Arbitrage, Liquidity and Price Informativeness

Stefan Voigt*

November 16, 2019

Click here to download the most recent version of this paper

Abstract

Limits to arbitrage impose costs on cross-market trading and harm informational efficiency in fragmented markets. I quantify the impact of blockchain-related trading frictions that arise from the time-consuming settlement process in the market for Bitcoin. The estimation rests on an error correction model that exploits the notion that arbitrageurs suspend their activity when arbitrage costs exceed price differences. I estimate substantial arbitrage costs that explain 63% of the observed price differences, where more than 75% of these costs can be attributed to settlement latency. I also find that a 10 bp decrease in technology-related arbitrage costs simultaneously results in a 3 bp increase of the spreads. I embed this finding in a theoretical model in which liquidity providers set larger spreads to cope with the higher adverse selection risks imposed by increased arbitrage activity. Consequently, efforts to reduce the latency of blockchain-based settlement might have unintended consequences for liquidity provision. In markets with substantial adverse selection risk, reduced technology-related arbitrage costs may thus even harm informational efficiency.

JEL Codes: G00, G10, G14

Keywords: Arbitrage, Market Frictions, Distributed Ledger, Blockchain

---

*Stefan Voigt, VGSF (Vienna Graduate School of Finance) and WU (Vienna University of Economics and Business), Department of Finance, Accounting and Statistics, Welthandelsplatz 1, Building D4, 1020 Vienna, Austria. Correspondence to: stefan.voigt@vgsf.ac.at (www.voigtstefan.me). I thank Torben Andersen, Thierry Foucault, Thomas Gehrig, Alois Geyer, Nikolaus Hautsch, Ravi Jagannathan, Gregor Kastner, Robert Korajczyk, Gyöngyi Lóránth, Albert Menkveld, Christoph Scheuch, Leopold Sögner, Günter Strobl, Viktor Todorov, Rossen Valkanov, Patrick Weiss, Christian Westheide, Kathy Yuan, Josef Zechner and seminar participants at Kellogg School of Management, the 4th Vienna Workshop on High-Dimensional Time Series in Macroeconomics and Finance 2019, the 5th Konstanz Lancaster Workshop on Finance and Econometrics and the VGSF Conference 2019 for helpful comments. I gratefully acknowledge financial support from the Austrian Science Fund (FWF project number DK W 1001-G16).

1 Introduction

In recent years, technological innovation and regulatory pressure spurred the emergence of many coexisting trading platforms in most asset classes. This market fragmentation may be beneficial for some participants, for instance due to increased competition between trading venues. However, it may also harm informational efficiency as long as limits to arbitrage prevent cross-market trading from equating prices across markets.[1] If frictions render arbitrage costly, arbitrageurs may refrain from exploiting price differences, thus giving up their pivotal role to restore the law of one price. Therefore, the speed and extent to which arbitrageurs exploit price differences depends on the magnitude of arbitrage costs and determines the informational value of prices in fragmented markets.

One particular asset that simultaneously exhibits substantial market fragmentation and considerable arbitrage costs is the market for Bitcoin, the most actively traded cryptocurrency to date. Whereas Bitcoin is traded worldwide without any trading pauses at hundreds of exchanges, persistent price differences across these exchanges consistently arise and cannot be reconciled solely with spreads or transaction costs.[2]

What distinguishes cross-market trading of Bitcoin from other assets (e.g., stocks or bonds) is the underlying decentralized settlement system. While traditional security markets are organized around trusted intermediaries, in decentralized settlement procedures no central clearing counterparty guarantees the ultimate delivery of the asset. Cross-market arbitrageurs are thus not able to dispose of their position until validators record the transaction on the blockchain. Hautsch et al. (2019) show that the associated

[1]Limits to arbitrage arise, for instance, as trading costs (Roll et al., 2007), holding costs (Pontiff (1996), Gagnon and Karolyi (2010)), constrained arbitrage capital (Shleifer and Vishny, 1997), or short sale constraints (Ofek et al., 2004). I refer to Gromb and Vayanos (2010) for an extensive survey of this literature.

[2]Makarow and Schoar (2019) document average daily price differences ratios between exchanges based in the US and Korea of 15%. Brauneis et al. (2019) find that the market for Bitcoin against USD is highly liquid in terms of bid-ask spreads but also document that markets are crossed most of the time. Hautsch et al. (2019) investigate 120 exchange-pairs and document average cross-market price differences of 63 basis points.

settlement latency resembles a novel technological friction that exposes arbitrageurs to the risk of adverse price movements and thus imposes limits to arbitrage.

In this paper, I analyze the implications of blockchain-related technological frictions on informational efficiency. I estimate the arbitrage costs for the Bitcoin market and show that they increase with settlement latency. This finding suggests that faster settlement reduces price risks for arbitrageurs and therefore improves informational efficiency. However, I also document that a lower settlement time (and hence lower limits to arbitrage) is associated with larger spreads. I embed this result in a theoretical model in which liquidity providers anticipate that arbitrageurs exploit their stale quotes more frequently if settlement time is low and thus set wider spreads to cope with the adverse selection risk. As a result, the direct effect of faster settlement on informational efficiency is offset by larger liquidity-related limits to arbitrage, close in spirit to Foucault et al. (2017).

The main econometric challenge to address the magnitude of technological frictions is that, in general, arbitrage costs or cross-market trading activities are not observable. Instead, the estimation of arbitrage costs requires an econometric model.

I provide a dynamic model for quoted prices of one asset at two markets. The framework rests on the assumption that arbitrageurs exploit price differences across both markets. If price differences occur, arbitrageurs aim to buy at the cheaper market and to sell at the more expensive market. As a result, price pressure from arbitrage activity enforces price adjustments at both markets towards the law of one price. Thus, price pressure from arbitrage capital (e.g., Ross (1976)) implies a cointegration relationship between the quotes at the two markets because price differences mean-revert towards the long-run equilibrium relationship (Engle and Granger (1987)).

This correction mechanism is suspended, however, whenever limits to arbitrage prevent profitable cross-market trading. In such a regime, price differences may persist and remain unexploited. As a result, adjustment of quotes towards the law of one price is non-linear in price differences and depends on arbitrage costs. More specifically, arbitrage costs determine a no-arbitrage regime during which arbitrageurs prefer to stay idle. From an econometric perspective, these considerations imply a threshold error correction model for quotes in which arbitrage costs determine the magnitude of the thresholds.[3] The estimated threshold defines the no-arbitrage regime and thus allows to back out arbitrage costs from the dynamics of quoted prices.

My econometric framework incorporates two important novel features: First, I parametrize the threshold as a function of (observed) proxies for market frictions such as quoted spreads. As a result, the no-arbitrage regimes can reflect time-varying arbitrage costs. Second, the parametrization allows me to decompose arbitrage costs into static latent (exchange-specific) and dynamic friction-specific features.[4]

I use the econometric model to estimate the no-arbitrage regime thresholds. Hereby, I exploit a large dataset of high-frequency orderbook snapshots of two of the largest cryptocurrency exchanges for trading Bitcoin versus US-Dollar from April 2018 until August 2019. The estimated arbitrage costs on average amount to 9 basis points. The estimated costs related to the time-consuming settlement process on a blockchain constitute 75% of the estimated threshold. To quantify the contribution of technological frictions to arbitrage costs, I use the number of transactions waiting for verification in the Bitcoin settlement network as a measure of network activity. Network activity varies considerably over time, where high activity increases the settlement latency of an individual transaction and therefore the price risk for arbitrageurs.[5]

I find that a one percent increase in network activity raises technology-related ar-

[3] Incorporating nonlinear adjustment processes to cointegrated variables goes back to, among others, Balke and Fomby (1997) and Hansen and Seo (2002).

[4] Ters and Urban (2019) propose a 3-regime threshold model to estimate latent arbitrage costs which are constant over time. Theissen (2012) incorporates quoted spreads into a similar regime model but does not estimate latent arbitrage costs.

[5] Arbitrageurs compete with other users of the Bitcoin network (e.g., motivated by consumption or private transactions) for limited settlement capacities. Blockchain congestion thus also increases arbitrage costs in terms of higher fees that provide validators incentives to prioritize the cross-market transaction (e.g., Easley et al. (2019)).

bitrage costs by 1.2 USD, whereas exchange-specific risks seem to play minor roles in preventing arbitrageurs from exploiting price differences. The increase of arbitrage costs due to higher network activity, however, is offset by a simultaneous adjustment of the quoted spreads. I find that an increase in network activity by one percent leads to a 0.4 USD decrease in spreads. This result is robust to controlling for trading volume and volatility.

I embed this finding in a theoretical framework that builds on two markets with asynchronous arrival of fundamental information about the same asset. Market fragmentation manifests itself in restrictions for some market participants to actively monitor and trade on information from both markets at the same time. Cross-market price differences occur if quotes on one of the two markets are based on concurrently outdated information. Whenever arbitrageurs observe profitable arbitrage opportunities and trade, they implicitly transmit information between the two markets. Two frictions impose limits to arbitrage: (i) exogenous technological frictions (e.g., settlement latency which may render arbitrage trades risky and therefore costly) and (ii) liquidity in terms of spreads which market makers set in anticipation of cross-market trading activity.

I derive quoted equilibrium spreads as a function of technology-related arbitrage costs and show that higher technology-related costs imply a lower adverse selection component in the spreads of the local market makers. Hereby, the adverse selection risk decreases because higher technology-related arbitrage costs reduce the likelihood of a profitable arbitrage opportunity. Consequently, quoted spreads are largest when costs due to technological frictions are absent. The overall effect of shifting technology-related arbitrage costs on informational efficiency, measured as the aggregate mispricing at both markets, is thus ambiguous: the change in spreads can even overcompensate the direct effect of technology-related costs such that a reduction (or even complete removal) of the technological friction *decreases* arbitrage activity and therefore also harms informational efficiency. Fragmentation of liquidity providers constitutes a friction on its own and implies arbitrage costs. Even more, in line with the empirical findings, the magnitude of the adverse selection related arbitrage costs depends inversely on the technology-related arbitrage costs.

In my sample, the economic magnitude of the adverse selection effect on informational efficiency is substantial: If network activity increases, technology-related arbitrage costs increase, but the adverse selection component in the spreads decreases, hampering the network effect on overall arbitrage costs by almost 30%. Therefore, the variation of the estimated total arbitrage costs is much smaller than the variation of the individual components. However, the decomposition reveals that, during periods of narrow spreads, market participants, who demand liquidity but are not arbitrageurs, benefit from the presence of technology-related arbitrage costs.

Overall, I argue that distributed ledgers impose a novel and economically significant technological friction for cross-market trading which differs substantially from the well-documented limits to arbitrage in markets for equities such as risk aversion (Green and Srivastava, 1985) or slow moving capital Mitchell et al. (2007). As blockchain-based settlement fundamentally differs from trading that involves (trusted) centralized clearing counterparties, there are still many unknowns when it comes to its microstructure implications. Abadi and Brunnermeier (2018) point out that blockchain-based settlement cannot simultaneously satisfy the demand for security, decentralization and cost efficiency and therefore centralized (trusted) intermediaries may dominate in some situations.[6] I provide a novel trade-off that is particularly relevant in the context of trading cryptocurrencies: due to the nature of the decentralized settlement process, fragmented markets cannot simultaneously achieve informational efficiency and narrow spreads. Instead, reducing or entirely removing the technological friction associated with decentralized settlement may

[6]For instance, Chiu and Koeppl (2019) estimate that the U.S. corporate debt market may benefit from blockchain-based settlement.

have unintended consequences in terms of local liquidity provision. Initiatives to reduce settlement latency should thus consider the extent to which fundamental information can become efficiently distributed across trading venues (see also, e.g., Foucault et al. (2003), Duffie and Manso (2007), Hagströmer and Menkveld (2019)).

Apart from the implications for blockchain-based trading, my paper also speaks to the effect of market frictions on informational efficiency and liquidity in other asset classes. Concerning the US equity markets, O'Hara and Ye (2011) conclude that Regulation National Market System Rule 611 fostered the coexistence of trading venues. At the same time, however, market consolidation fosters cross-market information dispersion. On the other end of the spectrum, European equity markets lack orderbook consolidation and substantial arbitrage costs might harm informational efficiency (see, e.g. Sagade et al. (2019)). Similarly, Benos et al. (2019) document arbitrage costs for interest rate swap contracts which are cleared both at the Chicago Mercantile Exchange (CME) and the London Clearing House (LCH). As a result, price differentials are costly to exploit because clearing house fragmentation imposes substantial collateral costs for dealers that are active at multiple markets.

Whereas it is well established that large spreads constitute limits to arbitrage (see, e.g. Stoll (1989)), the reverse direction – the interaction between technological frictions and adverse selection risks – received less attention.[7] However, it is important to understand this interaction to evaluate technological and regulatory changes that target cross-market trading activity, e.g., short-sale constraints or intentional latency delays (*speed bumps*). Crowding out cross-market liquidity takers provides market makers with the opportunity to set smaller spreads (see, e.g., Budish et al. (2015) and Brolley and Cimon (2019)).[8] Moreover, market-specific technological or regulatory changes that ham-

[7] Rösch (2013) empirically investigates adverse selection as a response to cross-market arbitrage activity for American Depositary Receipts (ADR).

[8] The implications of intentional latency delays are still under debate, see, e.g., Aldrich and Friedman (2017), Hu (2018), Woodward (2018) and Aoyagi (2018).

per liquidity-taking activities may harm price informativeness at other markets, imposing spill-over effects to overall market efficiency (see, e.g., Kyle and Xiong (2001)).

The structure of this paper is as follows: Section 2 theoretically studies the effect of technology-related arbitrage costs on liquidity provision. In Section 3, I provide the econometric framework to estimate arbitrage costs. In Section 4, I present the data, document price differences in the market for Bitcoin and provide the estimation results. Section 5 concludes.

2 Arbitrage and Liquidity Provision in Fragmented Markets

The following model establishes the baseline framework to analyze the effect of trading costs in fragmented markets on the trading activity of arbitrageurs and liquidity provision. In Section 3, I derive the quote dynamics implied by the theoretical framework as the basic pillar of the threshold error correction model.

2.1 Market structure and participants

I assume there are two markets i and j. One risky asset is traded on both markets simultaneously. The terminal value of the asset v_T is uncertain and revealed to all market participants at time T.[9]

Assumption 1. *The value of the asset v_t is a random variable which follows a Brownian motion*

$$v_t = v_0 + \int_0^t \sigma dW_s \tag{1}$$

where σ corresponds to the spot volatility and W_t denotes a Wiener process.

[9] I assume that the expected time until the terminal payoff realizes is long relative to the units of time in the setup below, similar in spirit to Baldauf and Mollner (2019).

Three groups of agents populate both markets: First, market participants with exogenous trading demand (noise traders) whose trading decisions are orthogonal to the asset value and therefore do not reveal any information to the market. Second, one arbitrageur who stands ready to exploit price differences between the two markets. The arbitrageur is the only participant able to trade at both platforms simultaneously. Third, at each market $k \in \{i, j\}$ there are competitive risk-neutral market makers specialized in trading the asset on their respective market.

Each market maker $k \in \{i, j\}$ continuously commits to buy or sell one unit of the asset at the prices she quotes, $\{a_t^k, b_t^k\}$, where a_t^k corresponds to the (ask) price at t at which market maker k is willing to sell, and b_t^k corresponds to the (bid) price at which she is willing to buy from sellers. Market makers determine quotes conditional on their (private) information regarding the asset value v_t.

I assume that the starting value v_0 at the initial date $t = 0$ is publicly known. For $t > 0$, however, new information about the value of v_t is only observable at randomly sampled discrete time points, not necessarily simultaneous on both markets.

Market makers update their beliefs regarding the terminal payoff if they receive new information. If an information event occurs on market k and time t, the current state of the price process, v_t, is revealed to the market makers at the respective market. For market participants on the other market, however, this information is not available in real-time. Instead, the arbitrageur is the only participant with the technology to monitor and act on both markets. Asynchronous information arrival resembles the core of market fragmentation in the theoretical model. It implies that informed investors are restricted in their access to multiple market venues and instead only act locally. [10] Fragmentation of information production can be an outcome of either geographical, regulatory or technical reasons. In the United States, for instance, fragmentation was an expected

[10] I do not provide a microfoundation for the actual process of information acquisition in this paper. Private information acquisition and how it is revealed through trading on local exchanges has been analyzed in depth (see, e.g. , Grossman and Stiglitz (1980) and Verrecchia (1982)).

outgrowth of Rule 611 (the trade-through rule) of Regulation National Market System (Reg NMS) (O'Hara and Ye (2011)). Asynchronous arrival of information may occur due to fragmented market participants (regional or in terms of trading motives) who either differ in terms of available information or required time to process information.[11] I relax the restricted monitoring capacities of market makers in the Appendix and I show that the results remain qualitatively similar when market makers are allowed to observe quote dynamics on the opposite market as long as some valuation uncertainty remains.

Information arrives on market k at time $\left\{0, t_1^k, \ldots, t_n^k\right\}$, whereas I denote the time between two information arrivals as $\tau_l^k := t_l^k - t_{l-1}^k$ for $l \in \{1, \ldots, n\}$. I put some structure on the sequence of random variables τ_l^k to obtain convenient analytical solutions in the following definition.

Assumption 2. *The sequence of information arrival times* $\left\{0, t_1^k, \ldots, t_n^k\right\}$ *follows a Poisson point process with parameter* λ_k*. Therefore, the inter-arrival times* $\{\tau_l^k\}_{l=1,\ldots,n}$ *are independent exponentially distributed variables with mean* $\mathbb{E}\left(\tau^k\right) = \frac{1}{\lambda_k}$ *and probability density function*

$$\pi\left(\tau^k\right) = \lambda_k \exp\left(-\lambda_k \tau^k\right). \tag{2}$$

Intuitively, if $\lambda_j > \lambda_i$, news arrive more frequently on market j. An alternative but equivalent interpretation of the information arrival process is the following: new information about the current state of v_t is revealed to the economy with inter-arrival times $\{\tau_1, \tau_2, \ldots, \tau_n\}$ which are exponentially distributed with parameter $\lambda := \lambda_i + \lambda_j$. If new information arrives, it is revealed only either to market i with probability $\frac{\lambda_i}{\lambda_i + \lambda_j}$ or to market j with probability $\frac{\lambda_j}{\lambda_i + \lambda_j}$. I provide a formal proof of the equivalence of this statement and the information arrival processes in Assumption 2 in the Appendix.

