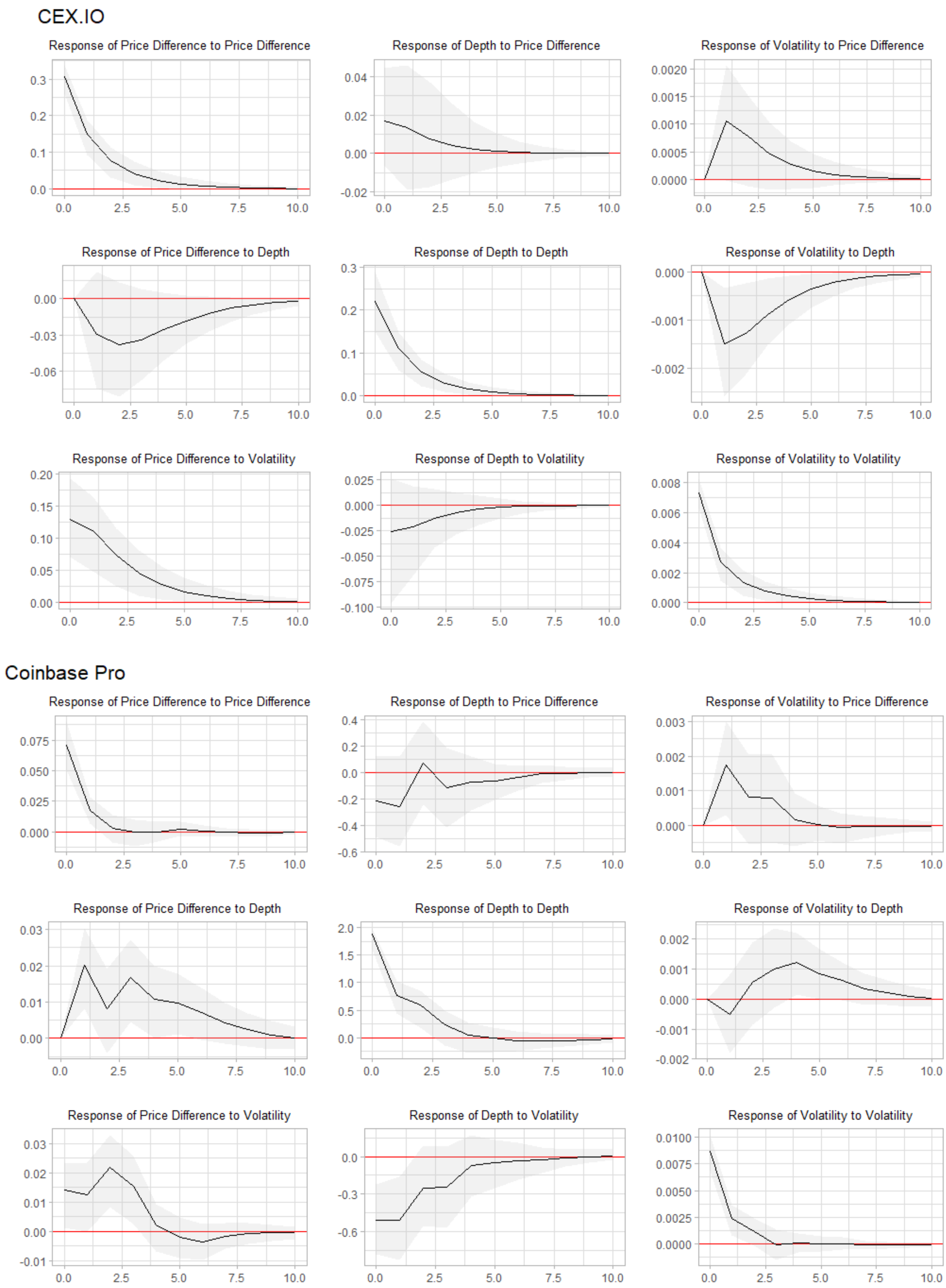
Notes: See the notes for Figure 36

Figure 39. Impulse response functions of Bitfinex and Bitstamp for the VAR with the daily price differences, quoted depth and volatility



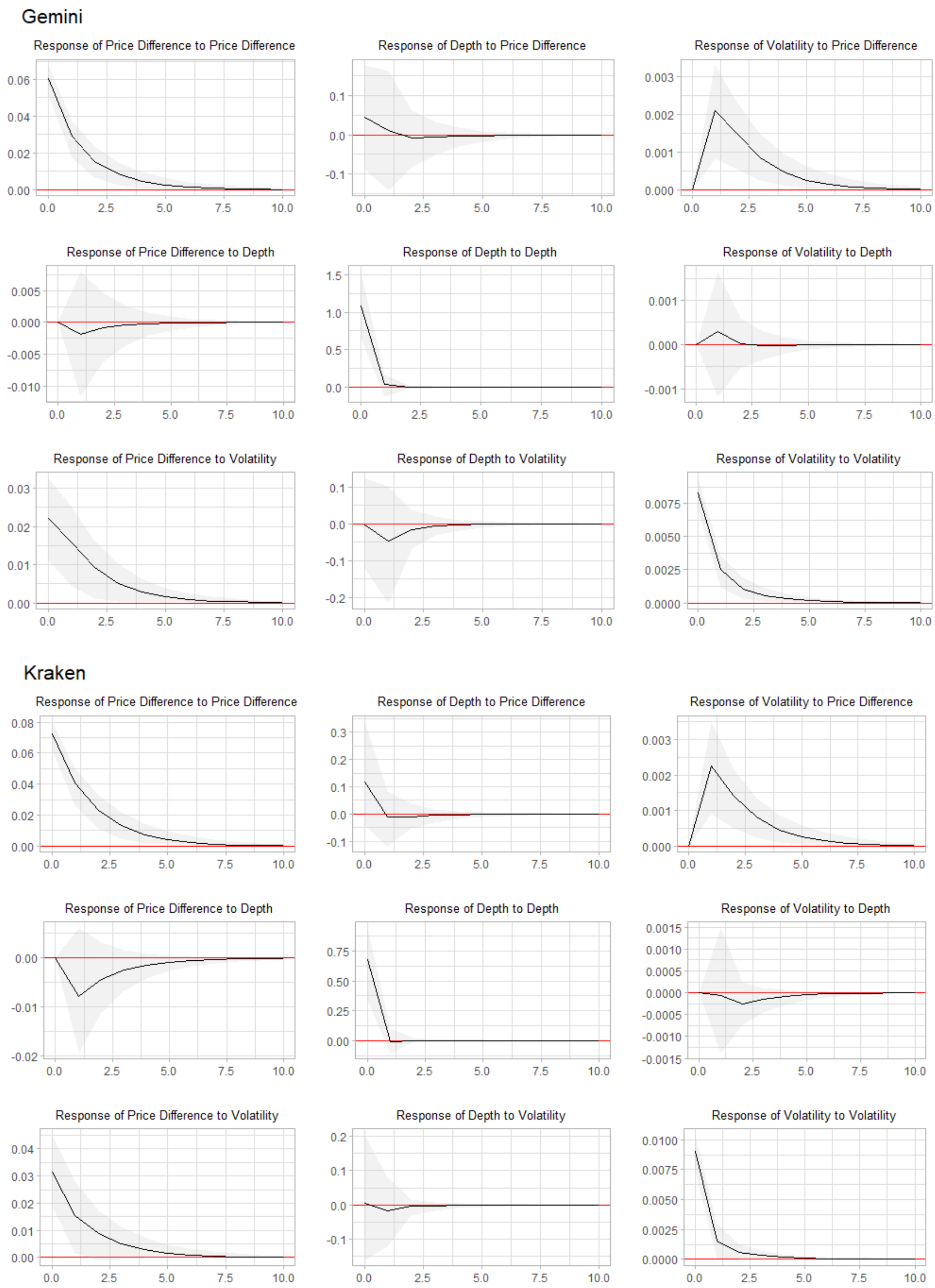
Notes: Impulse responses are presented for the vector autoregression that includes three endogenous variables: daily price differences, quoted depth and volatility on Bitfinex over the time period from March 29, 2018 to September 5, 2018 (all variables are adjusted for time trends and other regularities, see Section 6.1). Bootstrapped confidence intervals (0.95%) are provided along with the estimated responses.