After an information event at time t^k, market maker k considers the signal v_{t^k} to update her quotes. By construction, the terminal payoff follows a semi-martingale. More-

[11] In that sense, markets may lead price informativeness in the spirit of Hasbrouck (1995).

over, at time t^k, the best predictor of market maker k of the terminal payoff at T is $v_{t^k}^k = \mathbb{E}(v_T|v_{t^k})$. Further, the valuation $v_{t^k}^k$ does not change until the next information event which takes places at (random) time $t^k + \tau^k$.

Competitive pricing at both markets implies that the quotes reflect the valuation $\mathbb{E}(v_T|v_{t^k})$ at all times. Therefore, if market maker k received her last signal at t^k, her quotes at t (where $t^k \leq t < t^k + \tau^k$) are

$$a_t^k = v_{t^k}^k + S_t^k \quad \text{and} \quad b_t^k = v_{t^k}^k - S_t^k, \tag{3}$$

where S_t^k denotes the quoted bid-ask (half) spread.

Order flow stems from noise traders that arrive continuously on both markets and can be distinguished from the arbitrageur ex-post. I assume their expected arrival rate in an marginal unit of time is $2\lambda_L dt > 0$. Upon arrival, liquidity traders buy or sell one unit of the asset at one of the two markets with equal probabilities, independent of the efficient price v_t.

Absent any cross-market trading, market makers in expectation do not loose anything and therefore do not require any compensation for providing liquidity. However, cross-market arbitrage leads to adverse selection risk due to the inability of market makers to cancel mispriced quotes before arbitrageurs exploit them (see, e.g. Budish et al., 2015). I assume that the arbitrageur is able to monitor quoted prices in real-time at both markets, similar to Foucault et al. (2017). Whenever quoted prices imply a profitable arbitrage opportunity, the arbitrageur buys at the market quoting the lower price, transfers the asset to the other exchange and sells. An arbitrage opportunity occurs if markets are crossed, e.g., the bid at one market exceeds the ask of the other market. However, transaction costs or execution risk may prevent the arbitrageur from exploiting price differences. Further examples for impediments to arbitrage include risk aversion or financial constraints.[12]

Transaction costs $c \geq 0$ prevent the arbitrageur from trading if and only if the cross-market difference between bid and ask is below this threshold. If $c = 0$, arbitrageurs exploit price differences as soon as markets are crossed. However, if $c > 0$, limits to arbitrage can arise in the sense that although markets are crossed, the arbitrageur refrains from trading. Therefore, the arbitrageur trades at time t if, for instance, the sell price on market i exceeds the buy price on market j such that $b_t^i - a_t^j > c$ where $t^j \leq t < t^j + \tau^j$ and $t^i \leq t < t^i + \tau^i$. The reverse case, $b_t^j - a_t^i \geq c$ can be handled analogously.

Definition 1. *The arbitrageur exploits any profitable cross-market price difference. I define a profitable arbitrage opportunity as any situation in which*

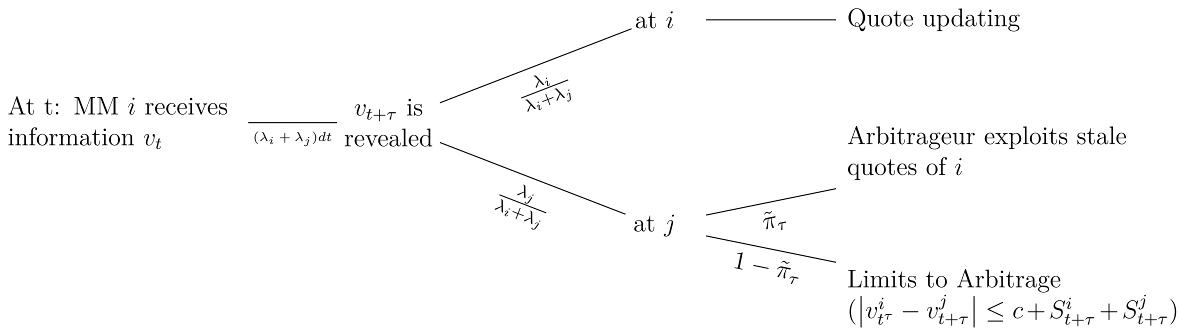
$$\left|v_t^i - v_t^j\right| > c + S_t^i + S_t^j \tag{4}$$

where $v_t^k = v_{t^k}$ *for* $t \in \left[t^k, t^k + \tau^k\right)$.

Therefore, arbitrage opportunities arise if the quotes are *stale* in the sense that, for instance, at t, the quotes of market maker i reflect current information but market maker j still offers quotes based on her last signal received at $t^j < t^i \leq t$. Such an event exposes the market maker with the risk of selling to (buying from) the arbitrageur an asset at a price which is too low (high). Although arbitrageurs do not collect information on their own and thus do not contribute to price discovery per se, they exploit mispricing fast and ensure that information is disseminated across markets, though, at the cost of wider spreads. I make the simplifying assumption that a trade by an arbitrageur resolves any information asymmetry and therefore also triggers a new valuation by the market maker. This assumption can be justified by direct competition or fast information revelation in relation to the time it takes to monitor and trade on cross-market price differences.

[12] For an extensive surveys on sources of limits to arbitrage I refer to Gromb and Vayanos (2010). In the empirical part of my paper I focus on price risk due to settlement latency which corresponds to risky arbitrage in the spirit of Bondarenko (2003).

Figure 1: Information revelation and decision making. This figure illustrates the major elements of the theoretical framework: After market maker i receives information at time t, the next information takes place either at the same exchange or market makers i set of information becomes stale. The cross-market arbitrageur exploits the potential price difference if and only if the trade resembles a profit, which happens with probability $\tilde{\pi}_\tau$ and depends on the decision of the market maker and characterizes the equilibrium conditions.



2.2 Equilibrium spreads

Next, I derive the equilibrium spreads at the fragmented markets in presence of an arbitrageur as a function of the trading costs c and the information arrival rates λ_i and λ_j. Figure 1 illustrates the relevant elements of the theoretical framework. Assume that a news event occurs on market i at time $t = t^i$. By Assumption 2, the time until the next information event, τ, is exponentially distributed with parameter $\lambda := \lambda_i + \lambda_j$ and expected inter-arrival time $\mathbb{E}(\tau) = \frac{1}{\lambda}$. First, during the time interval $[t, t+\tau]$ market maker i receives the order flow only from the noise traders with arrival intensity λ_L. At time $t+\tau$, new information arrives on one of the two markets and the corresponding market makers update their quotes.

I define the change of the signal regarding the terminal payoff of the asset during the time period $[t, t+\tau]$ as $\delta_{t,\tau} := v_{t+\tau} - v_t$. If information arrives for example on market j at time $t+\tau$ and $\delta_{t,\tau} > 0$, the mid-quotes on market$i$ will be too low. However, Definition1 implies that the arbitrageur trades only if the payoffs also exceed the trading costs, thus if $|\delta_{t,\tau}| > S^i_{t+\tau} + S^j_{t+\tau} + c$ as in Equation (4). In the particular example, the arbitrageur buys on market$i$, transfers the asset and sells on market$j$. The payoffs and the trading strategy are reversed if, instead, $\delta_{t,\tau} < 0$. Upon arrival of news at market j, adverse selection does only affect the opposite market maker i. From the perspective of the individual market maker, there is no threat of quoting outdated prices at the time when she receives information. Subsequently, at t^k, competitive spreads of market maker k are zero. Only the adverse selection component in the spreads of their own markets is relevant for the market makers when it comes to their equilibrium spreads.

In the following Lemma, I derive the probability of an arbitrage trade from the perspective of market maker $k \in \{i, j\}$, τ units of time after she updated her quotes for the last time.

Lemma 1. *Given Assumptions 1 and 2, the probability of an arbitrage trade at time* $t^k + \tau$ *is* $\pi_\tau^k \left(S_{t+\tau}, c, \sigma\right) := \mathbb{P}\left(|v_{t^k+\tau} - v_t| > S_{t+\tau} + c \mid v_t = v_t^k\right)$ *is*

$$\tilde{\pi}_\tau^k \left(S_{t+\tau}, c, \sigma\right) = 1 - \frac{2}{\sqrt{\pi}} \int_0^{\frac{c+S_{t+\tau}}{\sigma\sqrt{2\tau}}} e^{-z^2} dz. \tag{5}$$

Further, $\tilde{\pi}_\tau^k \left(S_{t+\tau}, c, \sigma\right)$ *exhibits the following characteristics*

$$\frac{\partial \tilde{\pi}_\tau^k}{\partial \tau} > 0, \quad \frac{\partial \tilde{\pi}_\tau^k}{\partial S_{t+\tau}} < 0, \quad \frac{\partial \tilde{\pi}_\tau^k}{\partial c} < 0, \quad \frac{\partial \tilde{\pi}_\tau^k}{\partial \sigma} > 0. \tag{6}$$

Proof. See Appendix. □

Large price changes $|\delta_{t,\tau}|$ are more likely if the time interval since last arrival of information, τ, is long, or if the volatility of the price process, σ is high. In particular, Assumption 1 and the random arrival rate of new information both imply that the unconditional volatility of the change of the value, $\mathbb{V}\left(|\delta_{t,\tau}|\right)$ is $\sigma_v := \sqrt{\frac{\sigma^2}{2\lambda}}$. As a consequence, the likelihood of a profitable arbitrage opportunity that exceeds the threshold $S_{t+\tau} + c$ increases with volatility and the expected waiting time, λ^{-1}.

Lemma 1 shows that the total trading costs, $S_{t+\tau} + c$, affect the activity of the arbitrageur in a straightforward manner: The probability that the valuation difference $|\delta_{t,\tau}|$

exceeds the threshold that makes the arbitrage opportunity profitable decreases with the quoted spread $S_{t+\tau}$ and c. Whereas c is exogenous, market makers control the spread based on their knowledge about the total trading costs paid by the arbitrageur at both markets. A higher spread makes it less likely that the change of v_t during the time interval $[t, t+\tau]$ exceeds the boundaries implied by the arbitrageurs participation constraint as of Definition 1. The extreme case of the spread $S_{t+\tau}$ approaching infinity corresponds to a zero probability event of arbitrageurs' activity. At the other extreme, even if the market marker quotes a zero spread, the presence of transaction costs, $c > 0$, imply a non-trivial probability that the absolute differences in valuation, $|\delta_{t,\tau}|$ do not exceed costs c.

Suppose now an information event occurs at time $t+\tau$. The price process is revealed with probability $\frac{\lambda_i}{\lambda_i+\lambda_j}$ on market i or with probability $\frac{\lambda_j}{\lambda_i+\lambda_j}$ on market j. In the first case, market maker i updates her quotes and shifts v_t^i correspondingly. As discussed above, this scenario does not leave any uncertainty for market maker i, and trading against the arbitrageur does not expose her to any adverse selection risk. In the second case, market maker j updates her quotes whereas market maker i is not fast enough to react. If the arbitrageur does not get active because the differences in valuation do not offset the quoted spreads, neither market maker i nor j earn or lose anything and price differences persist. However, with probability $\tilde{\pi}_\tau^i\left(S_{t+\tau}, c, \sigma\right)$, the arbitrageur exploits the price difference. Market maker i earns the spread but trades against the arbitrageur at stale quotes and her expected losses conditional on an arbitrage trade amount to $S_{t+\tau}^i - \mathbb{E}\left(|\delta_{t,\tau}| \middle| S_{t+\tau}^i + c \leq |\delta_{t,\tau}|\right)$. The following lemma summarises the expected profits of market maker i during the period of time $[t^i+\tau, t^i+\tau+dt)$ for small dt.

Lemma 2. *Under assumptions 1 and 2, the expected profits of market maker i at* $d(t^i+\tau)$

with spread S are

$$\mathbb{E}\left(\Pi_{i,t+\tau}\left(S\right)\right) = \underbrace{\lambda_L S}_{\substack{\textit{Profit from}\\ \textit{noise trading}}} + \frac{\lambda_j}{\lambda_i + \lambda_j}\left(\tilde{\pi}_\tau S - \underbrace{\sigma\sqrt{\frac{2\tau}{\pi}}\exp\left(-\frac{\left(S+c\right)^2}{2\tau\sigma^2}\right)}_{\textit{Expected loss from stale quote}}\right). \tag{7}$$

The expected profit of market maker i at $d(t+\tau)$ exhibit the following characteristics:

$$\frac{\partial\mathbb{E}\left(\Pi_{i,t+\tau}\left(S\right)\right)}{\partial\lambda_L} > 0, \qquad \frac{\partial\mathbb{E}\left(\Pi_{i,t+\tau}\left(S\right)\right)}{\partial\lambda_i} > 0, \quad \frac{\partial\mathbb{E}\left(\Pi_{i,t+\tau}\left(S\right)\right)}{\partial\lambda_j} < 0,$$
$$\frac{\partial\mathbb{E}\left(\Pi_{i,t+\tau}\left(S\right)\right)}{\partial\tau} < 0, \qquad \frac{\partial\mathbb{E}\left(\Pi_{i,t+\tau}\left(S\right)\right)}{\partial\sigma} < 0, \quad \frac{\partial\mathbb{E}\left(\Pi_{i,t+\tau}\left(S\right)\right)}{\partial c} > 0.$$

Proof. See Appendix. □

The expected gains from trading against the arbitrageur are always negative. As trading takes place only if the difference in the valuation exceeds at least the spreads, the losses conditional on this event always exceed the spread because

$$S^k_{t+\tau} + c \leq E\left(|\delta_{t,\tau}|\middle|S^k_{t+\tau} + c < |\delta_{t,\tau}|\right). \tag{8}$$

Expected profits can only become non-negative, if $\lambda_L > 0$ because noise trades compensate the market maker for providing liquidity even in anticipation of trading against the arbitrageur.

A higher spread S increases expected profits due to higher expected gains from trading with liquidity traders and it reduces the likelihood of arbitrage activity as shown in Lemma 1. Total trading costs, modelled as $S^k_{t+\tau} + c$, reduce the likelihood of the arbitrageur becoming active but they also increase the expected loss of market maker k conditional on arbitrageurs trading. Lemma 2 reflects that if information arrives on one market, profits would be strictly positive for $S^k_{t+\tau} > 0$ at that point in time ($\tau = 0$), due to the absence of any asymmetric information. Therefore, local competition forces market makers to set their spreads to zero at the time of information arrival.

From the perspective of market maker k at time $t^k + \tau$ it is uncertain, if the next information event will occur at her market or if she will be exposed to stale quote trading. Consequently, large λ_k decreases the likelihood of an (adverse) information event and therefore decreases expected losses of market maker k. Conversely, if the probability of information arrival on the opposite market increases, the expected profits decrease.

Next, I characterize the equilibrium spreads $\tilde{S}^k_{t+\tau}$. Equilibrium is characterized by both market makers setting their spreads such that at every point in time, the expected profits at $d(t^k+\tau)$ are zero due to local competition. Further, in equilibrium, arbitrageurs mechanically exploit price differences whenever profitable according to participation constraint in Definition 1.

Definition 2. *In equilibrium for $\lambda_k > 0$, after an information arrival on market k at time t^k, the competitive spread of market maker k at time $t^k + \tau$ is the maximum of zero and the unique root of the equation*

$$\mathbb{E}\left(\Pi_{k,t^k+\tau}\left(\tilde{S}_{t^k+\tau}\right)\right) = 0. \tag{9}$$

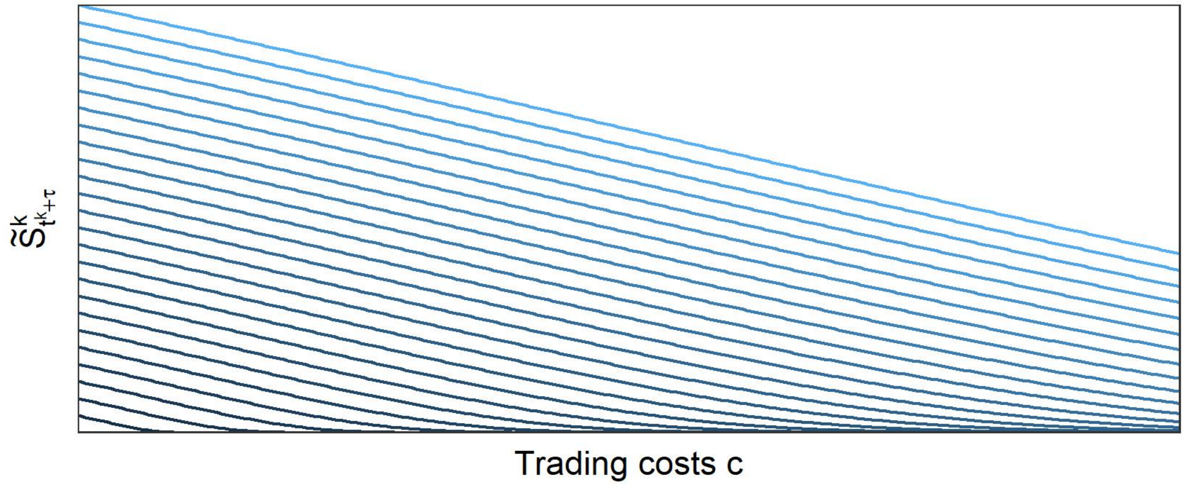
Equilibrium spreads $\tilde{S}^k_{t^k+\tau}$ are the minimum required spreads such that market makers earn zero expected profits. Requiring higher spreads is not a feasible solution due to competition among market makers on market k. Reversely, less compensation than $\tilde{S}^k_{t^k+\tau}$, both increases the likelihood of arbitrage trading and reduces profits from uninformed order flow.

By Lemma 2, the equilibrium spread is strictly positive but diverges to infinity for $\lambda_L \to 0$ which is in line with the common notion of the non-existence of an equilibrium in absence of belief dispersion (Grossman and Stiglitz (1980)). The arrival rate of noise traders determines the relation between losses due to trading activity from arbitrageurs and gains from trade without any subsequent price adjustment. For $\lambda_L \to \infty$, instead, the market maker is able to compensate her losses due to the high arrival rate of liquidity traders and equilibrium spreads $\tilde{S}^k_{t^k+\tau}$ converge towards zero (similar to vanishing price impact as described by Kyle (1985)). Further and in line with well-established results, equilibrium spreads increase with volatility σ (Easley and O'Hara (1987)).

The arrival rate of information, $\lambda = \lambda_i + \lambda_j$, constitutes an important parameter, both for the equilibrium spreads but also for the informativeness of quoted prices in general. Faster information arrival decreases the volatility σ and therefore reduces the adverse selection risk. However, from the perspective of the individual market maker, the probability of information arrival at the own market $\frac{\lambda_k}{\lambda_j+\lambda_i}$ is the relevant measure of adverse selection risk.

Finally, Figure 2 illustrates the trade-off between trading costs, c, and the endogenous adverse selection component of the spreads, $\tilde{S}^k_{t^k+\tau}$. Keeping everything else equal, higher costs c reduce the likelihood of an arbitrage event $\tilde{\pi}^k_\tau$ and reduce the threat of stale quote trading. However, conditional on an arbitrage event, higher costs c also increase the expected losses due to trading for the market maker. As characterized in Lemma 1, an increase of the trading costs c increases the expected profits of the market maker (and therefore the equilibrium spreads become smaller). Figure 3 illustrates the trade-off between a shift of trading costs, c, and the total costs for the arbitrageur, $\tilde{S}^i_t(c)+\tilde{S}^j_t(c)+c$. The figure shows the gradient of the total costs for the arbitrageur as a function of the trading costs, c. Increasing the trading costs c has two effects on the arbitrageurs' activity: first, it directly increases total trading costs as it shifts the minimum price difference required before the arbitrageurs prefers to trade. Second, the endogenous component of the costs decreases due to the reduced spreads in response to lower threat of adverse selection. The figure illustrates that regions exist for which the total costs of the arbitrageur even decrease if c increases. Intuitively, this seemingly puzzling result

Figure 2: Equilibrium spreads. This figure illustrates equilibrium spreads as a function of the exogenous limits to arbitrage c, keeping everything else equal. The different lines correspond to shifted values of the volatility σ_v of the efficient price process. Brighter lines denote higher volatility.



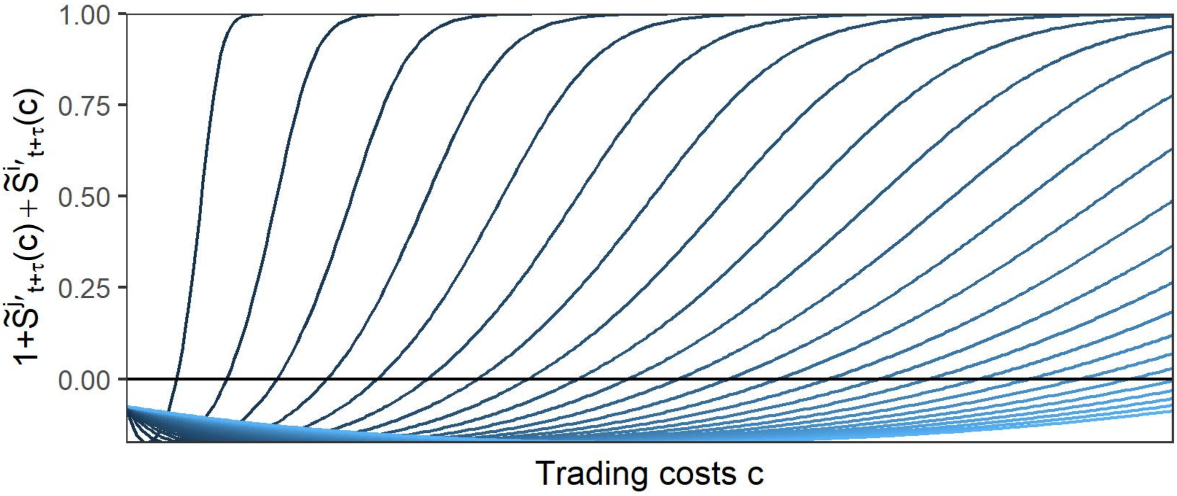
occurs when a marginal increase in c substantially decreases the probability of trading and thus simultaneously relaxes the zero expected profit constraints of the individual market makers.

2.3 Price informativeness

Quoted prices can deviate from the efficient price for two reasons. On the one hand, the efficient price process v_t is observable only at infrequent points in time. On the other hand, information is revealed asynchronous due to absence of (profitable) arbitrage opportunities.

The frequency of information arrival events, $(\lambda_i + \lambda_j)^{-1}$ determines the aggregate deviation from quoted prices from the underlying efficient price process. Arbitrageurs' activity, which depends on c and $\tilde{S}_{t+\tau}$, in fact only facilitates cross-market information aggregation and enforces price informativeness by updating prices across markets. Figure 4 illustrates both determinants of mispricing. All three panels are based on a simulated time series. The grey line corresponds to the latent efficient price process and the red

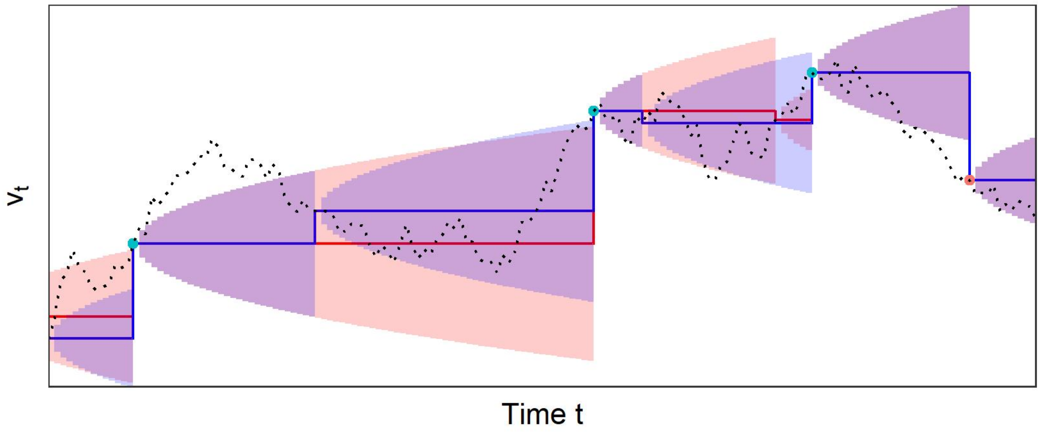
Figure 3: Gradient of total trading costs. This figure illustrates the effect of a marginal increase of the trading costs, c on the total costs for the arbitrageur $\tilde{S}^i_t(c) + \tilde{S}^j_t(c) + c$. The different lines correspond to shifted values of the volatility σ of the efficient price process. Brighter lines denote larger volatility.



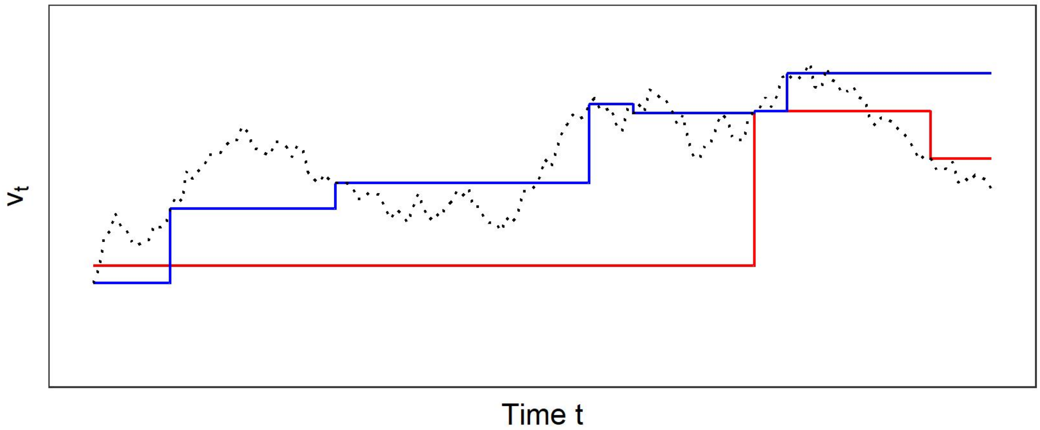
and blue lines denote the midquotes at both markets. The shaded areas correspond to the quoted spreads at both markets. The efficient price process evolve as a Brownian motion in line with Assumption 1 and information is revealed with independent exponential distributed waiting times as in Assumption 2. Jumps in the lines correspond to information arrivals and illustrate updated information sets of the market makers. During the remaining time no additional information is available regarding the current value of v_t, constituting one source of mispricing. Panel A illustrates the case without any trading costs , i.e. $c = 0$. Whereas spreads increase with the waiting time since the last information arrival event, price differences are bounded within narrow intervals as the arbitrageur continuously monitors and eventually performs cross-market trades. On the contrary, Panel B of Figure 4 corresponds to the case with prohibitively high trading costs. In this case the adverse selection component in the spreads is negligible because the probability of an arbitrage trade $\tilde{\pi}_\tau \left(\tilde{S}^k_{t^k+\tau}, c, \sigma\right)$ is close to zero. Price differences, however, can persist within wider bands. In fact, in the most extreme case both markets are entirely decoupled and quotes are only updated with (market-specific) intensities λ_i

Figure 4: Simulated Price Paths. This figure shows three outcomes based on one simulated Wiener process v_t (black dots). Time is plotted on the x-axis. Information inter-arrival times are exponentially distributed. Information is revealed at one of the two markets with the same probability. The blue (red) line correspond to the quoted mid-prices of the two markets. The shaded area corresponds to the corresponding (equilibrium) spreads. Arbitrage trades are indicated with green dots.

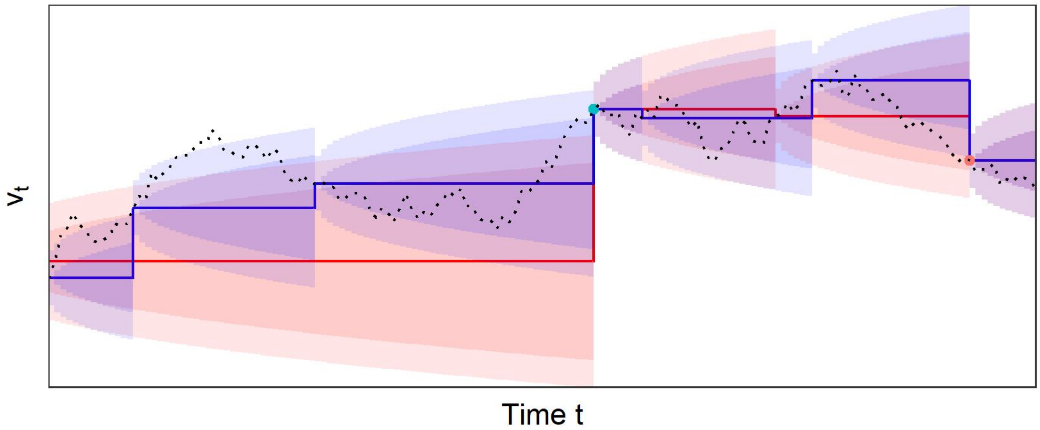
Panel A: No trading costs ($c = 0$).



Panel B: Absence of arbitrageur ($c \to \infty$).



Panel C: Intermediate trading costs ($0 < c < \infty$).



and λ_j. Panel C illustrates the intermediate case.

I analyze the aggregate level of price informativeness as the unconditional expected mispricing prevalent in the aggregated market which I define in the following proposition.

Proposition 1. *Under Assumption 1 and Assumption 2 the expected (L_1-) error of the aggregated quoted prices is*

$$\mathbb{E}\left(\left|v_t - v_t^i\right| + \left|v_t - v_t^j\right|\right) = \sqrt{\frac{\sigma^2}{2}\mathbb{E}\left(\tau\right)\Psi\left(c\right)} \tag{10}$$

where $\mathbb{E}\left(\tau\right) = \frac{1}{\lambda_i+\lambda_j}$ *and* $\Psi\left(c\right) := \sqrt{\frac{1+\lambda_j/\lambda_i}{1+\mathbb{E}\left(\tilde{\pi}_\tau^i\left(\tilde{S}^i_{t^i+\tau},c,\sigma\right)\right)\frac{\lambda_j}{\lambda_i}}} + \sqrt{\frac{1+\lambda_i/\lambda_j}{1+\mathbb{E}\left(\tilde{\pi}_\tau^j\left(\tilde{S}^j_{t^j+\tau},c,\sigma\right)\right)\frac{\lambda_i}{\lambda_j}}}$.

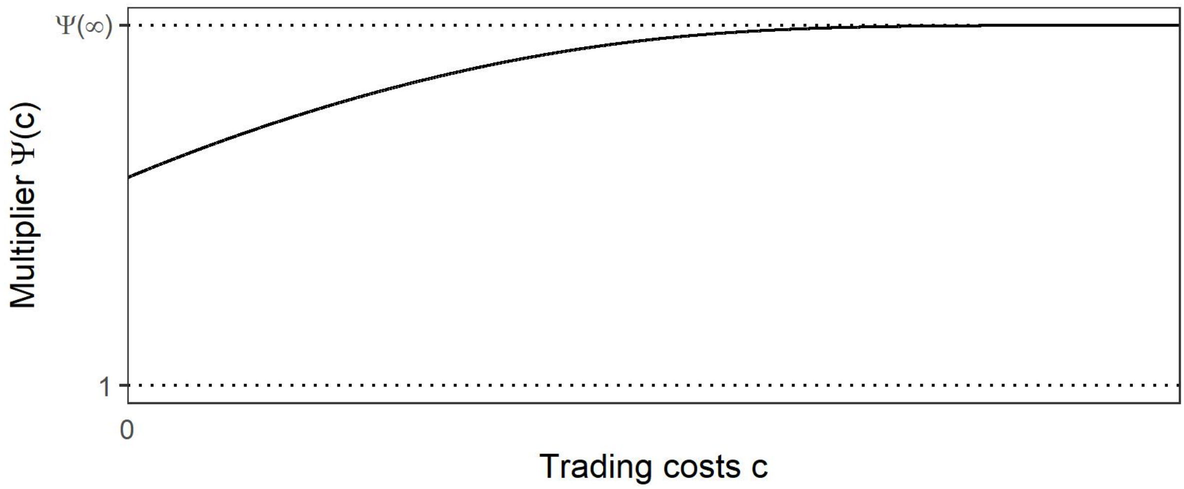
Proof. See Appendix. □

Proposition 1 illustrates the fundamental trade-off between price informativeness and endogenous spreads in a setting with exogenous cost of trading: First, the expected pricing error is always positive and increases with the volatility of the efficient price process σ and the expected inter-arrival times $\mathbb{E}\left(\tau\right)$. Both components increase the uncertainty with respect to the true value. The role of the arbitrageur can be understood as increasing the speed with which information is reflected at the individual markets. In the case of prohibitive high trading costs ($c \to \infty$) arbitrageurs never trade ($\tilde{\pi}^k = 0$) and the arrival rates of information at both market do not change. In that case $\lim_{c\to\infty} \Psi(c) = \left(\sqrt{\frac{\lambda_i+\lambda_j}{\lambda_i}} + \sqrt{\frac{\lambda_i+\lambda_j}{\lambda_j}}\right)$ and the expected pricing errors are

$$\lim_{c\to\infty} \mathbb{E}\left(\left|v_t - q_t^i\right| + \left|v_t - q_t^j\right|\right) = \sqrt{\frac{\sigma^2}{2}}\left(\sqrt{\mathbb{E}\left(\tau_i\right)} + \sqrt{\mathbb{E}\left(\tau_j\right)}\right). \tag{11}$$

On the contrary, if trading costs c are smaller, the probability of an arbitrage event on market i given information arrived at the opposite market may become positive ($\tilde{\pi}_\tau^i > 0$). In that case, the arrival rate of information on market i increases to $\tilde{\lambda}_i := \lambda_i + \mathbb{E}\left(\tilde{\pi}^i\right)\lambda_j$. The similar effect holds for market j such that $\tilde{\lambda}_j := \lambda_j + \mathbb{E}\left(\tilde{\pi}^j\right)\lambda_i$. Due to the adverse

Figure 5: Mispricing Multiplier. This figure illustrates the relation between the mispricing multiplier $\Psi(c)$ and the exogenous trading costs c as of Proposition 1.



selection component in the spreads the probability of an arbitrage event $\tilde{\pi}_\tau^k$ will never reach one even if trading costs c are 0.