Figure 40. Impulse response functions of CEX.IO and Coinbase Pro for the VAR with the daily price differences, quoted depth and volatility



Notes: See the notes for Figure 39

Figure 41. Impulse response functions of Gemini and Kraken for the VAR with the daily price differences, quoted depth and volatility



Notes: See the notes for Figure 39

8.2 Tables

Table 7. Augmented Dickey-Fuller and Phillips-Perron Test Results for the hourly series

	Augmented Dickey-Fuller				Phillips–Perron			
Exchange	PD	PS	$AvgD$	σ^2	PD	PS	$AvgD$	σ^2
adjusted time series, with drift, no trend								
Bitfinex (28, 26, 25, 24)	-2.38***	-2.70***	-7.19***	-3.10***	-58.23***	-55.84***	-26.78***	-6.52***
bitFlyer (28, 25, 25, 24)	-4.78***	-6.60***	-7.05***	-3.62***	-44.99***	-41.51***	-28.15***	-5.68***
Bitstamp (24, 29, 28, 24)	-4.08***	-7.04***	-7.51***	-4.55***	-23.50***	-55.93***	-15.69***	-5.92***
CEX.IO (25, 29, 26, 24)	-3.89***	-6.24***	-5.31***	-3.24***	-43.61***	-49.48***	-25.74***	-6.54***
Coinbase Pro (27, 23, 25, 24)	-5.53***	-5.32***	-7.34***	-3.36***	-60.83***	-36.22***	-25.74***	-6.83***
Gemini (23, 29, 29, 24)	-4.94***	-6.46***	-7.04***	-2.90***	-35.01***	-48.89***	-28.91***	-6.62***
Kraken (23, 29, 19, 24)	-4.76***	-9.05***	-8.11***	-4.90***	-42.78***	-48.54***	-24.46***	-6.89***
not adjusted time series, with drift and trend								
Bitfinex (9, 26, 25, 24)	-5.66***	-6.14***	-7.12***	-4.25***	-59.31***	-54.21***	-27.39***	-6.52***
bitFlyer (28, 25, 25, 24)	-5.23***	-7.71***	-7.34***	-3.64***	-43.33***	-41.50***	-28.31***	-6.89***
Bitstamp (24, 29, 28, 24)	-4.08***	-7.01***	-7.63***	-4.64***	-25.83***	-54.58***	-16.04***	-6.54***
CEX.IO (28, 29, 23, 29)	-3.90***	-5.90***	-6.91***	-5.80***	-46.20***	-49.22***	-25.16***	-7.67***
Coinbase Pro (27, 23, 25, 24)	-7.40***	-6.47***	-7.51***	-4.56***	-60.33***	-36.29***	-26.02***	-6.82***
Gemini (23, 29, 29, 24)	-6.83***	-7.94***	-7.00***	-4.82***	-35.07***	-48.88***	-28.99***	-6.66***
Kraken (23, 29, 19, 29)	-5.22***	-9.35***	-8.04***	-1.61***	-45.41***	-48.40***	-24.99***	-7.08***

Notes: The table reports the Augmented Dickey Fuller and Phillips-Perron unit root tests for the hourly price differences, quoted spread, quoted depth and bitcoin price volatility, where the null hypothesis is that a time series is $I(1)$. The hourly price difference of exchange i at time t is defined as the cross-sectional hourly average of differences between the price level, calculated as a midquote (see Equation 6), on each sell-side exchange and the midquote prices on all other exchanges. Quoted spread is defined in Equation 8, average quoted depth (BTC) is defined in Equation 9. For the adjusted series, the test includes drift but no linear trend with respect to time. For the not adjusted series is the one with drift but no linear trend. The lag order is chosen according to Akaike's Information Criterion (reported in brackets for each exchange), while ensuring that sufficient lags are included in the testing regression such that the error term becomes white noise. For the Phillips-Perron test the statistics Z_τ is reported.

Table 8. Augmented Dickey-Fuller and Phillips-Perron Test Results for the daily series

	Augmented Dickey-Fuller				Phillips–Perron			
Exchange	PS	$AvgD$	σ^2	V	PS	$AvgD$	σ^2	V
adjusted time series, with drift, no trend								
Bitfinex (1, 1, 1, 9)	-6.73***	-5.24***	-6.63***	-1.03	-4.48***	-6.44***	-6.03***	-7.72***
bitFlyer (10, 8, 1, 6)	-2.10**	-1.80**	-6.13***	-1.91**	-7.52***	-9.42***	-6.89***	-6.08***
Bitstamp (7, 7, 8, 7)	-1.90**	-2.07**	-2.67***	-2.34***	-4.39***	-7.83***	-7.35***	-6.26***
CEX.IO (1, 1, 1, 5)	-3.62***	-3.34***	-5.96***	-2.29***	-4.01***	-3.58***	-5.82***	-6.94***
Coinbase Pro (7, 4, 1, 12)	-2.91***	-3.34***	-6.49***	-3.12***	-7.99***	-9.92***	-6.57***	-6.90***
Gemini (1, 8, 1, 5)	-5.60***	-3.35***	-6.87***	-2.53***	-7.48***	-7.06***	-6.94***	-7.15***
Kraken (6, 1, 1, 1)	-1.84**	-5.41***	-6.50***	-5.26***	-4.45***	-7.65***	-6.31***	-7.41***
not adjusted time series, with drift and trend								
Bitfinex (9, 1, 1, 11)	-7.03***	-6.20***	-6.97***	-0.68	-5.08***	-7.41***	-6.55***	-7.97***
bitFlyer (1, 8, 8, 2)	-6.87***	-1.75	-3.92***	-3.88***	-9.70***	-9.40***	-7.50***	-7.17***
Bitstamp (7, 8, 8, 8)	-2.58	-2.74	-3.28*	-3.60**	-5.07***	-9.05***	-7.77***	-7.00***
CEX.IO (1, 1, 1, 5)	-3.76**	-3.10*	-5.92***	-2.53	-4.24***	-3.40**	-5.88***	-8.34***
Coinbase Pro (7, 4, 1, 12)	-2.90	-3.34*	-6.72***	-3.16*	-8.00***	-9.90***	-7.00***	-8.14***
Gemini (1, 8, 1, 5)	-5.89***	-3.47**	-7.08***	-2.49	-7.74***	-7.08***	-7.35***	-6.90***
Kraken (1, 1, 1, 1)	-6.18***	-5.23***	-6.88***	-5.27***	-7.37***	-7.50***	-6.93***	-7.59***

Notes: The table reports the Augmented Dickey Fuller and Phillips-Perron unit root tests for the daily price differences, quoted spread, quoted depth and bitcoin price volatility, where the null hypothesis is that a time series is $I(1)$. The hourly price difference of exchange i at time t is defined as the cross-sectional daily average of differences between the price level, calculated as a midquote (see Equation 6), on each sell-side exchange and the midquote prices on all other exchanges. Quoted spread is defined in Equation 8, average quoted depth (BTC) is defined in Equation 9. For the adjusted series, the test includes drift but no linear trend with respect to time. For the not adjusted series is the one with drift but no linear trend. The lag order is chosen according to Akaike's Information Criterion (reported in brackets for each exchange), while ensuring that sufficient lags are included in the testing regression such that the error term becomes white noise. For the Phillips-Perron test the statistics Z_τ is reported.