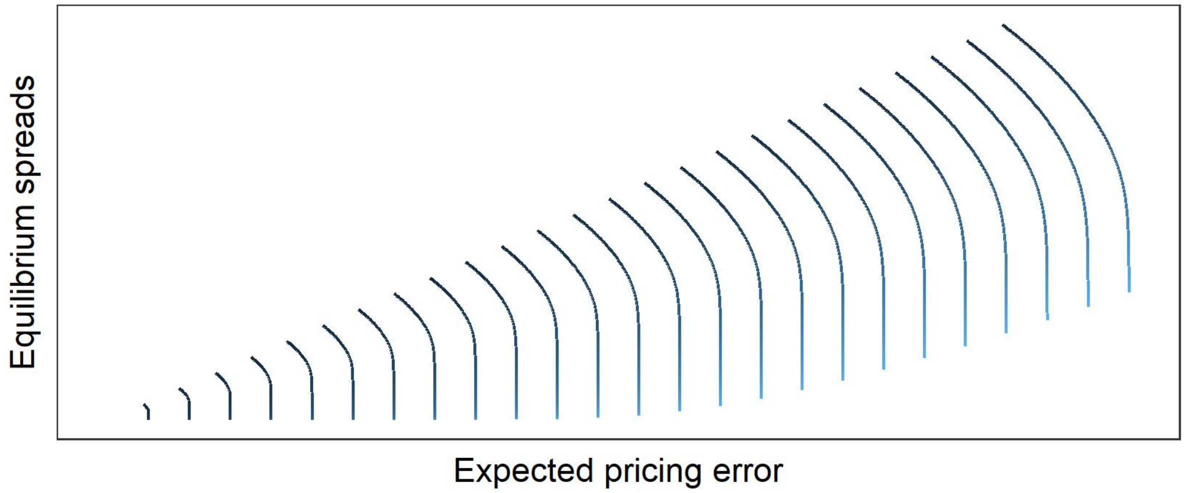
Removing the friction of fragmentation and instead allowing all participants to trade on a consolidated orderbook would make cross-market trading superfluous and could be interpreted as $\mathbb{E}\left(\tilde{\pi}_\tau^k\right) = 1$ which would result in the highest attainable value of price informativeness with

$$\mathbb{E}\left(\left|v_t - q_t^i\right| + \left|v_t - q_t^j\right|\right) \overset{\mathbb{E}\left(\tilde{\pi}_\tau^k\right)=1}{=} \sqrt{2\sigma^2 \mathbb{E}\left(\tau\right)}. \tag{12}$$

Figure 5 shows the value of the multiplier $\Psi(c)$ for intermediate values of c, resembling the case with exogenous trading costs and endogenous adverse selection component in the spreads. Without exogenous trading costs, endogenous limits to arbitrage induce mispricing. For increasing values of c, $\Psi(c)$ gradually converges towards entirely fragmented markets.

As a final step of the analysis, the theoretical framework allows to compare the joint equilibrium outcomes of liquidity (measured in the endogenous component of the spreads) and price informativeness (in terms of the inverse of expected aggregate pricing errors). As discussed in Lemma 2, an increase in the trading costs, c, reduces the adverse selection component of the spreads as the reduced probability of arbitrage trading relaxes the zero expected profits constraint of the market makers. On the other hand, price differences remain unexploited if trading costs are large and reduce the speed of information revelation across markets. As spreads increase with vanishing trading costs, market markers' anticipation of the arrival of arbitrageurs is already sufficient to generate limits to arbitrage in terms of higher total trading costs, $c + \tilde{S}_t^i + \tilde{S}_t^j$. Therefore, market fragmentation harms informational efficiency because the rate of information arrival is smaller than the hypothetical equivalent of integrated markets, i.e., $\tilde{\lambda}_k < \lambda_i + \lambda_j$. Figure 6 shows equilibrium spreads and pricing errors as a function of the trading costs c and shows that price discovery (in terms of lower expected aggregate pricing errors) can only be achieved if arbitrageurs' activity is high. The figure shows expected pricing errors on the x-axis and equilibrium spreads on the y-axis. Each line corresponds to the equilibrium values based on different values of the trading costs, c, keeping everything else equal. Bright colour corresponds to very high trading costs, dark colour indicates low values of c. The figure shows that for high trading costs (at the lower right hand side) the expected pricing error is generally higher than for lower values of c because the arrival rate of new information reaches its minimum, $\frac{1}{\lambda_i} + \frac{1}{\lambda_j}$. The corresponding equilibrium spreads, however, are lowest for high values of c due to the reduced threat of adverse selection. Decreasing c (at the upper left hand side) increases arbitrage activity, thus widens the quoted spreads of the market makers. Further, the expected pricing error decreases because $\Psi(c)$ converges to its lower limit. Note, that the end points of the lines at the left tails correspond to the extreme case with no trading costs, c. Here, total costs of the arbitrageur, $\tilde{S}_t^i + \tilde{S}_t^j$ arise entirely endogenous and therefore limit the expected pricing error from below. The different lines in the figure correspond to shifts in the volatility σ of the efficient price

Figure 6: Equilibrium Price Informativeness and Liquidity. This figure summarizes the (sub)-space of equilibrium outcomes for the expected pricing error and the equilibrium spread as functions of trading costs, c. The different lines correspond to shifted values of the volatility σ of the efficient price process. Colors denote the trading costs, c, whereas brighter colors correspond to larger values. Pricing error denotes the expected L_1 norm of aggregate mispricing according to Proposition 1 and the equilibrium spreads are according to Definition 2.



process. The lines to the left represent small volatilities which have two effects: first, smaller uncertainty regarding the efficient price corresponds to less adverse selection and smaller spreads and second, the expected pricing error decreases.

2.4 Settlement latency as a specific example for c

My empirical analysis rests on trading costs in terms of risky arbitrage payoffs due to settlement latency. In this section, I show that the time it takes to transfer an asset between two markets, implies costs for the arbitrageur which can be interpreted as trading costs c in the framework above.

Settlement latency is a market friction which arises, for instance, for any asset where legal change of ownership is recorded on a blockchain (Hautsch et al., 2019). Settlement latency prevents the arbitrageur from selling instantaneously, as the transfer of assets to the more expensive market and subsequent sale is only possible with a certain delay.

Latency $\tau_a > 0$ is the (possibly random) waiting time until settlement occurs and affects the relative speed of the arbitrageur. The longer the waiting time, the higher the likelihood that the market maker updates her quotes and therefore the arbitrage opportunity disappears.

Assumption 3. *Settlement latency $\tau_a > 0$ is the time it takes to transfer an asset between the two markets. An asset can be sold at a market only, if the arbitrageur is in possession of this asset, thus, only if the legal change of ownership has been completed.*

Consider for now the following example: new information arrives on market i at time t. Settlement latency then implies that during the time period $[t + \tau_a)$ the market maker on market i is not at risk of quoting stale quotes and trading against an arbitrageur at outdated prices. Furthermore, at time $t + \tau$ where $\tau > \tau_a$, the market maker knows that an arbitrageur may have exploited price differences which occurred $\tau - \tau_a$ periods after she updated her quotes last. Therefore, the implied volatility of price changes from the perspective of the market maker, is $\mathbb{V}\left(|v_{t+\tau_a} - v_t|\right) = \sqrt{\sigma^2 (\tau - \tau_a)}$ which is lower than it would have been without settlement latency ($\tau_a = 0$). The following lemma shows that from the perspective of the market maker, settlement latency can be interpreted as giving her a time advantage relative to the arbitrageur:

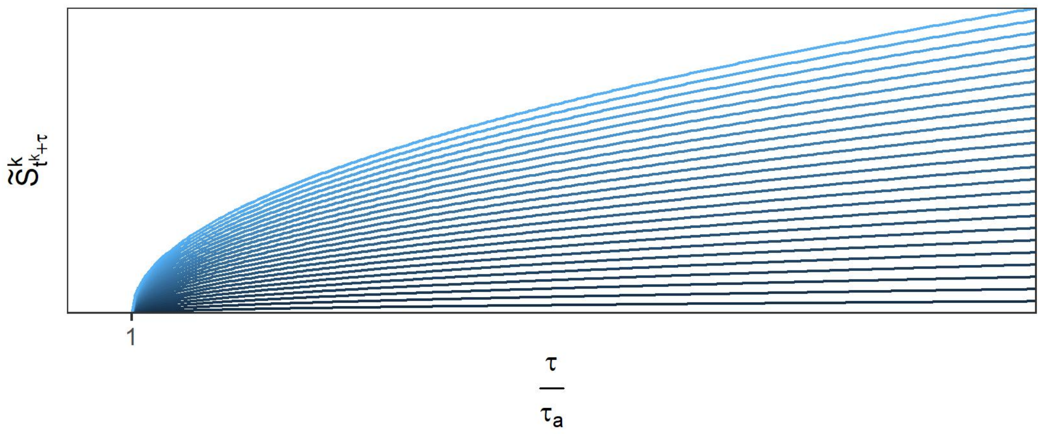
Lemma 3. *Under assumptions 1, 2 and 3 at $\tau > \tau_a$, the expected profits of market maker i at $d(t + \tau)$ are*

$$\mathbb{E}\left(\Pi^s_{i,t+\tau}(S)\right) = \lambda_L S + \frac{\lambda_j}{\lambda_i + \lambda_j}\left(\tilde{\pi}^s_{\tau-\tau_a} S - \sigma\sqrt{\frac{2(\tau-\tau_a)}{\pi}}\exp\left(-\frac{S^2}{2(\tau-\tau_a)\sigma^2}\right)\right). \tag{13}$$

Here, $\tilde{\pi}^s_{\tau-\tau_a} = 1 - \frac{2}{\sqrt{\pi}}\int_0^{\frac{S}{\sigma\sqrt{2(\tau-\tau_a)}}} e^{-z^2} dz$. For $\tau \leq \tau_a$,

$$\mathbb{E}\left(\Pi^s_{i,t+\tau}(S)\right) = 0 \ \Leftrightarrow S = 0. \tag{14}$$

Figure 7: Settlement Latency and Equilibrium Spread. This figure shows the effect of settlement latency τ_a on the equilibrium spreads. The case with no latency is equivalent to the benchmark case in Lemma 2 with $c = 0$. Brighter lines denote larger volatility.



Further, $\frac{\partial \mathbb{E}\left(\Pi^s_{i,t+\tau}(S)\right)}{\partial \tau_a} > 0$.

Proof. See Appendix. □

Lemma 3 takes into account that the (risk-neutral) arbitrageur upon observing a price difference at time t does only trade if expected profits are positive. This is the case, if $|\delta_{t,\tau}| > \mathbb{E}\left(S_{t+\tau+\tau_a}\right) \geq S_{t+\tau}$. Abstracting from any other costs for the arbitrageur, settlement latency imposes limits to arbitrage in the sense that price differences are exploited if and only if the hypothetical instantaneous returns exceed the (latency-adjusted) spreads.

Figure 7 illustrates the effects of settlement latency on the quoted spreads: the blue line shows the equilibrium spreads for different values of τ_a. Higher values indicate longer expected waiting times for the arbitrageur and correspond to higher price risk for the arbitrageur.

3 Threshold Vector Error Correction Model

From an empirical perspective, it is a challenge to quantify the trading costs for arbitrageurs which are generally not observable. Whereas often potential impediments to arbitrage can be identified, e.g., short selling restrictions or settlement latency, the magnitude and economic relevance of these frictions remains opaque. Further, the theoretical framework shows that the overall effect of a change in trading costs for arbitrageurs is diluted by the corresponding adjustments to the spreads.

In principle, one could try to draw indirect inference about arbitrageurs' activity by evaluating potential portfolio holdings. Recent literature inferred the long side of arbitrage trading by investigating hedge fund stock holdings (Brunnermeier and Nagel, 2004), the short side by investigating short-selling activity on stocks (Hanson and Sunderam, 2014) or combined both trading directions to examine net arbitrage trading (Chen et al., 2019). Instead, I utilize the implications of my theoretical model to identify otherwise latent trading costs c to quantify the extent to which arbitrageurs refrain from exploiting price differences due to market frictions.

I derive a structural vector error correction model for observed quoted prices based on the notion that cross-market arbitrage trading enforces the law of one price. Price pressure from arbitrage trades implies mean-reversion towards the efficient price. However, trading costs for arbitrageurs impose a threshold for price differences below which arbitrageurs prefer to stay idle instead of trading away price differences. Empirically, this strategic behaviour implies a three regime threshold model for the time series of price differences which allows to identify latent trading costs (see, e.g., Ters and Urban (2019)). The underlying theory is close in spirit to applications on exchange rates (Lo and Zivot, 2001), commodity markets (Park et al. (2007) and Stevens (2015)) and trading of futures (Dwyer et al. (1996) and Forbes et al. (1999)). However, I extend the framework to allow for time varying trading costs in response to shifts in potential proxies for the trading costs. Time-varying thresholds constitute one of the main differences of my framework compared to, among others, Balke and Fomby (1997) and Ters and Urban (2019). Whereas Theissen (2012) proposes to include time-varying liquidity in the regime specification, he abstracts from estimating latent trading costs. Dwyer et al. (1996) instead, do not allow for asymmetries depending on the direction of the trade.

More specifically, I exploit time-variation of a proxy for a particular friction and attribute its contribution to the trading costs for arbitrageurs. Then, the econometric model allows to back out c and the sensitivity of c with respect to the proxy using observed data.

3.1 Econometric Model

Although the efficient price v_t is not observable, the central law of one price defines a cointegration relationship of the midquotes. More specifically, information revelation at the two fragmented markets implies that the midquotes at time t reflect the (possibly) stale belief of the market makers about the current value of v_t. In light of my theoretical framework, midquote price differences between market i and market j, however, are stationary even in absence of any arbitrage trading because the following holds as of Assumption 1 for $\lambda_i > 0$ and $\lambda_j > 0$:

$$z_t := v_t^i - v_t^j = \left(v_t^i - v_t\right) - \left(v_t^j - v_t\right) = \int_{t^i}^{t} \sigma dW_s - \int_{t^j}^{t} \sigma dW_s = \pm \int_{t^i}^{t^j} \sigma dW_s \sim I(0). \quad (15)$$

Equation (15) confirms that the law of one price exhibits a valid cointegration relationship as long as the distribution of $t^i - t^j$ is stationary. More specifically, the distribution of the midquote price differences evolves as a random walk within information arrivals but exhibits mean reversion whenever quotes are updated, either due to a news event or due to arbitrage trades.[13] In line with the arbitrageurs' participation constraint in Definition 1, price differences $v_t^i - v_t^j$ are exploited only if they exceed the total trading costs. The crucial assumption behind the information revelation through arbitrage trading is price

[13]See Chan (1993) for necessary conditions for the stationarity of z_t.

pressure due to order flow as soon as prices deviate from their equilibrium relationship. Therefore, arbitrage activity reinforces the law of one price (see, e.g., Fama (1965) and Ross (1976)). In the theoretical framework, the presence of trading costs implies that price pressure is present only if price differences are large enough, which allows to identify the following three different regimes:

Definition 3. *Arbitrage activity implies three possible regimes r_t of the economy at time t: either the arbitrageur prefers to stay idle ($r_t = 0$), she sells at market i and buys at market j ($r_t = pos$), or vice versa ($r_t = neg$). Formally, given price difference $z_t = v_t^i - v_t^j$ at time t, the economy is in the following regime:*

$$r_t = \begin{cases} pos, & if\ z_t > S_t^i + S_t^j + c_t^{pos} \\ neg, & if\ -z_t > S_t^i + S_t^j + c_t^{neg} \\ 0, & else \end{cases} \tag{16}$$

where $c_t^{pos} \geq 0$ and $c_t^{neg} \geq 0$ correspond to the trading costs.

The regime r_t is central for the identification of the trading costs. In a frictionless market the theoretical no-arbitrage condition requires price differences to be exploited as soon as they arise, price dynamics would collapse to a 1-regime error correction model and price differences would be stationary and price pressure would be present continuously. However, market fragmentation imposes impediments to arbitrage and price differences may become sizeable which implies a non-linear adjustment process towards the long-run equilibrium given by the law of one price. The magnitude of the *no-trade* region determines if price differences persist because they do not correspond to a profitable arbitrage strategy. The boundaries of the no-trade region are defined as the minimum price differences required to make the arbitrageur indifferent between trading and staying idle ($r_t = 0$). If price differences z_t exceed the threshold, however, one can expect price pressure stemming from arbitrage trade to reinforce the equilibrium relationship represented by the law of one price. In other words, the law of one price may temporarily fail to hold whenever $r_t = 0$, whereas quoted prices exhibit mean reversion if and only if the economy is in regime $r_t = pos$ or $r_t = neg$. Definition 3 allows for two generalizations relative to the theoretical benchmark of Section 2: First, trading costs may be asymmetric in the sense that it can be costlier for the arbitrageur to perform cross-market trades in one direction rather than in the reverse direction ($c_t^{\text{pos}} \neq c_t^{\text{neg}}$). Potential reasons for asymmetries are, e.g., short-sale constraints or asymmetric buy and sell fees. Second, trading costs and liquidity can be time-varying such that during some periods the *no-trade* region widens.

Price pressure from arbitrageurs has the following implications: First, if $z_t > 0$, it is profitable to buy from market j and sell at market i. Consequently one can expect prices to adjust at least until $z_t = S_t^i + S_t^j + c_t^{\text{pos}}$. In the theoretical framework, I keep notational complexity small and assume that prices at the market where new information occurs are inelastic. For the empirical framework I relax this assumption and allow price pressure by arbitrage trading to increase prices at the buy side market and decrease prices at the sell side market. Further, price adjustments are presumably non-instantaneous. More specifically, the implied return dynamics $\Delta v_t^k := v_t^k - v_{t-1}^k$ are as follows:

$$\begin{pmatrix} \Delta v_t^i \\ \Delta v_t^j \end{pmatrix} = \begin{pmatrix} \mu_{i,t}^r \\ \mu_{j,t}^r \end{pmatrix} + \begin{pmatrix} \alpha_i^r \\ \alpha_j^r \end{pmatrix} \underbrace{\begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} v_{t-1}^i \\ v_{t-1}^j \end{pmatrix}}_{z_{t-1}} + \begin{pmatrix} u_t^i \\ u_t^j \end{pmatrix} \tag{17}$$

where $\mu_{k,t}^r$ corresponds to a potential time-varying mean specification, for instance in the presence of autoregressive dynamics and u_t^k denotes the innovation process. The theoretical framework as of Assumption 1 implies $\mu_{k,t}^r = 0 \quad \forall r$. Equation (17) corresponds to a vector error correction model (VECM) in the spirit of Engle and Granger (1987). The equation implies that price adjustments can occur due to idiosyncratic shocks u_t or due to the activity of arbitrageurs in response to profitable arbitrage opportunities. α_i^r and α_j^r correspond to the price adjustment in response to price pressure from arbitrage trades.

If, for instance, $r_t = \text{pos}$, the arbitrageur buys at exchange j and sells at the expensive exchange i. Consequently, prices at the buy side market should decrease ($\Delta v_t^i < 0$) and prices at the sell side market should increase ($\Delta v_t^j < 0$). For the adjustment terms α^{pos} this implies $\alpha_j^{\text{pos}} \geq 0$ and $\alpha_i^{\text{pos}} \leq 0$. The reverse case holds for $r_t = \text{neg}$, implying that $\Delta v_t^i > 0$ and $\Delta v_t^j < 0$. This is the case if, again, $\alpha_j^{\text{neg}} \geq 0$ and $\alpha_i^{\text{neg}} \leq 0$. However, although theoretically the sign of the adjustment terms should be identical for the regimes *pos* and *neg*, magnitudes of price adjustment may differ across the regimes. Most importantly, during the regime $r_t = 0$, no price adjustment due to arbitrage trading should be present and thus price dynamics only follow idiosyncratic shocks such that $\Delta v_t^k = u_t$.