Table 9. Price Difference Adjustment

Variable	Bitfinex	Bitstamp	bitFlyer	CEX.IO	Coinbase Pro	Gemini	Kraken
Monday	0.02***	0.0001	-0.0001	0.06**	0.01	0.01**	0.02***
Wednesday	0.01	0.02***	0.01*	0.06**	0.03***	0.02***	0.02***
Thursday	-0.001	-0.01*	0.01	-0.04	-0.013**	-0.01	-0.01
Friday	0.004	-0.004	0.01	-0.003	-0.004	-0.01**	0.01
Saturday	-0.003	-0.02***	0.01	-0.1***	-0.030***	-0.02***	-0.01
Sunday	0.06***	0.01*	0.01***	0.09***	0.003	0.02***	0.05***
March	0.2***	0.2***	0.02***	1.4***	0.2***	0.2***	0.39***
May	0.2***	0.06***	0.02***	0.7***	0.078***	0.08***	0.04***
June	0.2***	-0.1	0.03**	0.5***	0.02	0.05***	-0.04**
July	-0.05**	0.006***	0.03	-0.3***	-0.15***	-0.11***	-0.24***
August	-0.3***	-0.2***	0.04	-0.3***	-0.24***	-0.19***	-0.4***
September	-0.17	-0.2***	0.05	-0.3***	-0.2	-0.18***	-0.36***
01:00	-0.003	-0.003	0.003	0.01	0.003	0.01	0.002
02:00	0.002	0.007	0.01	0.02	0.01	0.01	0.006
03:00	0.009	0.02	0.02	0.04	0.01	0.01	0.01
04:00	0.01	0.02*	0.01	0.04	0.01	0.01	0.02
05:00	0.02	0.02*	0.02	0.05	0.02*	0.01	0.02
06:00	0.02	0.02**	0.01	0.1	0.02***	0.02	0.02
07:00	0.03*	0.02*	0.03	0.1	0.03***	0.01	0.02
08:00	0.03*	0.03*	0.04**	0.1	0.02***	0.01	0.03**
09:00	0.03*	0.03	0.03	0.1	0.02	0.01	0.03**
10:00	0.03**	0.01	0.05***	0.05	0.01	0.005	0.03**
11:00	0.03**	0.01	0.05**	0.03	0.02	0.013	0.02
12:00	0.02	-0.002	0.05***	-0.01	0.01	0.005	0.01
13:00	0.02	-0.007	0.03*	-0.02	0.003	0.005	0.01
14:00	0.01	-0.003	0.04**	-0.02	0.002	0.006	0.01
15:00	0.01	-0.002	0.02	-0.01	0.005	0.01	0.0005
16:00	0.01	0.006	0.04**	-0.02	0.008	0.007	-0.001
17:00	0.02	0.01	0.03	-0.01	0.01	0.01	0.00001
18:00	0.01	0.003	0.02	-0.01	0.007	0.01	-0.004
19:00	0.003	0.003	0.004	0.01	0.002	0.01	0.0005
20:00	0.002	0.005	-0.01	0.02	0.0001	0.01	0.0006
21:00	0.003	-0.001	-0.003	0.004	-0.0002	0.01	-0.01
22:00	0.0001	-0.003	-0.015	-0.004	-0.004	0.006	-0.01
23:00	0.001	-0.006	-0.013	-0.007	-0.002	0.003	-0.0005
Time	-0.0002***	$-4 \cdot 10^{-7}$	$-3 \cdot 10^{-5}$	$-5 \cdot 10^{-4***}$	$2 \cdot 10^{-5}$	$-2 \cdot 10^{-5}$	$4 \cdot 10^{-5**}$
Time^2	$-7 \cdot 10^{-8}$	$2 \cdot 10^{-8***}$	$-3 \cdot 10^{-5}$	$2 \cdot 10^{-7***}$	$2 \cdot 10^{-8***}$	$2 \cdot 10^{-8***}$	$2 \cdot 10^{-8***}$
Intercept	0.16***	0.10***	0.18***	0.61***	0.08***	0.1*** 0.11***	
Adjusted-R^2	0.33	0.26	0.11	0.37	0.29	0.30	0.34

Notes: This table reports OLS regressions for each exchange in the sample, the dependent variable is the hourly bitcoin price difference, defined in Equation 14. The time period is from 02:00:00, March 29, 2018 to 01:00:00, September 5, 2018. Time and Time^2 are linear and quadratic time trends, respectively. Dummy variables are included for hour-of-the-day, day-of-the week, and 7 calendar month dummies for March through September.

Table 10. Spread Adjustment

Variable	Bitfinex	Bitstamp	bitFlyer	CEX.IO	Coinbase Pro	Gemini	Kraken
Monday	0.003	0.165**	1.552***	0.010	-0.003	0.295***	0.204***
Wednesday	-0.002	0.097	1.901***	0.956***	0.004	0.026	0.151***
Thursday	-0.012	-0.060	0.628	-0.091	-0.014**	0.040	0.061
Friday	0.016*	0.195**	1.114*	0.659***	0.001	0.132	0.181***
Saturday	0.005	0.804***	-0.385	0.700***	0.01*	0.945***	0.370***
Sunday	0.012	0.911***	0.193	0.938***	0.01*	0.770***	0.354***
March	0.253***	1.351***	-3.768***	1.536***	0.055***	0.400**	0.540***
May	-0.033**	0.024	-0.791	-2.194***	-0.027**	0.958***	-0.107
June	-0.062***	-1.944***	-7.492***	-2.555***	-0.030	0.265	-1.077***
July	-0.070***	-3.852***	-7.573***	-3.101***	-0.064***	-0.125	-1.600***
August	-0.028	-3.892***	-7.919***	-6.732***	-0.047*	0.228	-1.615***
September	-0.053	-2.815***	-7.014**	-8.314***	-0.063**	0.526	-1.632***
01:00	-0.005	-0.033	-1.270	0.726	-0.027**	-0.289*	-0.367***
02:00	0.001	-0.189	0.119	0.810	-0.002	-0.381**	-0.176***
03:00	-0.023	-0.566***	0.241	0.308	-0.033***	-0.237	-0.402***
04:00	-0.037**	-0.645***	-0.784	0.278	-0.037***	-0.381**	-0.442***
05:00	-0.047***	-0.682***	-0.567	0.506	-0.033***	-0.293*	-0.485***
06:00	-0.037**	-0.800***	-1.050	0.105	-0.034***	-0.281*	-0.414***
07:00	-0.015	-0.851***	-0.888	-0.202	-0.034***	-0.176	-0.368***
08:00	-0.031**	-0.997***	-0.496	-0.932*	-0.034***	-0.145	-0.455***
09:00	-0.030**	-0.818***	1.111	-0.828*	-0.038***	0.005	-0.433***
10:00	-0.026*	-0.823***	1.291	-0.901*	-0.035***	-0.048	-0.363***
11:00	-0.020	-0.927***	0.122	-0.529	-0.034***	0.103	-0.378***
12:00	-0.003	-0.978***	0.785	-0.544	-0.032**	0.015	-0.415***
13:00	-0.023	-0.795***	0.675	-0.180	-0.028**	-0.131	-0.376***
14:00	-0.012	-0.934***	1.683	-0.175	-0.020	-0.474***	-0.408***
15:00	-0.019	-1.184***	0.578	-0.210	-0.023*	-0.518***	-0.560***
16:00	-0.016	-1.198***	0.409	-0.090	-0.019	-0.721***	-0.499***
17:00	-0.014	-1.006***	0.093	0.234	-0.023*	-0.778***	-0.603***
18:00	-0.005	-1.077***	0.287	0.214	-0.022*	-0.747***	-0.531***
19:00	-0.018	-1.096***	-1.144	-0.206	-0.016	-0.802***	-0.450***
20:00	-0.024	-1.229***	-1.912*	-8.8E-02	-0.021*	-0.876***	-0.471***
21:00	-0.027*	-1.149***	-0.479	-3.7E-01	-0.033***	-0.825***	-0.55***
22:00	-0.028*	-0.799***	-0.403	0.0004	-0.02	-0.67***	-0.45***
23:00	0.043***	-0.048	-1.152	-0.193	0.05***	0.43***	0.64***
Time	$-1 \cdot 10^{-4}$***	0.003***	-0.003**	0.001*	0.0001***	0.001***	0.0003
Time^2	$-1 \cdot 10^{-8}$***	$-1 \cdot 10^{-6}$***	$-5 \cdot 10^{-7}$	$-2 \cdot 10^{-7}$	$-1 \cdot 10^{-8}$***	$-2 \cdot 10^{-7}$	$-6 \cdot 10^{-8}$*
Tick Size	-0.011	-	-	-	0.011	-	-
Intercept	0.41***	4.05	23.95***	13.75***	0.01	1.24***	3.39***
Adjusted-R^2	0.27	0.55	0.22	0.17	0.04	0.18	0.41

Notes: This table reports OLS regressions for each exchange in the sample, the dependent variable is the hourly bid-ask spread, defined in Equation 7. The time period is from 02:00:00, March 29, 2018 to 01:00:00, September 5, 2018. Time and Time^2 are linear and quadratic time trends, respectively. Dummy variables are included for hour-of-the-day, day-of-the week, and 7 calendar month dummies for March through September.

Table 11. Estimation of ARIMA parameters for the hourly price differences. The lag order is chosen by AIC.

Exchange	Variable	Coefficient	Std. Error	Test statistic	Probability	R^2
Bitfinex	AR(1)	0.8248	0.0165	49.8893	< 0.01	0.8954
	AR(2)	0.0110	0.0215	0.5127	0.6082	
	AR(3)	0.0619	0.0215	2.8733	< 0.01	
	AR(4)	0.0045	0.0216	0.2080	0.8352	
	AR(5)	0.0303	0.0215	1.4063	0.1596	
	intercept	0.1889	0.0211	8.9632	< 0.01	
bitFlyer	AR(1)	0.5944	0.0165	36.0885	< 0.01	0.5246
	AR(2)	0.0466	0.0191	2.4377	0.0148	
	AR(3)	0.1143	0.0192	5.9587	< 0.01	
	AR(4)	0.0760	0.0166	4.5854	< 0.01	
	intercept	0.1828	0.0110	16.5344	< 0.01	
Bitstamp	AR(1)	0.7556	0.0165	45.7821	< 0.01	0.5691
	AR(2)	-0.0363	0.0208	-1.7474	0.0806	
	AR(3)	0.1029	0.0209	4.9255	< 0.01	
	AR(4)	0.0441	0.0209	2.1123	0.0347	
	AR(5)	-0.0096	0.0208	-0.4591	0.6461	
	AR(6)	0.0762	0.0166	4.5927	< 0.01	
	intercept	0.1369	0.0119	11.4737	< 0.01	
CEX.IO	AR(1)	1.0576	0.0165	64.1288	< 0.01	0.9403
	AR(2)	-0.2657	0.0241	-11.0412	< 0.01	
	AR(3)	0.1260	0.0245	5.1410	< 0.01	
	AR(4)	0.0049	0.0245	0.1994	0.8419	
	AR(5)	0.0515	0.0241	2.1372	0.0326	
	AR(6)	0.0011	0.0166	0.0659	0.9474	
	intercept	0.6560	0.0918	7.1484	< 0.01	
Coinbase Pro	AR(1)	0.9473	0.0180	52.5969	< 0.01	0.7953
	AR(2)	-0.0633	0.0184	-3.4409	< 0.01	
	intercept	0.1369	0.0080	17.1696	< 0.01	
Gemini	AR(1)	0.7389	0.0172	42.8915	< 0.01	0.828
	AR(2)	-0.0371	0.0215	-1.7285	0.0839	
	AR(3)	0.0982	0.0209	4.6919	< 0.01	
	AR(4)	0.0610	0.0212	2.8686	< 0.01	
	AR(5)	0.0729	0.0212	3.4350	< 0.01	
	AR(6)	0.0080	0.0167	0.4778	0.6328	
	intercept	0.1303	0.0136	9.5649	< 0.01	
Kraken	AR(1)	0.7790	0.0165	47.1991	< 0.01	0.8696
	AR(2)	-0.0057	0.0209	-0.2739	0.7842	
	AR(3)	0.0823	0.0209	3.9276	< 0.01	
	AR(4)	0.0961	0.0165	5.8264	< 0.01	
	intercept	0.1613	0.0166	9.7370	< 0.01	

Table 12. Estimation of ARIMA parameters for the daily price differences. The lag order is chosen by AIC.

Exchange	Variable	Coefficient	Std. Error	Test statistic	Probability	R^2
Bitfinex	AR(1)	0.7585	0.0517	14.6558	< 0.01	0.508
	intercept	0.1872	0.0313	5.9841	< 0.01	
bitFlyer	AR(1)	0.3006	0.1395	2.1551	0.0312	0.2513
	intercept	0.1586	0.0164	9.6400	< 0.01	
Bitstamp	AR(1)	0.8026	0.0697	11.509	< 0.01	
	MA(1)	-0.2455	0.1070	-2.295	0.0217	
	intercept	0.1364	0.0226	6.038	< 0.01	
CEX.IO	AR(1)	0.8160	0.0493	16.5373	< 0.01	0.6112
	intercept	0.6813	0.1409	4.8344	< 0.01	
Coinbase Pro	AR(1)	0.5189	0.0784	6.6146	< 0.01	0.2668
	AR(2)	0.2135	0.0791	2.6988	< 0.01	
	intercept	0.1365	0.0245	5.5639	< 0.01	
Gemini	AR(1)	0.7893	0.0508	15.5483	< 0.01	0.5238
	intercept	0.1280	0.0249	5.1425	< 0.01	
Kraken	AR(1)	0.8132	0.0789	10.3071	< 0.01	0.5703
	AR(2)	-0.0341	0.0818	-0.4168	0.6768	
	intercept	0.1566	0.0284	5.5084	< 0.01	

Table 13. Correlations between VAR innovations for the hourly series by exchange. Liquidity is measured by Quoted Spread.

Bitfinex	Not Adjusted (26 lags)			Adjusted (18 lags)		
	PD	QS	σ_t	PD	QS	σ_t
PD	1.00	0.06	0.08	1.00	0.07	0.13
QS	0.06	1.00	0.21	0.07	1.00	0.26
σ_t	0.08	0.21	1.00	0.13	0.26	1.00
Bitstamp	Not Adjusted (26 lags)			Adjusted (10 lags)		
PD	1.00	-0.04	0.10	1.00	-0.03	0.30
QS	-0.04	1.00	0.04	-0.03	1.00	0.02
σ_t	0.10	0.04	1.00	0.30	0.02	1.00
CEX.IO	Not Adjusted (29 lags)			Adjusted (28 lags)		
PD	1.00	0.06	0.10	1.00	0.06	0.11
QS	0.06	1.00	0.24	0.06	1.00	0.24
σ_t	0.10	0.24	1.00	0.11	0.24	1.00
Coinbase Pro	Not Adjusted (24 lags)			Adjusted (9 lags)		
PD	1.00	0.02	0.06	1.00	0.04	0.11
QS	0.02	1.00	0.15	0.04	1.00	0.18
σ_t	0.06	0.15	1.00	0.11	0.18	1.00
Gemini	Not Adjusted (25 lags)			Adjusted (9 lags)		
PD	1.00	-0.05	0.07	1.00	-0.02	0.11
QS	-0.05	1.00	0.15	-0.02	1.00	0.18
σ_t	0.07	0.15	1.00	0.11	0.18	1.00
Kraken	Not Adjusted (20 lags)			Adjusted (20 lags)		
PD	1.00	0.10	0.10	1.00	0.11	0.13
QS	0.10	1.00	0.21	0.11	1.00	0.24
σ_t	0.10	0.21	1.00	0.13	0.24	1.00

Notes: The table reports correlations in VAR innovations (residuals) for each exchange for the hourly series. The input data for VAR estimation are bitcoin price differences (PD), quoted bid-ask spread (QS) and spot volatility (σ_t). The number of lags p_i used in VAR for each exchange i is presented in parentheses. Two columns report the test results for adjusted and unadjusted series of the endogenous variables.

Table 14. Correlations in VAR innovations for the hourly series by exchange. Liquidity is measured by Quoted Depth.