Instead of imposing the law of one price directly, one could also aim at estimating the cointegration relationship $(1, \quad \beta)\left(v_t^i, \quad v_t^j\right)'$ which would significantly increase the estimation uncertainty (see, e.g., Hansen and Seo (2002), Seo (2011), Ters and Urban (2019)). Latter approaches focus on uncertain or unstable cointegration relationships, for instance in applications related to statistical arbitrage. I, instead, impose the law of one price as an economically motivated relationship to mitigate identification issues (see, e.g., Martens et al. (1998) and Stevens (2015)).

Equation (17) nests the price dynamics implied by the theoretical framework in Section 2. First, quote adjustments do not occur on market k continuously but at irregular frequencies τ_l^k. Therefore, when allowing for clock shifts in intrinsic event-time the noise process $u_{t^k}^k$ exhibits volatility $\mathbb{V}\left(u_t^k\right) = \sqrt{2\sigma^2\lambda_k^{-1}}$. Second, setting $\alpha_i^{\text{pos}} = 0, \alpha_j^{\text{pos}} = 1$ and $\alpha_i^{\text{neg}} = -1, \alpha_j^{\text{pos}} = 0$ ensures corresponding price pressure if news occurred either on market i or on market j. The resulting dynamics fully recover the dynamics of the theoretical model. In this case, if, for instance, information arrived on market i at time

t_l^i and triggered an arbitrage trade, it holds that

$$\begin{pmatrix} v^i_{t^i_{l-1}} \\ v^j_{t^i_{l-1}} \end{pmatrix} + \begin{pmatrix} \Delta v^i_{t^i_l} \\ \Delta v^j_{t^i_l} \end{pmatrix} = \begin{pmatrix} v^i_{t^i_{l-1}} \\ v^j_{t^i_{l-1}} \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} z_{t-1} = \begin{pmatrix} v^i_{t^i_{l-1}} \\ v^i_{t^i_{l-1}} \end{pmatrix}. \tag{18}$$

3.2 Estimation

In this section I derive the statistical properties of the model given by Equation (16) and Equation (17) and provide a (Bayesian) framework to estimate the trading costs c_t^r. First, I parametrize the time-varying trading costs c_t^r as a linear function of fixed costs and exposure to a potentially time-varying proxy of a source of exogenous trading costs.

Definition 4. *I parametrize the threshold c_t^r from Equation* (16) *as a function of the observable proxy for trading costs x_t with $\mathbb{V}(x_t) > 0$ such that*

$$c_t^r := \max\left(0, c_0^r + c_1 x_t\right). \tag{19}$$

Here, c_0^r corresponds to (unobservable) fixed trading costs which may depend on the trading direction. x_t is an vector of observations that proxy a time-varying source of trading costs for an arbitrageur. This could be, for instance, a time dummy to reflect changes in regulation (e.g., event fixed effect), trading fees, shorting costs or similar instruments. The parameter c_1 captures the effect of x_t on the latent trading costs. The parametrization allows to derive the sensitivity of the threshold c_t^r with respect to changes in the chosen instrument for trading costs by taking the following relationship into consideration

$$\frac{\partial(c_t^r + S_t^i + S_t^j)}{\partial x_t} = c_1 + \frac{\partial S_t^j}{\partial x_t} + \frac{\partial S_t^i}{\partial x_t}. \tag{20}$$

Here, $\frac{\partial S_t^k}{\partial x_t}$ can be modelled directly as the spreads S_t^k and the proxy x_t are both observable, and c_1 is the effect on the no-trade region in excess of (endogenous) spreads. The estimation can be performed either via concentrated maximum likelihood (see Tong (1983), Tsay (1998) and Hansen and Seo (2002)) or by means of Bayesian inference (see, e.g., Forbes et al. (1999) and Huber and Zörner (2019)). I employ a standard Bayesian approach that effectively circumvents issues related to the optimization of complex likelihood functions. The results are quantitatively similar for concentrated maximum likelihood methods.

The threshold vector error correction model in Equation (17) can be rewritten as a multivariate linear regression

$$\Delta V^r = X^r B^r + U^r \text{ where } r \in \{\text{neg, pos, } 0\} \tag{21}$$

with

$$\Delta V^r_{t_r} = \begin{pmatrix} \Delta v^i_{t_r} \\ \Delta v^j_{t_r} \end{pmatrix}' \text{ and } X^r_{t_r} = \begin{pmatrix} 1 \\ z_{t_r - 1} \end{pmatrix}'. \tag{22}$$

Here, t_r corresponds to the stacked dates of all observations in regime r. The parameters to estimate are

$$\theta = \left\{ \underbrace{\begin{pmatrix} \mu_1^{\text{neg}} & \mu_2^{\text{neg}} \\ \alpha_1^{\text{neg}} & \alpha_2^{\text{neg}} \end{pmatrix}}_{B^{\text{neg}}}, \begin{pmatrix} \mu_1^{\text{pos}} & \mu_2^{\text{pos}} \\ \alpha_1^{\text{pos}} & \alpha_2^{\text{pos}} \end{pmatrix}, \begin{pmatrix} \mu_1^{0} & \mu_2^{0} \\ \alpha_1^{0} & \alpha_2^{0} \end{pmatrix}, \begin{pmatrix} \Sigma^{\text{neg}} \\ \Sigma^{\text{pos}} \\ \Sigma^{0} \end{pmatrix}, \begin{pmatrix} c_0^{\text{neg}} \\ c_0^{\text{pos}} \\ c_1 \end{pmatrix} \right\}. \tag{23}$$

Under the assumption that the error terms U^r are zero-mean multivariate normal distributed with variance covariance matrix Σ^r, the likelihood of the data conditional on the parameters θ is given by

$$L(\Delta V|\theta, X) \propto \prod_{r \in \{\text{neg,pos,0}\}} |\Sigma^r|^{-\frac{T^r}{2}} \exp\left(-\frac{1}{2} \text{tr}\left(\left(\Sigma^r\right)^{-1} U^{r\prime}_\theta U^r_\theta \right) \right) \tag{24}$$

where T^r is the number of observations in regime r and $U_\theta^r = \Delta V_{t_r}^r - X^r B^r$. I specify non-informative prior distributions for $\beta^i := \text{vec}\left(B^i\right), \Sigma^i$, c_0 and c_1 as follows:

$$p(\Sigma^r) \sim IW(V, v), p(\text{vec}(B^r)|\Sigma^r) \sim MN(0, C \otimes \Sigma^r), p(c) \sim U(-\infty, \infty) \tag{25}$$

where $IW(\cdot, v)$ corresponds to an inverse Wishart Distribution with v degrees of freedom and positive scale matrix V, $MN(\cdot)$ corresponds to a multivariate normal distribution and $U(\cdot)$ corresponds to a uniform distribution with unbounded support. In the empirical analysis, I use $v = 2$, $V = 10^{-5}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $C = V^{-1}$ as hyperparameters. I perform inference on θ using Monte Carlo Markov Chain methods. I provide a detailed description of the sampling algorithm in Appendix C.

4 Trading Costs in Cryptocurrency Markets

I perform the empirical investigation with data on the cryptocurrency market. More specifically, I use a dataset of minute-level orderbook snapshots for large cryptocurrency exchanges offering trading Bitcoin versus US-Dollar. The data is complemented by granular information on settlement latency which constitutes a relevant source of price risk for cross-market arbitrageurs in any decentralized market. The focus on cryptocurrency markets is relevant for at least two reasons: First, substantial obstacles to cross-market arbitrage seem to persist and hamper price informativeness in these rather novel markets. Examples for well-documented frictions include regulations, withdrawal restrictions and exchange risks (see, e.g. Choi et al. (2018) and Makarow and Schoar (2019)). Therefore, the quantification of the trading costs allows to address the economic relevance of the present frictions and to evaluate the corresponding persistence of price differences.

Second, one salient feature of any blockchain-based settlement procedure is the trust-less verification process which replaces fast but potentially inefficient intermediaries. Recently, however, it has been documented that the limited capacities of proof-of-work consensus protocols and the large intra-daily variation of transactions waiting for settlement results in costly and time-consuming competition for the service of verification (see, e.g. Biais et al. (2019) and Easley et al. (2019)). Such latency implies trading costs due to non hedgeable price risk for cross-market arbitrageurs (Hautsch et al., 2019). Further, I exploit an proxy for latency risks to investigate the adverse selection channel of trading costs on liquidity.

4.1 Data

I use data from the Bitcoin network, one of the most popular decentralized protocols since Nakamoto (2008) published the concept and the underlying code. Bitcoin can be traded continuously on more than 400 markets that differ substantially in terms of location, fee structure and investors access. I employ hand collected orderbook information from the public application interfaces (APIs) of the largest cryptocurrency exchanges that feature BTC versus USD trading.[14]

For the analysis at hand I use minute-level orderbook data from the exchanges *Bitstamp* and *Gemini* which both do not allow any form of margin trading. Both exchanges comply with the virtual currency license of the New York State Department of Financial Services (DFS). The sample ranges from March 1st, 2018 until August 1st, 2019. Daily trading volume at the two exchanges varies considerably, ranging from 1.9 to almost 180 Million USD at *Gemini* and exceeding 450 Million USD at *Bitstamp* during periods of high trading activity. The time-series of midquote price differences is computed each

[14]A more detailed analysis of the underlying dataset is provided in Hautsch et al. (2019).

Table 1: Summary statistics of the price differences.
This table provides summary statistics for the quoted prices for Bitcoin in USD on *Bitstamp* and *Gemini*. The sample is based on minute level information starting from March 1st, 2018 until August 1st, 2019. z_t corresponds to the midquote price differences (Bitstamp - Gemini as of Equation (26)). δ_t is the spread-adjusted midquote price differential as of Equation (28). Basis points (*bp*) are always computed by scaling with the average midquote across both exchanges. *% Fraction of Excess Price Differences* corresponds to minutes in which spread-adjusted price differences imply an arbitrage opportunity. Trading *volume* is computed in million USD per day.

variable	Mean	SD	Min	5%	Median	95%	Max
z_t (USD)	0.22	7.20	-296.36	-7.65	-0.20	9.53	499.31
Spread (Bitstamp)	3.28	3.14	0.01	0.01	2.31	9.47	53.65
Spread (Gemini)	1.68	2.47	0.01	0.01	0.71	6.40	87.03
δ_t (USD)	0.22	6.08	-293.51	-5.05	-0.00	6.80	497.06
$\|z_t\|$ (bp)	5.90	8.78	0.00	0.39	4.38	15.80	951.86
Spreads (bp)	4.06	3.25	0.01	0.23	3.35	10.07	118.30
$\|\delta\|$ (bp)	3.15	8.35	0.00	0.00	0.90	12.20	947.57
% Fraction of Excess Price Differences	58	49	-	-	-	-	-
Volume (Gemini, million USD)	21.87	19.31	1.91	3.42	16.73	56.43	176.03
Volume (Bitstamp, million USD)	66.02	56.30	5.37	14.09	48.72	173.00	477.67

minute as follows

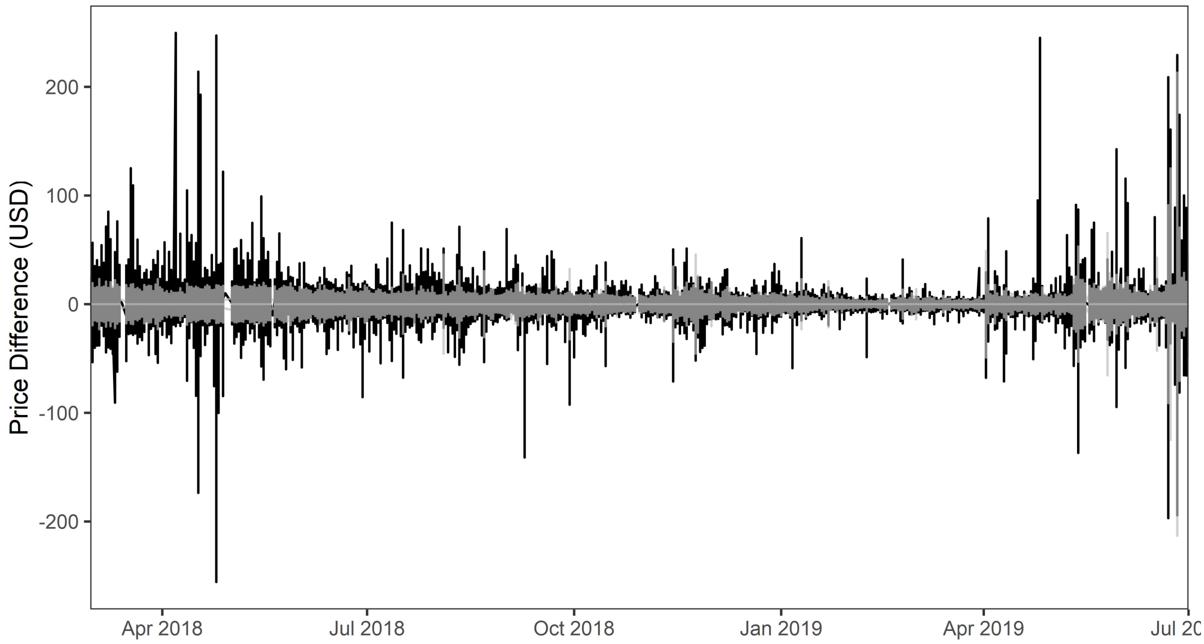
$$z_t := q_t^{\text{Bitstamp}} - q_t^{\text{Gemini}} = \frac{1}{2}\left(a_t^{\text{Bitstamp}} - a_t^{\text{Gemini}} + b_t^{\text{Bitstamp}} - b_t^{\text{Gemini}}\right) \tag{26}$$

where a_t^k and b_t^k are the quoted best ask and best bid, respectively. Spreads at the best level denote the difference between the ask and bid price on each market.

$$\begin{pmatrix} S_t^{\text{Bitstamp}} \\ S_t^{\text{Gemini}} \end{pmatrix} := \frac{1}{2}\begin{pmatrix} a_t^{\text{Bitstamp}} - b_t^{\text{Bitstamp}} \\ a_t^{\text{Gemini}} - b_t^{\text{Gemini}} \end{pmatrix}. \tag{27}$$

Table 1 provides summary statistics of the quoted prices and spreads. Midquote price differences are centred around zero but reveal substantial variation over time, sometimes reaching almost 500 USD (1000 bp). Quoted spreads at the two exchanges are relatively small but exhibit substantial variation over time. During some periods, spreads at *Gemini* spike at almost 90 USD. However, average spreads are around 2 bp which indicates that *Gemini* and *Bitstamp* resemble rather liquid markets, also in comparison to US equity

Figure 8: Quoted price differences and spreads. This figure visualizes the minute-level price differences between *Bitstamp* and *Gemini.* The black line corresponds to the midquote price differences (z_t) in USD whereas a positive value indicates that *Bitstamp* quotes a higher price than *Gemini.* The grey area corresponds to the spreads required to buy and sell a marginal unit of Bitcoin at the two markets.



markets (Brogaard et al., 2014). Price differences in excess of the spreads constitute potential arbitrage opportunities. The average spread-adjusted price difference is

$$\begin{aligned}\delta_t :=& \max\left\{0, |z_t| - \left(S_t^{\text{Bitstamp}} + S_t^{\text{Gemini}}\right)\right\} \\ =& \max\left\{0, b_t^{\text{Gemini}} - a_t^{\text{Bitstamp}}, b_t^{\text{Bitstamp}} - a_t^{\text{Gemini}}\right\}. \end{aligned} \tag{28}$$

Average midquote differences adjusted for spreads are around 3bp and positive during 58% of all minutes in the sample period.

Figure 8 visualizes price differences and quoted spreads at the two exchanges during the sample period. The black line corresponds to the midquote valuation differences z_t in USD and the grey shaded area corresponds to the minute-level sum of the quoted spreads $S_t^i + S_t^j$. Consequently, the area between the two lines corresponds to δ_t. The figure suggests that arbitrage opportunities may exist, however, deviations from the law of one price do not persist, instead, mean-reversion seems to play a role. To quantify the extent of arbitrage activity, I estimate the threshold vector error correction model as of above. I quantify the impact of the number of transactions waiting for settlement on the *no-trade* region of arbitrageurs.

To do so, I gather transaction-specific information from `blockchain.com`, a popular provider of Bitcoin network data and download all blocks verified during my sample period. I extract information about all verified transactions in this period. Each transaction contains a unique identifier, a timestamp of the initial announcement to the network, and, among other details, the fee (per byte) the initiator of the transaction offers validators to verify the transaction. Any transaction in the Bitcoin network, irrespective of its origin, has to go through the so-called *mempool* which is a collection of all unconfirmed transactions. These transactions wait in the mempool until they are picked up by validators and get verified. However, the Bitcoin protocol restricts the number of transactions that can enter a single block and therefore induces competition among the originators of transactions who can offer higher settlement fees to make it attractive for validators to include transactions in the next block.

Validators bundle transactions that wait for verification and try to solve a computationally expensive problem which involves numerous trials until the first validator finds the solution. For the Bitcoin protocol, validators successfully find a solution and append a block on average every 10 minutes. The number of transactions waiting for verification serves as a proxy for the usage of the Bitcoin network. The average number of transactions waiting for verification is above 8,400 and temporarily exceeds 39,000. As on average only around 1,000 enter a single block, the queue of transaction waiting for verification induces settlement latency. The probability of being included in the next block decreases with the number of transactions that wait for settlement. Table 2 shows

Table 2: Descriptive Statistics of the Bitcoin Network
This table reports descriptive statistics of our Bitcoin transaction data. The sample contains all transactions settled in the Bitcoin network from March 1st, 2018, until August 31, 2019. *Fee per Byte* is the total fee per transaction divided by the size of the transaction in bytes in Satoshi where 100,000,000 Satoshi are 1 Bitcoin. *Fee per Transaction* is the total settlement fee per transaction (in USD). I approximate the USD price by the average minute-level midquote across all exchanges in our sample. *Latency* is the time until the transaction is either validated or leaves the mempool without verification (in minutes). *Transaction Size* denotes the size of the transaction in bytes. *# Waiting transactions* is the number of transactions waiting for verification (per minute).