Bitfinex	Not Adjusted (19 lags)			Adjusted (19 lags)		
	PD	$AvgD_t$	σ_t	PD	$AvgD_t$	σ_t
PD	1.00	0.02	0.08	1.00	0.02	0.14
$AvgD_t$	0.02	1.00	0.04	0.02	1.00	0.05
σ_t	0.08	0.04	1.00	0.14	0.05	1.00
Bitstamp	Not Adjusted (12 lags)			Adjusted (12 lags)		
PD	1.00	0.05	0.10	1.00	0.06	0.10
$AvgD_t$	0.05	1.00	0.11	0.06	1.00	0.09
σ_t	0.10	0.11	1.00	0.10	0.09	1.00
CEX.IO	Not Adjusted (21 lags)			Adjusted (20 lags)		
PD	1.00	0.03	0.10	1.00	0.02	0.14
$AvgD_t$	0.03	1.00	0.04	0.02	1.00	0.05
σ_t	0.010	0.04	1.00	0.14	0.05	1.00
Coinbase Pro	Not Adjusted (24 lags)			Adjusted (20 lags)		
PD	1.00	-0.05	0.06	1.00	-0.04	0.08
$AvgD_t$	-0.05	1.00	-0.03	-0.04	1.00	-0.06
σ_t	0.06	-0.03	1.00	0.08	-0.06	1.00
Gemini	Not Adjusted (24 lags)			Adjusted (20 lags)		
PD	1.00	-0.04	0.07	1.00	-0.04	0.08
$AvgD_t$	-0.04	1.00	-0.04	-0.04	1.00	-0.07
σ_t	0.07	-0.04	1.00	0.08	-0.07	1.00
Kraken	Not Adjusted (24 lags)			Adjusted (16 lags)		
PD	1.00	-0.01	0.10	1.00	-0.005	0.11
$AvgD_t$	-0.01	1.00	0.02	-0.005	1.00	0.02
σ_t	0.10	0.02	1.00	0.11	0.02	1.00

Notes: The table reports correlations in VAR innovations (residuals) for each exchange for the hourly series. The input data for VAR estimation are bitcoin price differences (PD), quoted depth $_t$ and spot volatility (σ_t). The number of lags p_i used in VAR for each exchange i is presented in parentheses. Two columns report the test results for adjusted and unadjusted series of the endogenous variables.

Table 15. Granger Causality Tests with Quoted Depth for the hourly series

Bitfinex	Not Adjusted (26 lags)			Adjusted (18 lags)		
	PD	$AvgD_t$	σ_t	PD	$AvgD_t$	σ_t
PD	-	19.85 (0.40)	21.8 (0.29)	-	9.25 (0.16)	39.98*** (< 0.01)
$AvgD_t$	16.34 (0.63)	-	38.95*** (0.004)	3.73 (0.71)	-	15.17** (0.01)
σ_t	18.25 (0.50)	59.93*** (< 0.01)	-	6.09 (0.42)	50.92*** (< 0.01)	-
Bitstamp	Not Adjusted (26 lags)			Adjusted (10 lags)		
PD	-	29.39 (0.21)	18.91 (0.76)	-	12.04 (0.28)	11.78 (0.32)
$AvgD_t$	21.97 (0.58)	-	23 (0.52)	8.09 (0.62)	-	7.94 (0.63)
σ_t	13.39 (0.95)	34.52* (0.07)	-	1.19 (0.98)	4.46 (0.61)	-
CEX.IO	Not Adjusted (29 lags)			Adjusted (28 lags)		
PD	-	34.07** (0.03)	31.95** (0.04)	-	28.76 (0.42)	47.84*** (< 0.01)
$AvgD_t$	14.94 (0.78)	-	53.64*** (< 0.01)	20.08 (0.86)	-	62.03*** (< 0.01)
σ_t	25.60 (0.18)	115.03*** (< 0.01)	-	25.95 (0.57)	105.95*** (< 0.01)	-
Coinbase Pro	Not Adjusted (24 lags)			Adjusted (9 lags)		
PD	-	27.43 (0.28)	35.67* (0.06)	-	18.11** (0.03)	37.32*** (< 0.01)
$AvgD_t$	32.03 (0.13)	-	28.37 (0.24)	15.02* (0.09)	-	10.81 (0.29)
σ_t	29.94 (0.17)	36.56** (0.05)	-	15.47* (0.08)	44.22*** (< 0.01)	-
Gemini	Not Adjusted (25 lags)			Adjusted (9 lags)		
PD	-	20.47 (0.43)	28.15 (0.11)	-	4.92 (0.42)	45.91*** (< 0.01)
$AvgD_t$	13.59 (0.85)	-	30.21* (0.07)	2.61 (0.76)	-	683.63*** (< 0.01)
σ_t	32.14 (0.15)	36.56* (0.06)	-	9.49* (0.09)	15.50*** (0.008)	-
Kraken	Not Adjusted (20 lags)			Adjusted (20 lags)		
PD	-	31.90 (0.13)	39.35** (0.02)	-	9.92 (0.13)	35.24*** (< 0.01)
$AvgD_t$	21.03 (0.64)	-	22.27 (0.56)	2.00 (0.92)	-	27.27*** (< 0.01)
σ_t	24.13 (0.45)	72.77*** (< 0.01)	-	13.77** (< 0.01)	59.32*** (0.001)	-

Notes: The table reports Granger-causality tests. Null hypothesis: row variable does not Granger-cause column variable, more precisely, it is tested whether the coefficients on all lags of the causing variable i are jointly zero when j is dependent in the VAR. Both chi-square statistics and p-values (in parentheses) of pairwise Granger-causality tests between the endogenous variables are presented. p-values are based on heteroscedasticity-consistent (HC) standard errors (White (1980) corrected standard errors). The lags in parentheses represent the lag order chosen by AIC and SC information criteria.

Table 16. Correlations in VAR innovations for the daily series by each exchange. Liquidity is measured by Quoted Spread.

Bitfinex	Not Adjusted (1 lags)			Adjusted (1 lag)		
	PD	QS	σ_t	PD	QS	σ_t
PD	1.00	0.18	0.22	1.00	0.26	0.29
QS	0.18	1.00	0.05	0.26	1.00	0.46
σ_t	0.22	0.05	1.00	0.29	0.46	1.00
Bitstamp	Not Adjusted (9 lags)			Adjusted (6 lags)		
PD	1.00	0.01	-0.007	1.00	-0.005	0.06
QS	0.01	1.00	0.05	-0.005	1.00	0.02
σ_t	-0.007	0.05	1.00	0.06	0.02	1.00
CEX.IO	Not Adjusted (2 lags)			Adjusted (2 lag)		
PD	1.00	0.22	0.32	1.00	0.27	0.39
QS	0.22	1.00	0.16	0.27	1.00	0.25
σ_t	0.32	0.16	1.00	0.39	0.25	1.00
Coinbase Pro	Not Adjusted (2 lags)			Adjusted (1 lag)		
PD	1.00	0.17	0.09	1.00	0.21	0.19
QS	0.17	1.00	0.42	0.21	1.00	0.51
σ_t	0.09	0.42	1.00	0.19	0.51	1.00
Gemini	Not Adjusted (2 lags)			Adjusted (4 lags)		
PD	1.00	0.15	0.26	1.00	0.14	0.36
QS	0.15	1.00	0.11	0.14	1.00	0.28
σ_t	0.26	0.11	1.00	0.36	0.28	1.00
Kraken	Not Adjusted (4 lags)			Adjusted (3 lags)		
PD	1.00	0.29	0.27	1.00	0.27	0.36
QS	0.29	1.00	0.30	0.27	1.00	0.28
σ_t	0.27	0.30	1.00	0.36	0.28	1.00

Notes: The table reports correlations in VAR innovations (residuals) for each exchange for the daily series. The input data for VAR estimation are bitcoin price differences (PD), quoted bid-ask spread (QS) and spot volatility (σ_t). The number of lags p_i used in VAR for each exchange i is presented in parentheses. Two columns report the test results for adjusted and unadjusted series of the endogenous variables.