	Mean	SD	5 %	25 %	Median	75 %	95 %
Transaction Size	514.00	2169.86	192.00	225.00	248.00	372.00	962.00
Fee per Byte (Satoshi)	23.17	203.80	1.36	4.01	9.17	22.52	87.80
Fee per Transaction (USD)	0.60	8.08	0.02	0.07	0.15	0.40	2.00
Latency	30.54	165.31	0.73	3.58	8.85	20.28	90.87
# Waiting Transactions	8437.26	14438.03	324.00	1336.00	3429.50	8064.50	39415.00

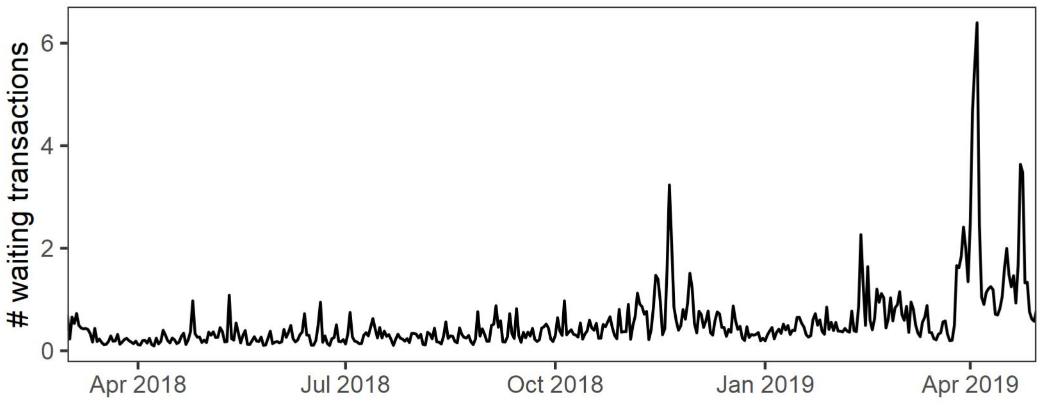
relevant summary statistics of the sample of Bitcoin transactions. The time until verification of a transaction in the Bitcoin network on average exceeds 30 minutes and exhibits substantial fluctuation. The costs of transferring Bitcoin from one wallet to another are on average 0.60 USD, irrespective of the trading size.

Panel A of Figure 9 illustrates the time-series of outstanding transactions during the sample period. Besides of regular intraday fluctuation patterns, periods of high network utilization occurred during December 2018 and since April 2019. The intra-daily variation of the Bitcoin network utilization is large. Panel B of Figure 9 illustrates the average number of transactions waiting for verification during the day, divided into intervals of 15 minutes in Central European Time. Network activity starts to spike at around 2pm CET which corresponds to 9am EST. As postulated, for instance, by Easley et al. (2019) and Biais et al. (2019), the number of transactions waiting for verification increases the latency of all transaction waiting for verification. I illustrate this relationship in Figure 10. On days with many transactions waiting for verification, the average latency in minutes until settlement increases, reflecting the competition for settlement services in the decentralized network.[15] The relationship between the number of transactions waiting

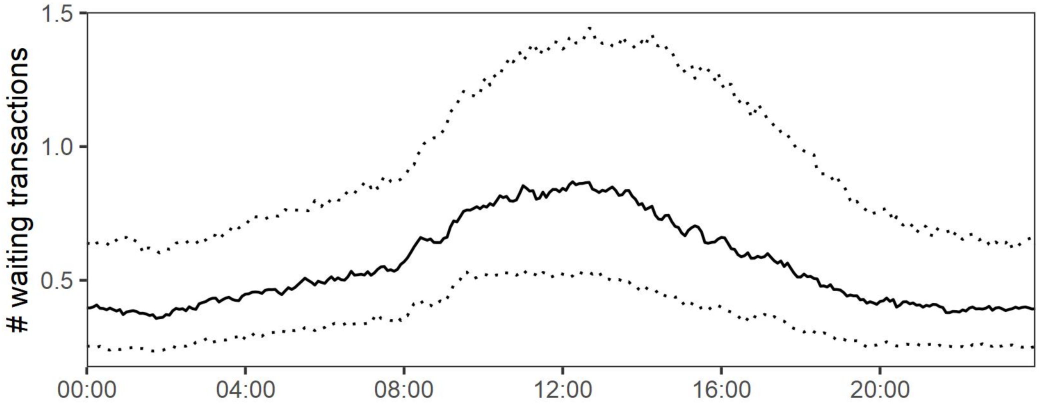
[15]I provide a more detailed analysis of the statistical relationship between the number of transaction waiting for verification and the settlement latency in the Appendix. For further information, see also

Figure 9: Blockchain congestion. These figures illustrate the distribution of the number of transactions waiting for verification across time and within the day.

Panel A: Number of transactions waiting for verification. This figure shows the time series of daily average number of transactions waiting for verification (in 10,000 transactions waiting for verification).



Panel B: Intra-daily fluctuations. This figure shows the average number of transactions waiting for verification during 15 minute intervals over the day.

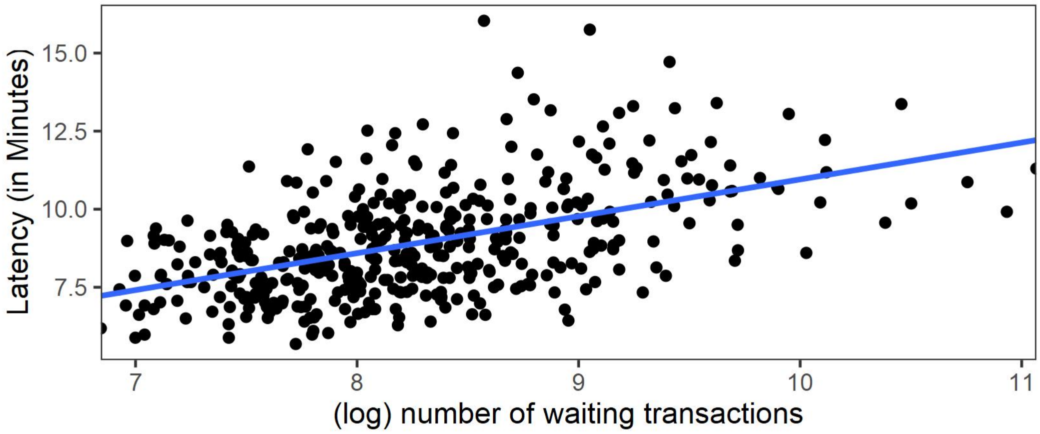


for verification and settlement latency implies non-hedgeable price risks. Therefore, if price risk due to settlement latency plays a role for the activity of arbitrageurs one would expect the no-trade region of arbitrageurs to widen during times of high network utilization. For the analysis, two observations are important: First, the exchanges *Gemini* and *Bitstamp* both net trades internally, effectively circumventing settlement latency for quote updating activities. Therefore, the number of transactions waiting for verification

---

Hautsch et al. (2019).

Figure 10: Transactions waiting for verification and settlement latency. This figure visualizes the relationship between the number of transactions waiting for verification and the latency of subsequently verified transactions. The scatterplot shows with (log) number of transactions waiting for verification on the x-axis and the average daily waiting time (in minutes) of all transactions verified on that particular day. The blue line indicate the OLS estimator of the slope and constant of regressing waiting time on log of the number of waiting transactions.



is relevant only for cross-market arbitrageurs which are forced to use the blockchain to funds between wallets. Second, blockchain congestion due to arbitrageurs transactions waiting for verification is unlikely to be of any concern. The Bitcoin network is utilized for transactions of any purpose, including consumption or financing of illegal activities (see, e.g., Foley et al. (2019)). Cross-market trades with the purpose of exploiting price differences are of negligible relevance and therefore no feedback effect from the presence of price differences on settlement latency dilutes the results. I examine a novel dataset that allows to identify potential arbitrage transactions and I find that such transactions only represent a minor fraction of all transactions waiting for verification and further do not differ with respect to the relevant summary statistics from transactions unrelated to cross-market arbitrage activity. To quantify the magnitude of impediments to arbitrage due to settlement latency I use the (log of) the number of transactions waiting for verification as proxy x_t in the econometric model of Equation (17).

4.2 Estimation Results

Next, I turn to the estimated trading costs of arbitrageurs. Table 3 summarises the posterior distribution of the parameters. Here, the estimation is based on setting $\mu^r_{i,t}$ to a constant such that the econometric model resembles Equation (21). Inference is drawn based on Gibbs sampling procedures with Metropolis Hasting steps as illustrated in the Appendix. The sample of parameters is generated with 50 parallel chains of length 20.000 after discarding 20.000 burn-in iterations.

As of Panel A of Table 3, the estimated parameters $\hat{c}^r_0$ are between 3.7 USD and 5.05 USD. The exposure of the trading costs with respect to the number of unconfirmed transactions, $\hat{c}_1$ is positive. Thus, higher blockchain activity imposes additional costs for cross-market trading.

The average estimated total trading costs, $\hat{c}^r := \max\left(0, \hat{c}^r_0 + \hat{c}^r_1 \bar{x}_t\right)$ where $\bar{x}_t$ is the time-series average of the (log) number of transactions outstanding amount to 15.11 USD in the case where *Bitstamp* is the sell-side market and 13.81 USD in the reverse direction, hinting at substantial trading costs in excess of the spreads. Therefore, pushing forces back to the equilibrium pricing relationship are present only, if price differences exceed these thresholds *after* already having adjusted for the spreads. Further, sampling from the posterior distribution allows to evaluate the distribution of the difference $c^{\text{pos}}_0 - c^{\text{neg}}_0$ which is positive, thus trading costs are different depending on the direction of the trade. In relation to the average absolute price difference, $|\delta_t|$, the magnitudes of $\hat{c}^r$ are substantial, covering almost 63% of all observed price differences in the sample. Relative to the liquidity, latent trading costs resemble more than 75% of the no-trade regions implied by the data. Moreover, an increase in the number of outstanding transactions by 1% implies an average increase in the no-trade region by more than 2bp. Therefore, fluctuations in the network activity, as for instance characterized by intra-daily variation as of Figure 9 imply considerable impediments for arbitrageurs and may impose severe

Table 3: Posterior parameter estimates.
This table summarises the posterior distribution of the main parameters of interest of the model as of Equation (21) and Equation (16). Panel A contains posterior means of the threshold parameters c_0^r and c_1^r. The values in brackets correspond to the 99% credible regions. Panel B contains posterior means and credible regions for the adjustment parameters α_k^r and the mean values μ from the different regimes r. Inference is drawn based on Gibbs sampling procedures with Metropolis Hasting steps as illustrated in the Appendix. The sample of parameters is generated with 50 parallel chains of length 20.000 after discarding 20.000 burn-in iterations.

Panel A: Posterior estimates of c_t^r. *Diff* corresponds to the posterior distribution of the difference of the two fixed cost parameters, $c_0^{\text{pos}} - c_0^{\text{neg}}$.

r	Neg	Pos	Diff	
c_0	3.761	5.055	1.294	
	(3.45, 3.79)	(4.99, 5.62)	(1.21, 2.17)	
c_1				1.271
				(1.23, 1.67)

Panel B: Posterior estimates of α_r and μ^r.

r	Pos	0	Neg
α_{bitstamp}	-0.789	-0.049	-0.655
	(-0.8, -0.77)	(-0.15, 0.14)	(-0.67, -0.62)
α_{gemini}	0.084	0.004	0.091
	(0.027, 0.091)	(-0.08, 0.09)	(0.081, 0.095)
μ_{bitstamp}	-0.77	0.011	0.36
	(- 0.78,-0.71)	(-0.01, 0.03)	(0.29, 0.38)
μ_{gemini}	0.847	-0.001	-0.848
	(0.74, 0.95)	(-0.02, 0.01)	(-0.94, -0.66)

effects on price informativeness. Panel B of Table 3 contains summary statistics of the posterior distribution of the remaining parameters of interest, $\hat{\alpha}_k^r$ and $\hat{\mu}_k^r$. The adjustment parameters α_k^r exhibit the expected sign and reflect pushing forces back to equilibrium whenever price differences indicate a regime different from $r_t = 0$. Positive price difference ($r_t = \text{pos}$) imply that *Bitstamp* serves as a sell-side exchange and thus prices are expected to decrease at *Bitstamp* and to increase at the buy-side exchange *Gemini*. The credible regions of $\hat{\alpha}_k^0$ for a regime without arbitrage trading both contain zero. The spot drift terms μ_k^r reflect behaviour in line with the adjustment parameters: In periods where trading is profitable for arbitrageurs, price differences tend to decrease at both market. Only exception is the negative sign for $\mu_{\text{gemini}}^{\text{neg}}$. The price dynamics further indicate that

Gemini adjusts to price differences at a lower rate than its competitor *Bitstamp*. This could be due to at least two reasons: Either the price impact of arbitrageurs at *Gemini* is mostly absorbed or the rate of information arrival at *Gemini* leads the market with *Bitstamp* subsequently adjusting its current quotes.

Next, I investigate the relationship between trading costs for arbitrageurs and the liquidity component in the spreads. Based on the framework developed in Section 2, adverse selection should play a role in fragmented markets. Then, shifts in trading costs should be reflected in the dynamics of the quoted spreads. The empirical analysis rests on the observed spread at the two markets in the sample. That is, I estimate:

$$\begin{pmatrix} S_t^{\text{Bitstamp}} \\ S_t^{\text{Gemini}} \end{pmatrix} = \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} + \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} x_t + \begin{pmatrix} \rho_1^1 & \rho_2^1 \\ \rho_1^2 & \rho_2^2 \end{pmatrix} \begin{pmatrix} S_{t-1}^{\text{Bitstamp}} \\ S_{t-1}^{\text{Gemini}} \end{pmatrix} + \Gamma\beta + \begin{pmatrix} u_t^1 \\ u_t^2 \end{pmatrix} \tag{29}$$

where x_t is the (log) number of transactions waiting for verification at minute t as a proxy for the trading costs and u_t^k are potentially correlated normally distributed error terms. Γ contains a host of control variables to explain quoted spreads. As controls I use trading volume V_t^k (see, e.g. Lin et al. (1995) and Stoll (1989)), the lagged cross-market average midquote and the cross-market average minute level spot volatility σ_t (see, e.g. Easley and O'Hara (1987)). I estimate the spot volatility using the procedure proposed by Kristensen (2010).[16] Table 4 illustrates the parameter estimates of the vector autoregressive structure as of Equation (29). The parameters ω_1 and ω_2 are both negative, indicating that an increase in the number of transactions waiting for verification decreases quoted spreads in line with the adverse selection proposed in Section 2. Further, the table reveals that spreads increase with spot volatility as a measure of uncertainty, in line with Lemma 2.

The empirical analysis reveals two different effects of a shift in the trading costs for

[16] I refer to Hautsch et al. (2019) for further information regarding estimating σ_t.

Table 4: Spread decomposition. Parameter estimates of Equation (29). x_t denotes the (log) number of transactions waiting for verification, σ_t denotes the minute level spot volatility, $\bar{q}_t$ is the cross-market average midquote, V_t^k corresponds to trading volume (daily) at exchange k. Values in brackets denote p values of the estimated parameters.

	Dependent variable:	
	S_t^{Bitstamp}	S_t^{Gemini}
$S_{t-1}^{\text{Bitstamp}}$	0.455***	0.028***
	(0.001)	(0.001)
S_{t-1}^{Gemini}	0.051***	0.552***
	(0.001)	(0.001)
γ	−13.593***	−2.683***
	(0.115)	(0.095)
x_t	−0.114***	−0.051***
	(0.003)	(0.002)
σ_t	5.818***	6.420***
	(0.136)	(0.112)
$\bar{q}_{t-1}$	1.824***	0.419***
	(0.013)	(0.011)
V_t^{Gemini}	0.0003	0.004***
	(0.0004)	(0.0003)
V_t^{Bitstamp}	0.005***	−0.003***
	(0.0002)	(0.0001)
Observations	588,716	

arbitrageurs on the arbitrageurs participation constraint as of Equation (4): First, Panel A of Table 3 illustrates a direct cost effect. More transactions waiting for verification increase the price risks cross-market arbitrageurs in the Bitcoin market are exposed to and consequently widen the no-trade region which constitutes the positive coefficient c_1. Second, if cross-market arbitrageurs constitute an adverse selection risk for local liquidity providers, the theoretical model implies than an increase in x_t narrows spreads. This expectation is confirmed as of Table 4. After controlling for other potential factors determining liquidity provisioning, the number of transactions waiting for verification decreases quoted spreads.

5 Conclusions

Fragmented trading characterizes nowadays financial market infrastructure - similar assets are traded on multiple venues that differ with respect to transparency, order types and, more generally, access for investors. Informational efficiency, however, requires that cross-market arbitrageurs monitor and exploit price differences such that quoted prices at all markets ultimately incorporate information regarding the fundamental value of the underlying asset. This pivotal role of arbitrage activity evaporates when market frictions impose limits to arbitrage that render cross-market trading costly.

Bitcoin is one particular asset that exhibits substantial market fragmentation and simultaneously imposes considerable technology-related arbitrage cost. In this paper, I analyze the implications of blockchain-related settlement latency (see Hautsch et al. (2019)) on informational efficiency. I find that faster settlement latency reduces technology-related arbitrage costs but at the same time increases quoted spreads. Thus, the effect of reducing the technological frictions on arbitrage costs is partially offset by an increase in costs related to liquidity-related frictions. In extreme cases, this substitution effect can even predominate and therefore *harm* informational efficiency.

The main econometric challenge hereby is to estimate arbitrage costs, which are generally not observable. When arbitrageurs enter the market and start to exploit price differences, however, associated price pressure towards the law of one price should reveal their activity – arbitrage trading therefore implies a cointegration relationship between markets. Whenever arbitrage costs do not render arbitrage trades profitable (e.g., price differences fall within a no-arbitrage regime) the correction mechanism should evaporate.

I exploit high-frequency orderbook data of two of the largest cryptocurrency exchanges to estimate the no-arbitrage regimes. More specifically, I provide a (Bayesian) estimation procedure to parametrize the thresholds as functions of latent exchange-specific arbitrage costs and time-varying observable proxies for arbitrage costs. The parametrization allows decomposing arbitrage costs into liquidity and technology related components. The number of transactions waiting for verification serves as a measure of network activity that increases the price risks of arbitrageurs. The quoted spreads on both markets denote liquidity.

I show that faster settlement reduces arbitrage costs. More specifically, an increase of one percent in the number of transactions waiting for verification implies around 2 basis points wider no-trade regions for arbitrageurs. Most importantly, however: This finding does not necessarily imply that developing faster consensus protocols ultimately benefits informational efficiency. Instead, I also document that faster settlement is associated with larger spreads. In fact, my findings suggest that a 10 basis point decrease in technology-related arbitrage costs is associated with a 3 basis point *increase* in spreads. The offsetting effects suggest that efforts to reduce the latency of blockchain-based settlement might have unintended consequences for liquidity provision which – in extreme cases – could even *harm* informational efficiency.

I back the empirical evidence by theoretical reasoning and embed my results in a model in which liquidity providers anticipate that arbitrageurs exploit their stale quotes more frequently if settlement time is low and thus set wider spreads to cope with the adverse selection risk. As a result, the direct effect of faster settlement on informational efficiency is offset by larger liquidity-related limits to arbitrage, close in spirit to Foucault et al. (2017).

Ultimately, the empirical analysis of cryptocurrency data addresses the fundamental question how settlement procedures that rely on distributed ledger technologies affect market efficiency. In the recent past, for instance, the number of cryptocurrency exchanges across the globe has grown significantly. At the same time, technological frictions due to the time-consuming settlement latency impose limits to arbitrage that cause deviations from the law of one price to persist. The limited capacities of distributed consensus protocols served as motivation to decrease blockchain related frictions (e.g. "Segregated Witness", an implemented soft fork change in the transaction format for Bitcoin was also intended to mitigate the transaction speed problem). However, the findings of this paper suggest that informational efficiency *and* narrow spreads are hard to achieve if market makers fear adverse selection.

Reversely, the consolidation of the highly fragmented US equity market landscape (Regulation National Market System Rule 611) seems to be a way to foster price informativeness by reducing technology-related frictions for cross-market trading. However, the growing debate on intentional access delays, better known as speed bumps, as a means to hamper cross-market liquidity taking activities sheds light from a different perspective on a very similar trade-off between informational efficiency and adverse selections risks. In this particular example, however, liquidity providers criticize the abundance of technology-related costs and claim that adverse selection risks impose disproportionally high costs on all market participants.

It should be emphasized, however, that the theoretical discussion abstracts from endogenous effects of the market structure on the arrival rate of information, or, in other words, the incentives for information acquisition in general. Thus, before turning to welfare analysis in order to determine *optimal* values of technology-related frictions, a full equilibrium model would be required to model the order flow of the informed investors to subsequently investigate how information revelation is affected when trading costs are shifted (see, e.g. Pagnotta and Philippon (2018)). Such an analysis would certainly be interesting, but is left for future research.

References

Abadi, J. and M. Brunnermeier (2018). Blockchain Economics. Working paper.

Aldrich, E. M. and D. Friedman (2017). Order Protection through Delayed Messaging. Working paper.

Aoyagi, J. (2018). Strategic Speed Choice by High-Frequency Traders Under Speed Bumps. Working paper.

Baldauf, M. and J. Mollner (2019). Trading in Fragmented Markets. *The Journal of Financial and Quantitative Analysis (forthcoming).*

Balke, N. S. and T. B. Fomby (1997). Threshold Cointegration. *International Economic Review 38*(3), 627–645.

Benos, E., W. Huang, A. J. Menkveld, and M. Vasios (2019). The cost of clearing fragmentation. Working paper.

Biais, B., C. Bisiere, M. Bouvard, and C. Casamatta (2019). The Blockchain Folk Theorem. *The Review of Financial Studies 32*(5), 1662–1715.

Bondarenko, O. (2003). Statistical Arbitrage and Securities Prices. *The Review of Financial Studies 16*(3), 875–919.

Brauneis, A., R. Mestel, R. Riordan, and E. Theissen (2019). A High-Frequency Analysis of Bitcoin Markets. Working paper.

Brogaard, J., T. Hendershott, and R. Riordan (2014). High-Frequency Trading and Price Discovery. *The Review of Financial Studies 27*(8), 2267–2306.

Brolley, M. and D. Cimon (2019). Order Flow Segmentation, Liquidity and Price Discovery: The Role of Latency Delays. *Journal of Financial Quantitative Analysis (forthcoming).*

Brunnermeier, M. and S. Nagel (2004). Hedge Funds and the Technology Bubble. *The Journal of Finance 59*(5), 2013–2040.

Budish, E., P. Cramton, and J. Shim (2015). The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response. *Quarterly Journal of Economics 130*(4), 1547–1621.

Chan, K. S. (1993). Consistency and Limiting Distribution of the Least Squares Estimator of a Threshold Autoregressive Model. *The Annals of Statistics 21*(1), 520–533.

Chen, Y., Z. Da, and D. Huang (2019). Arbitrage Trading: The Long and the Short of It. *The Review of Financial Studies 32*(4), 1608–1646.

Chiu, J. and T. V. Koeppl (2019). Blockchain-based Settlement for Asset Trading. *The Review of Financial Studies 32*(5), 1716–1753.

Choi, K. J., A. Lehar, and R. Stauffer (2018). Bitcoin Microstructure and the Kimchi Premium. Working paper.

Duffie, D. and G. Manso (2007). Information Percolation in Large Markets. *The American Economic Review 97*(2), 203–209.

Dwyer, G. P., P. Locke, and W. Yu (1996). Index Arbitrage and Nonlinear dynamics Between the S&P 500 Futures and Cash. *The Review of Financial Studies 9*(1), 301–332.

Easley, D. and M. O'Hara (1987). Price, Trade Size, and Information in Securities Markets. *Journal of Financial Economics 19*(1), 69–90.

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

Engle, R. F. and C. Granger (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica 55*(2), 251.

Fama, E. F. (1965). The Behavior of Stock-Market Prices. *The Journal of Business 38*(1), 34–105.

Foley, S., J. Karlsen, and T. J. Putninš (2019). Sex, Drugs, and Bitcoin: How Much Illegal Activity is Financed Through Cryptocurrencies? *The Review of Financial Studies 32*(5), 1789–1853.

Forbes, C. S., G. R. J. Kalb, and P. Kofhian (1999). Bayesian Arbitrage Threshold Analysis. *Journal of Business & Economic Statistics 17*(3), 364–372.

Foucault, T., R. Kozhan, and W. W. Tham (2017). Toxic Arbitrage. *The Review of Financial Studies 30*(4), 1053–1094.

Foucault, T., A. Röell, and P. Sandås (2003). Market Making with Costly Monitoring: An Analysis of the SOES Controversy. *The Review of Financial Studies 16*(2), 345–384.

Gagnon, L. and A. G. Karolyi (2010). Multi-market Trading and Arbitrage. *Journal of Financial Economics 97*(1), 53–80.

Green, R. C. and S. Srivastava (1985). Risk Aversion and Arbitrage. *The Journal of Finance 40*(1), 257.

Gromb, D. and D. Vayanos (2010). Limits of Arbitrage. *Annual Review of Financial Economics 2*(1), 251–275.

Grossman, S. J. and J. E. Stiglitz (1980). On the Impossibility of Informationally Efficient Markets. *The American Economic Review 70*(3), 393–408.

Hagströmer, B. and A. J. Menkveld (2019). Information Revelation in Decentralized Markets. *The Journal of Finance (forthcoming).*

Hansen, B. E. and B. Seo (2002). Testing for two-regime Threshold Cointegration in Vector Error-Correction Models. *Journal of Econometrics 110*(2), 293–318.

Hanson, S. G. and A. Sunderam (2014). The Growth and Limits of Arbitrage: Evidence from Short Interest. *The Review of Financial Studies 27*(4), 1238–1286.

Hasbrouck, J. (1995). One Security, Many Markets: Determining the Contributions to Price Discovery. *The Journal of Finance 50*(4), 1175–1199.

Hautsch, N., C. Scheuch, and S. Voigt (2019). Trust takes Time: Limits to Arbitrage in Decentralized Markets. Working paper.

Hu, E. (2018). Intentional Access Delays, Market Quality, and Price Discovery: Evidence from IEX Becoming an Exchange. Working paper.

Huber, F. and T. O. Zörner (2019). Threshold Cointegration in International Exchange Rates: A Bayesian approach. *International Journal of Forecasting 35*(2), 458–473.

Kristensen, D. (2010). Nonparametric Filtering of the Realized Spot Volatility: A Kernel-Based Approach. *Econometric Theory 26*(1), 60–93.

Kyle, A. S. (1985). Continuous Auctions and Insider Trading. *Econometrica 53*(6), 1315.

Kyle, A. S. and W. Xiong (2001). Contagion as a Wealth Effect. *The Journal of Finance 56*(4), 1401–1440.

Lin, J.-C., G. C. Sanger, and G. G. Booth (1995). Trade Size and Components of the Bid-Ask Spread. *The Review of Financial Studies 8*(4), 1153–1183.

Lo, M. C. and E. Zivot (2001). Threshold Cointegration and nonlinear adjustment to the law of one price. *Macroeconomic Dynamics 5*(4), 533–576.

Makarow, I. and A. Schoar (2019). Trading and Arbitrage in Cryptocurrency Markets. *Journal of Financial Economics (forthcoming).*

Martens, M., P. Kofman, and T. C. F. Vorst (1998). A Threshold Error-Correction Model for Intraday Futures and Index Returns. *Journal of Applied Econometrics 13*(3), 245–263.

Mitchell, M., L. H. Pedersen, and T. Pulvino (2007). Slow Moving Capital. *The American Economic Review 97*(2), 215–220.

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

Ofek, E., M. Richardson, and R. F. Whitelaw (2004). Limited Arbitrage and Short Sales Restrictions: Evidence from the Options Markets. *Journal of Financial Economics 74*(2), 305–342.

O'Hara, M. and M. Ye (2011). Is Market Fragmentation harming Market Quality? *Journal of Financial Economics 100*(3), 459–474.

Pagnotta, E. S. and T. Philippon (2018). Competing on Speed. *Econometrica 86*(3), 1067–1115.

Park, H., J. W. Mjelde, and D. A. Bessler (2007). Time-varying Threshold Cointegration and the Law of one Price. *Applied Economics 39*(9), 1091–1105.

Pontiff, J. (1996). Costly Arbitrage: Evidence from Closed-End Funds. *Quarterly Journal of Economics 111*(4), 1135–1152.

Roll, R., E. Schwartz, and A. Subrahmanyam (2007). Liquidity and the Law of one Price: the Case of the Futures–Cash Basis. *The Journal of Finance 62*(5), 2201–2234.

Rösch, D. (2013). The Impact of Arbitrage on Market Liquidity. Working paper.

Ross, S. A. (1976). The Arbitrage Theory of Capital Asset Pricing. *Journal of Economic Theory 13*(3), 341–360.

Sagade, S., S. Scharnowski, E. Theissen, and C. Westheide (2019). A Tale of Two Cities-Inter-Market Latency, Market Integration, and Market Quality. Working paper.

Seo, B. (2011). Nonparametric Testing for Linearity in Cointegrated Error-Correction Models. *Studies in Nonlinear Dynamics & Econometrics 15*(2).

Shleifer, A. and R. W. Vishny (1997). The Limits of Arbitrage. *The Journal of Finance 52*(1), 35–55.

Stevens, J. (2015). Do Transaction Costs prevent Arbitrage in the Market for Crude Oil? Evidence from a Threshold Autoregression. *Applied Economics Letters 22*(3), 169–172.

Stoll, H. R. (1989). Inferring the Components of the Bid-Ask Spread: Theory and Empirical Tests. *The Journal of Finance 44*(1), 115.

Ters, K. and J. Urban (2019). Estimating Unknown Arbitrage Costs: Evidence from a 3-regime Threshold Vector Error Correction Model. *Journal of Financial Markets (forthcoming).*

Theissen, E. (2012). Price Discovery in Spot and Futures Markets: a Reconsideration. *The European Journal of Finance 18*(10), 969–987.

Tong, H. (1983). Threshold Models. In H. Tong (Ed.), *Threshold Models in Non-linear Time Series Analysis*, Lecture Notes in Statistics, pp. 59–121. New York, NY: Springer New York.

Tsay, R. S. (1998). Testing and Modeling Multivariate Threshold Models. *Journal of the American Statistical Association 93*(443), 1188–1202.

Verrecchia, R. E. (1982). Information Acquisition in a Noisy Rational Expectations Economy. *Econometrica 50*(6), 1415–1430.

Woodward, M. (2018). Bumping Up the Competition: The Influence of IEX's Speed Bump in US Financial Markets. Working paper.

Appendix

A Proofs

Proof of Lemma 1. First, I derive the distribution of $|v_{t+\tau} - v_t|$ given information available at time t. Assumption 1 implies that given v_t and τ, the distribution of $\delta_{t,\tau} = v_{t+\tau} - v_t$ is Gaussian with mean 0 and variance $\sqrt{\tau}\sigma$. Then, $|\delta_{t,\tau}|$ corresponds to a half-normal distribution with probability density function

$$\pi(|\delta_{t,\tau}|) = \frac{\sqrt{2}}{\sigma\sqrt{\tau\pi}} \exp\left(-\frac{|\delta_{t,\tau}|^2}{2\tau\sigma^2}\right). \tag{30}$$

The half-normal distribution has mean $\sqrt{\frac{2\tau}{\pi}}\sigma$. The probability of $|\delta_{t,\tau}| > z$ for $z > 0$ is

$$\pi\left(|\delta_{t,\tau}| > z\right) = 1 - \operatorname{erf}\left(\frac{z}{\sqrt{2\tau}\sigma}\right) \tag{31}$$

where $\operatorname{erf}(\cdot)$ corresponds to the Gauss error function defined as

$$\operatorname{erf}\left(\frac{z}{\sqrt{2\tau}\sigma}\right) := \frac{2}{\sqrt{\pi}} \int_0^{\frac{z}{\sqrt{2\tau}\sigma}} e^{-x^2} dx. \tag{32}$$

The derivatives follow immediately from

$$\frac{\partial \operatorname{erf}(x)}{\partial x} = \frac{2}{\sqrt{\pi}} \exp^{-x^2}. \tag{33}$$

$\square$

Proof of Lemma 2. First, I derive the expected losses of the market maker k *conditional on an information event* at the opposite market. After an information event occurs, the arbitrageur trades if the difference in valuation, $|\delta_{t,\tau}|$ exceeds the limits to arbitrage given by the threshold $c + S^k_{t+\tau}$. Therefore, it holds that

$$E\left(|\delta_{t,\tau}| \mid c + S_{t+\tau} < |\delta_{t,\tau}|\right) = \frac{\int_{c+S^k_{t+\tau}}^{\infty} |\delta_{t,\tau}|\pi(|\delta_{t,\tau}|)d|\delta_{t,\tau}|}{1 - P(|\delta_{t,\tau}| < c + S^k_{t+\tau})}. \tag{34}$$

From Lemma 1, $|\delta_{t,\tau}|$ follows a half-normal distribution. Thus, the denominator of Equation (34) can easily be derived as $\tilde{\pi}_t$, whereas for the nominator I make use of the fact that $\pi(|\delta_{t,\tau}|) = 2\phi\left(\frac{\delta_{t,\tau}}{\sigma\sqrt{\tau}}\right)$ for $\delta_{t,\tau} \geq 0$, where $\phi(\cdot)$ is the probability density function of the standard normal distribution. Then, I get

$$\int_{c+S^k_{t+\tau}}^{\infty} |\delta_{t,\tau}|\pi(|\delta_{t,\tau}|)d|\delta_{t,\tau}| = \sqrt{\frac{2}{\tau\pi\sigma^2}} \int_{c+S^k_{t+\tau}}^{\infty} \delta \exp\left(-\frac{\delta^2}{2\sigma^2\tau}\right) d\delta = \sqrt{\frac{2\tau\sigma^2}{\pi}} \exp\left(-\frac{\left(S^k_{t+\tau} + c\right)^2}{2\sigma^2\tau}\right). \tag{35}$$

Therefore, the expected loss from trading against arbitrageurs is

$$\tilde{\pi}_\tau \left(S^k_{t+\tau} - \frac{1}{\tilde{\pi}_\tau}\sqrt{\frac{2\tau\sigma^2}{\pi}} \exp\left(-\frac{\left(S^k_{t+\tau} + c\right)^2}{2\sigma^2\tau}\right)\right). \tag{36}$$

Note, that for $\frac{\partial \mathbb{E}(\Pi_{i,t+\tau}(S))}{\partial c}$ the following holds (based on Lemma 1):

$$\frac{\partial \tilde{\pi}_{i,\tau}(S)}{\partial c} = -\sqrt{\frac{2}{\pi\tau\sigma^2}} \exp\left(-\frac{\left(S^k_{t+\tau} + c\right)^2}{2\sigma^2\tau}\right). \tag{37}$$

Then, it holds that

$$\frac{\partial \mathbb{E}\left(\Pi_{i,t+\tau}(S)\right)}{\partial c} = \sqrt{\frac{2}{\pi\tau\sigma^2}} \exp\left(-\frac{\left(S^k_{t+\tau} + c\right)^2}{2\sigma^2\tau}\right)\left(S_{t+\tau} + c - S_{t+\tau}\right) > 0 \tag{38}$$