Table 17. Granger Causality Tests with Quoted Spread for the daily series

Bitfinex	Not Adjusted (1 lags)			Adjusted (1 lag)		
	PD	QS	σ_t	PD	QS	σ_t
PD	-	0.11 (0.73)	0.32 (0.56)	-	0.42 (0.51)	3.18* (0.07)
QS	0.03 (0.86)	-	5.42** (0.02)	0.21 (0.64)	-	0.81 (0.36)
σ_t	1.49 (0.22)	0.001 (0.97)	-	0.54 (0.46)	0.07 (0.79)	-
Bitstamp	Not Adjusted (9 lags)			Adjusted (6 lags)		
PD	-	2.60 (0.95)	3.18 (0.92)	-	1.62 (0.20)	0.001 (0.97)
QS	11.11 (0.19)	-	4.02 (0.85)	0.04 (0.83)	-	0.10 (0.74)
σ_t	18.46** (0.02)	25.69*** (0.001)	-	1.14 (0.28)	0.02 (0.88)	-
CEX.IO	Not Adjusted (2 lags)			Adjusted (2 lag)		
PD	-	2.23 (0.32)	3.60 (0.16)	-	2.33 (0.31)	3.42 (0.18)
QS	3.51 (0.17)	-	2.78 (0.24)	3.55 (0.17)	-	1.79 (0.41)
σ_t	1.22 (0.54)	0.74 (0.68)	-	1.21 (0.54)	0.71 (0.70)	-
Coinbase Pro	Not Adjusted (2 lags)			Adjusted (1 lag)		
PD	-	1.89 (0.38)	2.51 (0.28)	-	5.02** (0.02)	13.58*** (< 0.01)
QS	1.81 (0.40)	-	2.62 (0.26)	2.68 (0.10)	-	1.81 (0.18)
σ_t	0.19 (0.91)	2.24 (0.32)	-	4.58** (0.03)	0.39 (0.53)	-
Gemini	Not Adjusted (2 lags)			Adjusted (4 lags)		
PD	-	1.96 (0.37)	3.85 (0.14)	-	8.28* (0.08)	11.12** (0.02)
QS	2.01 (0.36)	-	0.60 (0.73)	8.32* (0.08)	-	1.23 (0.87)
σ_t	1.56 (0.45)	0.46 (0.79)	-	10.37** (0.03)	10.29** (0.03)	-
Kraken	Not Adjusted (4 lags)			Adjusted (3 lags)		
PD	-	7.98* (0.09)	10.52** (0.03)	-	0.001 (0.97)	16.34*** (< 0.01)
QS	8.33* (0.08)	-	5.99 (0.19)	0.07 (0.78)	-	2.60 (0.10)
σ_t	10.37** (0.03)	3.69 (0.44)	-	0.41 (0.51)	0.65 (0.41)	-

Notes: The table reports Granger-causality tests. Null hypothesis: row variable does not Granger-cause column variable, more precisely, it is tested whether the coefficients on all lags of the causing variable i are jointly zero when j is dependent in the VAR. Both chi-square statistics and p-values (in parentheses) of pairwise Granger-causality tests between the endogenous variables are presented. p-values are based on heteroscedasticity-consistent (HC) standard errors (White (1980) corrected standard errors). The lags in parentheses represent the lag order chosen by AIC and SC information criteria.

Table 18. Correlations in VAR innovations for the daily series by exchange. Liquidity is measured by Quoted Depth.

Bitfinex	Not Adjusted (1 lags)			Adjusted (1 lags)		
	PD	$AvgD_t$	σ_t	PD	$AvgD_t$	σ_t
PD	1.00	-0.04	0.20	1.00	-0.003	0.30
$AvgD_t$	-0.04	1.00	0.006	-0.003	1.00	0.01
σ_t	0.20	0.006	1.00	0.30	0.01	1.00
Bitstamp	Not Adjusted (2 lags)			Adjusted (2 lags)		
PD	1.00	0.002	0.03	1.00	0.04	0.1
$AvgD_t$	0.002	1.00	0.1	0.04	1.00	-0.04
σ_t	0.03	0.1	1.00	0.1	-0.04	1.00
CEX.IO	Not Adjusted (1 lag)			Adjusted (1 lag)		
PD	1.00	0.07	0.33	1.00	0.02	0.35
$AvgD_t$	0.07	1.00	-0.08	0.02	1.00	-0.15
σ_t	0.33	-0.08	1.00	0.35	-0.15	1.00
Coinbase Pro	Not Adjusted (7 lags)			Adjusted (3 lags)		
PD	1.00	-0.16	0.10	1.00	-0.16	0.20
$AvgD_t$	-0.16	1.00	-0.16	-0.16	1.00	-0.28
σ_t	0.10	-0.16	1.00	0.20	-0.28	1.00
Gemini	Not Adjusted (1 lag)			Adjusted (1 lag)		
PD	1.00	0.03	0.22	1.00	0.03	0.35
$AvgD_t$	0.03	1.00	0.07	0.03	1.00	-0.0002
σ_t	0.22	0.07	1.00	0.35	-0.0002	1.00
Kraken	Not Adjusted (1 lag)			Adjusted (1 lag)		
PD	1.00	0.04	0.35	1.00	0.15	0.39
$AvgD_t$	0.04	1.00	-0.08	0.15	1.00	0.007
σ_t	0.35	-0.08	1.00	0.39	0.007	1.00

Notes: The table reports correlations in VAR innovations (residuals) for each exchange for the daily series. The input data for VAR estimation are bitcoin price differences (PD), quoted depth ($AvgD_t$) and spot volatility (σ_t). The number of lags p_i used in VAR for each exchange i is presented in parentheses. Two columns report the test results for adjusted and unadjusted series of the endogenous variables.

Table 19. Granger Causality Tests with Quoted Depth for the daily series

Bitfinex	Not Adjusted (1 lags)			Adjusted (1 lags)		
	PD	$AvgD_t$	σ_t	PD	$AvgD_t$	σ_t
PD	-	0.42 (0.52)	0.42 (0.51)	-	0.66 (0.41)	3.60* (0.06)
$AvgD_t$	0.01 (0.91)	-	5.69** (0.31)	0.91 (0.02)	-	1.04 (0.33)
σ_t	3.28* (0.07)	12.50*** (<0.01)	-	0.46 (0.49)	8.90*** (0.002)	-
Bitstamp	Not Adjusted (2 lags)			Adjusted (2 lags)		
PD	-	2.59 (0.27)	0.86 (0.65)	-	2.81 (0.25)	0.14 (0.93)
$AvgD_t$	7.26** (0.03)	-	4.82* (0.09)	9.67*** (0.01)	-	0.23 (0.89)
σ_t	0.36 (0.83)	0.71 (0.70)	-	1.90 (0.38)	4.98* (0.08)	-
CEX.IO	Not Adjusted (1 lag)			Adjusted (1 lag)		
PD	-	3.42* (0.06)	0.73 (0.39)	-	2.80 (0.59)	22.21*** (<0.01)
$AvgD_t$	0.99 (0.32)	-	8.55*** (<0.01)	1.04 (0.90)	-	12.47*** (0.01)
σ_t	2.07 (0.15)	8.33*** (<0.01)	-	5.99 (0.19)	8.80* (0.07)	-
Coinbase Pro	Not Adjusted (7 lags)			Adjusted (3 lags)		
PD	-	0.53 (0.77)	2.35 (0.30)	-	2.60 (0.45)	9.63** (0.02)
$AvgD_t$	2.04 0.36)	-	0.55 (0.75)	10.55*** (0.01)	-	2.34 (0.50)
σ_t	1.21 (0.54)	1.34 (0.50)	-	5.84 (0.12)	3.02 (0.38)	-
Gemini	Not Adjusted (1 lag)			Adjusted (1 lag)		
PD	-	1.28 (0.26)	4.16** (0.04)	-	0.007 (0.93)	9.78*** (<0.01)
$AvgD_t$	0.68 (0.41)	-	0.41 (0.52)	0.29 (0.59)	-	0.14 (0.71)
σ_t	0.77 (0.38)	2.76 (0.10)	-	0.92 (0.34)	0.47 (0.49)	-
Kraken	Not Adjusted (1 lag)			Adjusted (1 lag)		
PD	-	1.04 (0.31)	7.30*** (<0.01)	-	0.22 (0.63)	13.72*** (<0.01)
$AvgD_t$	0.43 (0.51)	-	0.07 (0.80)	0.29 (0.59)	-	0.01 (0.90)
σ_t	0.47 (0.49)	0.03 (0.86)	-	0.14 (0.78)	0.07 (0.49)	-

Notes: The table reports Granger-causality tests. Null hypothesis: row variable does not Granger-cause column variable, more precisely, it is tested whether the coefficients on all lags of the causing variable i are jointly zero when j is dependent in the VAR. Both chi-square statistics and p-values (in parentheses) of pairwise Granger-causality tests between the endogenous variables are presented. p-values are based on heteroscedasticity-consistent (HC) standard errors (White (1980) corrected standard errors). The lags in parentheses represent the lag order chosen by AIC and SC information criteria.