□

Proof of Proposition 1. First, I derive the distribution of the pricing error $|z^k_t| := |v^k_t - v_t|$. Assumption 1 implies that given $v^k_t = v_{t-\tau}$ for known τ, the distribution of $z^k_t = v_t - v_{t+\tau}$ is Gaussian with mean 0 and variance $\sigma^2\tau$. However, τ is a random variable which follows an exponential distribution with parameter λ_k. Therefore, the characteristic function of the stopped Wiener process z_t^k is

$$\varphi_{z_t^k} : \mathbb{R} \to \mathbb{C} \tag{39}$$

$$\varphi_{z_t^k}(s) = E\left(e^{isz_t^k}\right) = E\left(E\left(e^{isz_t^k} \mid \tau\right)\right) \tag{40}$$

$$= E\left(e^{-\tau\sigma^2 s^2/2}\right) \tag{41}$$

$$= \lambda_k \int_0^\infty e^{-(\lambda_k+\sigma^2 s^2/2)t}\, dt \tag{42}$$

$$= \frac{1}{1+\frac{\sigma^2 s^2}{2\lambda_k}}. \tag{43}$$

Equation (43) corresponds to the characteristic function of a Laplace distribution with expected value $\mathbb{E}\left(z1k_t\right) = 0$, scale parameter $\sqrt{\frac{\sigma^2}{2\lambda_k}}$ and corresponding probability density function

$$\pi\left(z_t^k\right) = \sqrt{\frac{\lambda_k}{2\sigma^2}} \exp\left(-\sqrt{\frac{2\lambda_k}{\sigma^2}}\left|z_t^k\right|\right). \tag{44}$$

By equation (44), the distribution of $|z_t^k| \propto \exp\left(-\sqrt{\frac{2\lambda_k}{\sigma^2}}\left|z_t^k\right|\right)$ corresponds to the kernel of an exponential distribution with rate parameter $\sqrt{\frac{2\lambda_k}{\sigma^2}}$. Therefore, the expected (L_1) difference between quoted price v_t^k and efficient price v_t is

$$\mathbb{E}\left(|z_t|\right) = \sqrt{\frac{\sigma^2}{2}\mathbb{E}\left(\tau\right)}. \tag{45}$$

Then, for independent information arrivals ($\tilde{\pi}_t = 0$) the expected pricing error is additive:

$$\mathbb{E}\left(\left|v_t - v_t^i\right| + \left|v_t - v_t^j\right|\right) = \frac{\sigma}{\sqrt{2}}\left(\frac{1}{\sqrt{\lambda_i}} + \frac{1}{\sqrt{\lambda_j}}\right) \tag{46}$$

$$= \sqrt{\frac{\sigma^2}{2}\mathbb{E}\left(\tilde{\tau}\right)}\left(\sqrt{\frac{\lambda_i+\lambda_j}{\lambda_i}} + \sqrt{\frac{\lambda_i+\lambda_j}{\lambda_j}}\right). \tag{47}$$

Here, $\mathbb{E}\left(\tilde{\tau}\right) := \frac{1}{\lambda_i+\lambda_j}$ is the expected inter-arrival of information at the economy and the adjustment terms $\left(\sqrt{\frac{\lambda_i+\lambda_j}{\lambda_i}}+\sqrt{\frac{\lambda_i+\lambda_j}{\lambda_j}}\right)>1$ are related to the odds of information arriving on market$k$. Next, arbitrage activity increases the arrival rate of information at the individual markets. More specifically, recall that Assumption 2 implies that information arrives at the economy with rate $\tilde{\lambda}:=\lambda_i+\lambda_j$ and is then revealed on market$k$ with probability $\frac{\lambda_k}{\lambda_i+\lambda_j}$. Given an information event on market$j$, the probability of a trade is characterized by $\tilde{\pi}^i_\tau$ as of Lemma 1 which depends on the trading costs c and the equilibrium spreads $\tilde{S}^i_\tau$.

Therefore, by the properties of the exponential distribution, the (expected) arrival rate of information on market$i$ is $\tilde{\lambda}_i=\lambda_i+\lambda_j\mathbb{E}\left(\tilde{\pi}^i\right)>\lambda_i$ and similar $\tilde{\lambda}_j=\lambda_j+\lambda_i\mathbb{E}\left(\tilde{\pi}^j\right)>\lambda_j$. Replacing λ_i and λ_j in the adjustment term of Equation (47) concludes the proof. □

Proof of Lemma 3. The proof of Lemma 3 follows immediately from Lemma 2 in combination with Assumption 3. □

Extensions

A Market makers with cross-market monitoring capacities

Fragmentation in my theoretical framework resembles a strict separation of the market participants at the two markets which can be overcome only by arbitrageurs. More specifically, I restrict local market makers to observe quotes at the other market, respectively. Arguably, a more realistic trading infrastructure presumably imposes less severe restrictions on the monitoring capacities of local market makers. Instead, assume that market makers can observe their own quotes and the quotes of the competing market at all times but they are restricted from providing liquidity at both markets. Then, market maker $k\in\{i,j\}$ instantaneously reacts to a shift in the quotes of market $k'\neq k$ and limited price informativeness due to market fragmentation does not play a role. However, if I instead impose noisy information revelation in the spirit of Foucault et al. (2017), price differences may persist and the thread of adverse selection remains active. More specifically, assume that at random times $\tilde{\tau}_l^k$, market makers at k receive a private valuation shock of magnitude $\gamma \sim N(0, \sigma_\gamma)$. Then from the perspective of market maker k, quoted prices at the opposite market only reveal information about the aggregate valuation $v_t^{k'} + \gamma$. Further, from the perspective of market maker k' the uncertainty with respect to the efficient price process v_t remains present and causes her to set spreads accordingly. In line with Foucault et al. (2017) such a setup does not reveal the full information set of market maker k' to market maker k and therefore price differences remain until they are either dissolved by new information or by an arbitrage trade.

B Details regarding assumption 2

Proposition 2. *Assumption 2 implies an equivalent distribution of arrival times at the individual markets as assuming that information arrives at times* $\{0, t_1, \ldots, t_n\}$ *and is then revealed on marketk* $\in \{i, j\}$ *with probability* $\lambda_k/(\lambda_i + \lambda_j)$.

Proof. In the following I show that the probability density function of the event $\tau^k = x$, $x > 0$ is equivalent to an exponential distribution with parameter λ_k. Define $\lambda; = \lambda_i + \lambda_j$. First $\tau^k = x$ holds if and only if all information arrivals until τ^k occurred at the opposite market. Further, the sum of s independent and identical exponential distributed variables with scale parameter λ follows the Erlang distribution with probability density function

$\pi(x) = \frac{\lambda^s x^{s-1}\exp(-\lambda x)}{(s-1)!}$. Then it holds that

$$\pi\left(\tau^k = x\right) = \sum_{s=1}^{\infty} \pi\left(s \text{ trials}\right) \pi\left(\sum_{i=1}^{s} \tau_i = x\right) \tag{48}$$

$$= \sum_{s=1}^{\infty} \left(1 - \frac{\lambda_k}{\lambda}\right)^{s-1} \frac{\lambda_k}{\lambda} \frac{\lambda^s}{(s-1)!} x^{s-1} \exp^{-\lambda x} \tag{49}$$

$$= \lambda_k \exp^{-\lambda x} \sum_{s=1}^{\infty} \left(\lambda - \lambda_k\right)^{s-1} x^{s-1} \tag{50}$$

$$= \lambda_k \exp^{-\lambda x} \exp^{(\lambda - \lambda_k)x} \tag{51}$$

$$= \lambda_k \exp^{-\lambda_k x}. \tag{52}$$

Therefore, $\pi\left(\tau^k = x\right)$ corresponds to an exponential distributed random variable with scale parameter λ_k which concludes the proof. □

C MCMC Algorithm

The following section illustrates the Monte Carlo Markov Chain Algorithm to calibrate the three-regime threshold vector error correction model. In line with the notation of Equation (21) the model can be rewritten as a (stationary) multivariate linear regression

$$\Delta V^r = X^r B^r + U^r \tag{53}$$

where

$$\Delta V^r_{t_r} = \begin{pmatrix} \Delta v^i_{t_r} \\ \Delta v^j_{t_r} \end{pmatrix}' \text{ and } X^r_{t_r} = \begin{pmatrix} 1 \\ z_{t_r - 1} \end{pmatrix}'. \tag{54}$$

Here, t_r corresponds to the stacked dates of all observations in regime r. Therefore, the data is of the form $\Delta V \in \mathbb{R}^{T\times 2}$ and $X \in \mathbb{R}^{T\times 2}$. The data is separated into three regimes by the thresholds c_t^{neg} and c_t^{pos} such that

$$\Delta V^{\text{neg}} := \{Y_t : z_{t-1} < c_t^{\text{neg}}\} \text{ and } X^{\text{neg}} := \{X_t : z_{t-1} < c_t^{\text{neg}}\} \tag{55}$$

$$\Delta V^{\text{pos}} := \{Y_t : z_{t-1} > c_t^{\text{pos}}\} \text{ and } X^{\text{pos}} := \{X_t : z_{t-1} > c_t^{\text{pos}}\} \tag{56}$$

$$\Delta V^{0} := \{Y_t : c_t^{\text{neg}} < z_{t-1} < c_t^{\text{pos}}\} \text{ and } X^{0} := \{X_t : c_t^{\text{neg}} < z_{t-1} < c_t^{\text{neg}}\} \tag{57}$$

The respective size of the partitioned matrices is $\Delta V^0 \in \mathbb{R}^{T^0 \times 2}$, $\Delta V^{\text{pos}} \in \mathbb{R}^{T^{\text{pos}} \times 2}$ and $\Delta V^{\text{neg}} \in \mathbb{R}^{T^{\text{neg}} \times 2}$ with $T^0 + T^{\text{pos}} + T^{\text{neg}} = T$. Thus, for $i \in \{\text{neg}, 0, \text{pos}\}$, the underlying data-generating process takes the form:

$$\Delta V^r = X^r \beta^r + U^r \text{ with } U_t^R \sim MN(0, \Sigma^r) \tag{58}$$

where $MN(\cdot)$ corresponds to the multivariate normal distribution with probability density function with zero mean

$$\pi\left(x\right) = \pi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \left(2\pi\right)^{-1} \det\left(\Sigma^r\right)^{-\frac{1}{2}} \exp\left(-\left('\Sigma^{r-1} x\right)\right). \tag{59}$$

Conditional on the parameters $c_0^{\text{pos}}, c_0^{\text{neg}}$ and c_1, standard inference from Bayesian multivariate linear regression models applies. First, the likelihood takes the form

$$\mathcal{L}(\Delta V|\theta, X) \propto \prod_{r \in \{\text{neg, 0, pos}\}} |\Sigma^r|^{-T^r/2} \exp\left(-\frac{1}{2}\text{tr}\left(\left(\Sigma^r\right)^{-1} U_\theta^{r\prime} U_\theta^r\right)\right) \tag{60}$$

Then, conditional conjugate priors for $\beta^r = \text{vec}\,(B^r)$ and Σ^r are chosen such that for suitable hyperparameters $\beta = \text{vec}(B), \Psi$, and ν I have:

$$\pi(\Sigma^r) \sim IW(\Psi,\nu) = \frac{|\Psi|^{\nu/2}}{2^{\nu p/2} c_p(\frac{\nu}{2})} |\Sigma^r|^{-(\nu+p+1)/2} \exp\left(-\frac{1}{2}\,\text{tr}(\Psi\,(\Sigma^r)^{-1})\right) \tag{61}$$

$$\pi(\beta^r|\Sigma^r) \sim MN(\beta, \Sigma^r \otimes \Lambda^{-1}) \tag{62}$$

The priors for c_0^r and c_1 are uniform such that

$$\pi(c_1) \sim \pi(c_0^r) \sim U\,(-\infty,\infty)\,. \tag{63}$$

A standard Gibbs sampling scheme applies for the posterior when conditioning on the threshold variables c_t^{pos} and c_t^{neg}. Given initial (or sampled) values of $\{c_t^{\text{pos}}, c_t^{\text{neg}}\}$ the algorithm works as follows:

- Separate the data ΔV and X into $\Delta V^{\text{neg}}, \Delta V^0, \Delta V^{\text{pos}}$ and $X^{\text{neg}}, X^0, X^{\text{pos}}$.
- For each of the three regimes $r \in \{\text{neg, 0, pos}\}$ generate a draw Σ^r from the Inverse Wishart distribution based on the posterior distribution

$$\pi(\Sigma^r|\Delta V, X) \sim IW(\tilde{\Psi}, \tilde{\nu}) \tag{64}$$

where

$$\tilde{\Psi} = \Psi + (\Delta V - X\tilde{B})'(\Delta V - X\tilde{B}) + (\tilde{B} - B)'\Lambda(\tilde{B} - B) \tag{65}$$

$$\tilde{\nu} = \nu + T^r \tag{66}$$

$$\tilde{B} = (X'X + \Lambda)^{-1}(X'\Delta V + \Lambda B) \tag{67}$$

$$\tag{68}$$

- For each of the three regimes $r \in \{\text{neg, 0, pos}\}$ generate a draw β from the multivariate normal distribution based on the (conditional) posterior distribution

$$\pi\left(\beta^r|\Sigma^r\right) \propto MVN\left(\text{vec}\left(\tilde{B}\right), \Sigma^r \otimes \left(X^{r\prime}X^r + \Lambda\right)^{-1}\right) \tag{69}$$

- Random walk Metropolis Hastings step within Gibbs: Sample $\tilde{c}_0^{\text{neg}}, \tilde{c}_0^{\text{pos}}$ and $\tilde{c}_1$ independently from a normal distribution with means $c_0^{\text{neg}}, c_0^{\text{pos}}$ and c_1 and sampling variances σ_0 and σ^1. Compute the acceptance ratio

$$\alpha(\{\tilde{c}_0^{\text{neg}}, \tilde{c}_0^{\text{pos}}, \tilde{c}_1\}|\{c_0^{\text{neg}}, c_0^{\text{pos}}, c_1\}) = \min\left(1, \frac{\mathcal{L}(\Delta V|\tilde{c}_t, \beta^0, \beta^{\text{pos}}, \beta^{\text{neg}}, \Sigma^0, \Sigma^{\text{neg}}, \Sigma^{\text{pos}}, X)}{\mathcal{L}(\Delta V|c_t, \beta^0, \beta^{\text{pos}}, \beta^{\text{neg}}, \Sigma^0, \Sigma^{\text{neg}}, \Sigma^{\text{pos}}, X)}\right). \tag{70}$$

  Letting $\tilde{T}^r$ the sample size of ΔV based on the new segmentation due to $\tilde{c}_t$ and $\tilde{U}^r$ the corresponding residuals $\Delta\tilde{V}^r - \tilde{X}^r\beta^r$, I get:

$$\log\left(\alpha(\{\tilde{c}_0^{\text{neg}}, \tilde{c}_0^{\text{pos}}, \tilde{c}_1\}\right) = \frac{1}{2}\sum_r\left(\left(T^r - \tilde{T}^r\right)\log|\Sigma^r| + \text{tr}(\Sigma^{r-1}(U^{r\prime}U^r - \tilde{U}^{r\prime}\tilde{U}^r))\right). \tag{71}$$

D Network activity and cross-market trading

Settlement latency affects cross-market arbitrageurs because they cannot dispose of their position before the sell-side exchange accepts their Bitcoin deposit as valid. Whereas the Bitcoin blockchain is public, trades are usually not disclosed directly. However, to provide more information regarding the cross-market Bitcoin flows, I collect a list of wallets which are likely under the control of the exchanges in my sample. Although Bitcoin transactions are pseudonymous in the sense that the transactions publicly reveal all addresses associated with a transaction, but it is hard to map these addresses to their

Table A1: Descriptive Statistics of the Bitcoin Network
This table reports descriptive statistics of our Bitcoin transaction data. The sample contains all transactions settled in the Bitcoin network from March 1st, 2018, until August 31, 2019. *Fee per Byte* is the total fee per transaction divided by the size of the transaction in bytes in Satoshi where 100,000,000 Satoshi are 1 Bitcoin. *Fee per Transaction* is the total settlement fee per transaction (in USD). I approximate the USD price by the average minute-level midquote across all exchanges in our sample. *Latency* is the time until the transaction is either validated or leaves the mempool without verification (in minutes). *Transaction Size* denotes the size of the transaction in bytes. *Mempool Size* is the number of other transactions in the mempool at the time a transaction of our sample enters the mempool.

		Mean	SD	5 %	25 %	Median	75 %	95 %
Standard	Fee per Byte (Satoshi)	19.23	180.29	1.32	3.59	7.81	18.21	62.47
	Fee per Transaction (USD)	0.52	5.70	0.02	0.06	0.13	0.31	1.76
	Latency	24.69	100.48	0.73	3.53	8.67	19.48	78.73
	Transaction Size	519.14	2191.75	192.00	225.00	247.00	372.00	964.00
Arbitrage	Fee per Byte (Satoshi)	44.52	78.38	3.39	8.22	14.46	39.69	238.66
	Fee per Transaction (USD)	4.91	27.84	0.06	0.17	0.74	2.97	9.86
	Latency	12.39	24.70	0.50	3.00	7.27	14.70	36.53
	Transaction Size	1708.47	4342.36	223.00	249.00	424.00	1172.00	5220.00

respective physical or legal owners.[17] The wallet IDs allows me to identify all transactions in which funds have presumably been moved between two exchange-controlled wallets. All cross-market arbitrage activity should therefore be a subset of the transactions included in this sample (in principle, agents could also move funds between trading platforms due to reasons which are not related to exploiting price differences). Table A1 provides comparable statistics of the underlying characteristics of the two subsets of transactions. *Standard* comprises of all transactions of the Bitcoin network and *Arbitrage* comprises of all transactions which are flagged as being potential cross-market fund flows. The table reveals that in general cross-exchange flows settle faster, the average latency in minutes only comprises about 50% of the latency of the entire network. However, this does not mean that arbitrageurs can move funds between two exchanges more efficient than other market participants. The table reveals that fees paid for cross-exchange flows are on average about 5 USD, roughly 10 times as much as the average transaction fee in the entire network. Therefore, cross-exchange flows are settled faster because they are more valuable from the perspective of miners. However, these settlement fees resemble trading

[17]More specific information on the data, I refer to Hautsch et al. (2019).

costs for arbitrageurs which should reflect the trade-off between paying higher fees and getting faster settlement.