References

Demsetz, H. (1968). "The Cost of Transacting." In: *The Quarterly Journal of Economics.*

Box, G. and G. Jenkins (1970). *Time Series Analysis: Forecasting and Control.*

Stoll, H. (1978a). "The pricing of security dealer services: An empirical study of NASDAQ stocks." In: *Journal of Finance.*

— (1978b). "The supply of dealer services in securities markets." In: *Journal of Finance.*

Amihud, Y. and H. Mendelson (1980). "Dealership market: Market-making with inventory." In: *Journal of Financial Economics.*

White, H. (1980). "A heteroskedasticity consistent covariance matrix estimator and a direct test for heteroskedasticity." In: *Econometrica.*

Copeland, D. and T. Galai (1983). "Information effects on the bid-ask spread." In: *Journal of Finance.*

Glosten, L. and P. Milgrom (1985). "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders." In: *Journal of Financial Economies.*

Kyle, A. S. (1985). "Continuous Auctions and Insider Trading." In: *Econometrica.*

Amihud, Y. and H. Mendelson (1986). "Asset pricing and the bid-ask spread." In: *Journal of Financial Economics.*

O'Hara, M. and G. Oldfield (1986). "The microeconomics of market making." In: *Journal of Financial and Quantitative Analysis.*

Rosenthal, L. and c. Young (1990). "The Seemingly Anomalous Price Behavior of Royal Dutch/Shell and Unilever." In: *Journal of Financial Economics, Vol. 26.*

Dufour, Jean-Marie and David Tessier (1992). "On the relationship between impulse response analysis, innovation accounting and Granger causality." In: *Economics Letters.*

Mcinish, T.H. and Wood R.A. (1992). "An analysis of intraday patterns in bid/ask spreads for NYSE stocks." In: *Journal of Finance.*

Subrahmanyam, A. (1994). "Circuit breakers and market volatility: A theoretical perspective." In: *Journal of Finance.*

Hasbrouck, J. (1995). "Many markets: determining the contributions to price discovery." In: *Journal of Finance.*

Holden, C. (1995). "Index arbitrage as cross sectional market making." In: *Journal of Futures Markets.*

Longin, F. (1997). “The threshold effect in expected volatility: A model based on asymmetric information.” In: *Review of Financial Studies.*

Madhavan, A., M. Richardson and M. Roomans (1997). “Why do prices change? A transaction-level analysis of NYSE stocks.” In: *Review of Financial Studies.*

Shleifer, A. and R. W Vishny (1997). “The limits of arbitrage.” In: *Journal of Finance.*

Torre, N. (1997). *Barra market impact model handbook.*

Froot, K.A. and E.M. Dabora (1999). “How are stock prices affected by the location of trade?” In: *Journal of Financial economics.*

Shleifer, A. (2000). “Clarendon Lectures: Inefficient Markets.” In: *Oxford University Press, Oxford.*

Chordia, T., R. Roll and A. Subrahmanyam (2002). “Order imbalance, liquidity, and market returns.” In: *Journal of Financial Economics 65.*

Evans, M. D. D. and R. K. Lyons (2002). “Order flow and exchange rate dynamics.” In: *Journal of Political Economy 110.*

Shen, P. and R.M. Starr (2002). “Market-makers’ supply and pricing of financial market liquidity.” In: *Economics Letters.*

Barberis, N. and R. Thaler (2003). “A survey of behavioral finance”. In: *Handbook of the Economics of Finance.*

Brandt, M. W. and K. A. Kavajecz (2004). “Price discovery in the US Treasury market: the impact of order flow and liquidity on the yield curve.” In: *Journal of Finance 59.*

Ofek, E., M. Richardson and R. F. Whitelaw (2004). “Limited Arbitrage and Short Sales Restrictions: Evidence from the Options Markets.” In: *Journal of Financial Eco- nomics.*

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis.*

Pontiff, J. (2006). “Costly Arbitrage and the Myth of Idiosyncratic Risk.” In: *Journal of Accounting and Economics.*

Roll, R. W., E. S. Schwartz and A. Subrahmanyam (2007). “Liquidity and The Law of One Price: The Case of the Cash/Futures Basis.” In: *The Journal of Finance.*

Berger, D. et al. (2008). “Order flow and exchange rate dynamics in electronic brokerage system data.” In: *Journal of International Economics 75.*

Easley, D. et al. (2008). “Time-varying arrival rates of informed and uninformed trades.” In: *Journal of Financial Econometrics.*

Nakamoto, S. (2008). “Bitcoin: a peer-to-peer electronic cash system.” In:

Goyenko, R. Y., C. W. Holden and C. A. Trzcinka (2009). "Do liquidity measures measure liquidity?" In: *Journal of Financial Economics 92*.

Hafner, C.M. and H. Herwatz (2009). "Testing for linear vector autoregressive dynamics under multivariate generalized autoregressive heteroskedasticity." In: *Statistica Neerlandica*.

Rosenthal, L., Abe De Jong and Mathijs A. Van Dijk (2009). "The Risk and Return of Arbitrage in Dual-Listed Companies." In: *Review of Finance, Vol. 13*.

Comerton-Forde, C. et al. (2010). "Time Variation in Liquidity: The Role of Market-Maker Inventories and Revenues." In: *Journal of Finance*.

Gagnon, L. and G. A. Karolyi (2010). "Multi-Market Trading and Arbitrage." In: *Journal of Financial Economics, Vol. 97*.

Gatheral, J. (2010). "No-dynamic-arbitrage and market impact." In: *Quantita-tive finance*.

Gromb, D. and D. Vayanos (2010). "Limits of Arbitrage". In: *Annual Review of Financial Economics*.

Kristensen, D. (2010). "Nonparametric Filtering of the Realized Spot Volatility: A Kernel-Based Approach." In: *Econometric Theory*.

Deuskar, P. and T. C. Johnson (2011). "Market liquidity and flow-driven risk." In: *Review of Financial Studies 24*.

Kotz, S., T. Kozubowski and K. Podgorski (2012). "The Laplace Distribution and Gen- eralizations: a Revisit with Applications to Communications, Economics, Engineering, and Finance." In: *Springer*.

Gould, M. D. et al. (2013). "Limit order books". In: *Quantitative Finance*.

KRAKENFX (2013). *Kraken Forms Exclusive Partnership with Fidor Bank*. URL: https://blog.kraken.com/post/303/kraken-forms-exclusive-partnership-with-fidor-bank/.

Moore, Tyler and Nicolas Christin (2013). "Beware the MiddleMan: Empirical Analysis of Bitcoin-Exchange Risk." In: *Financial Cryptography and Data Security*.

Brogaard, J., T. Hendershott and R. Riordan (2014). "High frequency trading and price discovery." In: *Review of Financial Studies*.

Hendershott, T. and A. J. Menkveld (2014). "Price pressures." In: *Journal of Financial Economics 114*.

Holden, C. and S. Jacobsen (2014). "Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions." In: *Journal of Finance*.

Kannenberg, A. (2014). *Bitcoin-Börse in Abwärtsspirale: "Mt. Gox has left the building.* URL: https://www.heise.de/newsticker/meldung/Bitcoin-Boerse-in-Abwaertsspirale-Mt-Gox-has-left-the-building-2119550.html.

Records, Guinness World (2014). *Largest cryptocurrency hack.* URL: https://www.guinnessworldrecords.com/world-records/696467-largest-cryptocurrency-hack.

Böhme, R. et al. (2015). "Bitcoin: Economics, technology, and governance." In: *Journal of Economic Perspectives.*

Fourel, V. et al. (2015). "Common factors, order flows, and exchange rate dynamics. Unpublished working paper." In: *MIT, City University London, BI Norwegian Business School, and Goethe University Frankfurt.*

Nasdaq (2015). *A $480 Million Mystery: The Saga of Mt. Gox.* URL: https://www.nasdaq.com/articles/480-million-mystery-saga-mt-gox-2015-08-10.

Rubisov, Anton. D. (2015). "Statistical Arbitrage Using Limit Order Book Imbalance." In: *University of Toronto Institute for Aerospace Studies.*

Russell, J. (2015). "Bitstamp Suspends Its Service After Hackers Snatch Nearly $5M in Bitcoins." In: *TechCrunch.*

BBC (2016). *Bitcoin drops more than 10% after security breach.* URL: https://www.bbc.com/news/business-36962254.

Tepper, F. (2016). "Hacked Bitcoin Exchange Bitfinex will reduce balances by 36% to distribute losses amongst all users." In: *TechCrunch.*

Catalini, C. and J. S. Gans (2017). "Some Simple Economics of the Blockchain." In: *Rotman School of Management Working Paper No. 2874598; MIT Sloan Research Paper No. 5191-16.*

Chochan, U. W. (2017). "The Double Spending Problem and Cryptocurrencies." In: *Banking Insurance Journal.*

Foucault, T., R. Kozhan and W. W. Tham (2017). "Toxic Arbitrage." In: *Review of Financial Studies.*

Kroeger, A. and A. Sarkar (2017). "The Law of One Bitcoin Price?" In: *Federal Reserve Bank of New York.*

Lütkepohl, H. and L. Kilian (2017). *Structural Vector Autoregressive Analysis.*

U.S. Attorney's Office (2017). *Russian National And Bitcoin Exchange Charged In 21-Count Indictment For Operating Alleged International Money Laundering Scheme And Allegedly Laundering Funds From Hack Of Mt. Gox.* URL: https://www.justice.gov/usaondca/pr/russian-national-and-bitcoin-exchange-charged-21-count-indictment-operating-alleged.

Yermack, D. (2017). "Corporate Governance and Blockchains." In: *Review of Finance.*

Bitfinex (2018). *Fees schedule.* URL: https://www.bitfinex.com/fees/.

CEX.IO (2018). *Fees schedule.* URL: https://broker.cex.io/fees.

Dyhrberg, A. H., S. Foley and J. Svec (2018). "How investible is Bitcoin? Analyzing the liquidity and transaction costs of Bitcoin markets." In: *Economics Letters.*

Goldenberg, T. (2018). *Watch out Crypto Exchanges, Decentralization is coming.* URL: https://www.coindesk.com/future-crypto-exchanges-decentralization-coming.

Golstein, S. (2018). *Research Shows that Most Major Cryptocurrency Providers Not KYC Compliant.* URL: https://www.financemagnates.com/cryptocurrency/news/research-shows-major-cryptocurrency-providers-not-kyc-compliant/.

Kharpal, A. (2018). *Over $100 billion wiped off global cryptocurrency market following talk of South Korea trading ban.* URL: https://www.cnbc.com/2018/01/11/bitcoin-ripple-ethereum-prices-fall-after-south-korea-trading-ban-talk.html.

Peh, B. (2018). *Why are crypto exchanges charging exorbitant withdrawal fees?* URL: https://www.linkedin.com/pulse/why-crypto-exchanges-charging-exorbitant-withdrawal-fees-bernard-peh?trk=related_artice_Why%20are%20crypto%20exchanges%20charging%20exorbitant%20withdrawal%20fees%3F_article-card_title.

Russolino, S. and Eun-Young Jeong (2018). "Cryptocurrency Exchanges Are Getting Hacked Because It's Easy." In: *The Wall Street Journal.*

Wenxin, D., A. Tepper and A. Verdelhan (2018). "Deviations from Covered Interest Rate Parity." In: *The Journal of Finance.*

Zach, R. (2018). *Complete Guide to Institutional-Grade Crypto-Exchanges.* URL: https://medium.com/@reinhardzach/complete-guide-to-institutional-grade-crypto-exchanges-293b8e8d28ef.

Biais, B. et al. (2019). "The Blockchain Folk Theorem." In: *The Review of Financial Studies.*

Casey, M. J. (2019). *Banking, Bitfinex and the Hidden Irony of Crypto's Newest Controversy.* URL: https://www.coindesk.com/banking-bitfinex-and-the-hidden-irony-of-cryptos-newest-controversy.

Copeland, T. (2019). *Decrypt.* URL: https://decrypt.co/6824/crypto-capital-follow-up-article.

CryptoCompare (2019). *Bitcoin Premium on Bitfinex Spikes After Allegations of $850 Million 'Cover-up'*. URL: https://blog.cryptocompare.com/bitcoin-premium-on-bitfinex-spikes-after-allegations-of-850-million-cover-up-9667004b87eb.

Easley, D., M. O'Hara and S. Basu (2019). "From Mining to Markets: The Evolution of Bitcoin Transaction Fees." In: *Journal of Financial Economics*.

Huang, K. (2019). *CCXT Pro*. URL: https://github.com/hzjken/crypto-arbitrage-framework.

James, L. (2019). *Attorney General James Announces Court Order Against "Crypto" Currency Company Under Investigation For Fraud*. URL: https://ag.ny.gov/press-release/2019/attorney-general-james-announces-court-order-against-crypto-currency-company.

Morris, D. Z. (2019). *BREAKERMAG*. URL: https://breakermag.com/crypto-capital-is-the-common-thread-between-bitfinex-quadrigacx-and-other-troubled-exchanges/.

New York County Clerk (2019). "Affirmation of Brian M. Whitehurst in Support of OAG's Ex Parte Application for an Order Pursuant to General Business Law §354". In: *Supreme Court of the State of New York*.

Shi, Fa-Bin et al. (2019). "Anomaly detection in Bitcoin market via price return analysis." In: *PLoS ONE*.

Vigna, P. (2019). *Lack of Banking Options a Big Problem for Crypto Businesses*. URL: https://www.wsj.com/articles/lack-of-banking-options-a-big-problem-for-crypto-businesses-11558092600.

Zhao, W. (2019). *Bitfinex has applied to a U.S. court to subpoena a former banking executive of a trust company that the crypto exchange believes can help it recover over $850 million in frozen funds*. URL: https://www.coindesk.com/bitfinex-files-for-subpoena-in-bid-to-recover-880-million-in-frozen-funds.

bitcoinity.org (2020). *Bitcoin trading volume*. URL: http://data.bitcoinity.org/markets/volume/30d?c=e&t=b..

Blockchain (2020). *Marktkapitalisierung (USD)*. URL: https://www.blockchain.com/de/charts/market-cap?timespan=all.

Cheng, Y. (2020). *Alleged 'shadow banker' for crypto exchanges moves toward guilty plea on unlicensed money transmission business charge*. URL: https://finance.yahoo.com/news/alleged-shadow-banker-crypto-exchanges-224702169.html.

FDIC (2020). *About the FDIC*. URL: https://www.fdic.gov/.

Hamilton, J. (2020). *Blackbird*. URL: https://github.com/butor/blackbird.

Hautsch, N., C. Scheuch and S. Voigt (2020). "Building Trust takes Time: Limits to Arbitrage in Blockchain-Based Markets." In: *Working paper*.

Huang, K. (2020). *Crypto Arbitrage Framework*. URL: https://github.com/butor/blackbird.

Makarov, I. and A. Schoar (2020). "Trading and Arbitrage in Cryptocurrency Markets." In: *Journal of Financial Economics*.

Voigt, S. (2020). "Liquidity and Price Informativeness in Blockchain-Based Markets." In: *Working Paper*.

Belongia, Michael. T. and Peter N. Ireland (2021). "The Transmission of Monetary Policy Shocks Through the Markets for Reserves and Money." In:

Coinbase (2021). *Coinbase expands international payments options for institutional customers*. URL: https://blog.coinbase.com/coinbase-expands-international-payments-options-for-institutional-customers-1c92ae7283af.

Pound, J. (2021). *Bitcoin hits $1 trillion in market value as cryptocurrency surge continues*. URL: https://www.cnbc.com/2021/02/19/bitcoin-hits-1-trillion-in-market-value-as-cryptocurrency-surge-continues.html.

Rösch, D. (2021). "The impact of arbitrage on market liquidity." In: *Journal of Financial Economics*.

Kroitor, I. (2022). *CCXT – CryptoCurrency eXchange Trading Library*. URL: https://github.com/ccxt/ccxt.