Limits to Arbitrage in Markets with Stochastic Settlement Latency

Nikolaus Hautsch
University of Vienna
Vienna Graduate School of Finance

Christoph Scheuch
Vienna Graduate School of Finance

Stefan Voigt
Vienna Graduate School of Finance
What makes a blockchain market different from a ”traditional” market?

- No intermediaries (e.g., for clearing)
- Settlement much faster
- Trust through proof of work
- Security through complexity of verification
- On a public blockchain, transactions are (semi-)transparent
- ...
Blockchain technology and stochastic latency

In "classical" markets, process of trading and process of ownership transfer are disconnected.

- Both transacting parties face counterparty risks during settlement period (1-2 trading days).
- (Third-party) central clearing allows for continuous (high-frequency) trading on non-settled positions.

In (pure) distributed ledger systems, transfer of ownership is connected to trading.

- Funds cannot be further transferred before transaction is validated.
- Process of validation takes time (>30min): Consensus algorithms introduce stochastic latency

What are the implications of stochastic latency for arbitrage?
Stochastic latency through the blockchain

Market $b$
Low Price

Market $s$
High Price

Arbitrageur
Stochastic latency through the blockchain
Stochastic latency through the blockchain

Stochastic Latency

Transfer + Settlement

Market_b
Low Price

Market_s
High Price

Arbitrageur

Buy Order
Instantaneous

Sell Order
Instantaneous
Introduction

Arbitrage possibilities in BTC-USD trading on May 25, 2018?
This paper

Research question:

▶ How much of observed cross-market price differences are explained by limits to arbitrage due to (stochastic) latency?

This paper:

▶ Derivation of arbitrage bounds for markets with stochastic latency
▶ In-depth analysis of Bitcoin network based on limit order book data of 16 exchanges
▶ Estimation of arbitrage bounds and analysis of arbitrage opportunities in the BTC vs. USD market
2. Stochastic Latency and Limits to Arbitrage
Market $i \in \{1, \ldots, N\}$ continuously provides buy quotes (ask) $A^i_t$ and sell quotes (bid) $B^i_t$ for one unit of the asset where $B^i_t \leq A^i_t$ at time $t$.

No short selling, margin trading or derivatives

Arbitrageur continuously monitors the quotes on markets $b$ and $s$.

Instantaneous trading: Arbitrageur exploits price differences if

$$B^s_t - A^b_t > 0$$

Stochastic latency $\tau$ is the random waiting time until a transfer of the asset between markets is settled.

Profit of arbitrageur’s trading decision if

$$\mathbb{P}\left( B^s_{t+\tau} \leq A^b_t \mid \mathcal{I}_t \right) > 0.$$
2. Stochastic Latency and Limits to Arbitrage

Log return of arbitrageur’s strategy

\[
\begin{align*}
\log r_{(t:t+\tau)}^{b,s} &= b_{t+\tau}^s - a_t^b = \\
&= \underbrace{\delta_t^{b,s}}_{\text{instantaneous return}} + \underbrace{b_{t+\tau}^s - b_t^s}_{\text{exposure to price risk}},
\end{align*}
\]

where \( \delta_t^{b,s} = b_t^s - a_t^b \)

**Assumption 1.** For given latency \( \tau \), we model the log price changes on the sell-side \( b_{t+\tau}^s - b_t^s \) as a Brownian motion with drift \( \mu_t^s \) such that

\[
\log r_{(t:t+\tau)}^{b,s} = \delta_t^{b,s} + \tau \mu_t^s + \int_t^{t+\tau} \sigma_t^s dW_k^s,
\]

We assume that \( \mu_t^s \) and \( \sigma_t^s \) are locally constant over the interval \([t, t + \tau]\).

**Assumption 2.** Stochastic latency \( \tau \in \mathbb{R}_+ \) is a random variable equipped with a (conditional) probability distribution \( \pi_t(\tau) := \pi(\tau | \mathcal{I}_t) \). We assume that all moments of \( \tau \) are finite.
Lemma. Under Assumptions 1 and 2, the arbitrage returns follow a normal mean-variance mixture with probability distribution

\[
\pi_t \left( r^{b,s}_{(t:t+\tau)} \right) = \int_{\mathbb{R}_+} \pi_t \left( r^{b,s}_{(t:t+\tau)} | \tau \right) \pi_t (\tau) \, d\tau
\]

and characteristic function \( \varphi_x(t) = \mathbb{E}[e^{itx}] \),

\[
\varphi_{r^{b,s}_{t:t+\tau}} (u) = e^{iu\delta^{b,s}_t} m_\tau \left( iu\mu^s_t - \frac{1}{2}u^2(\sigma^s_t)^2 \right),
\]

where \( m_\tau (u) := \mathbb{E}_t (e^{u\tau}) \) is the moment-generating function of \( \pi_t (\tau) \).
Example: Exponential distribution

- P.d.f. of stochastic latency: $\pi_t(\tau) = \lambda_t e^{-\lambda_t \tau}$, with moment generating function $m_{\tau}(u) = \left(1 - \lambda_t^{-1} u\right)^{-1}$.
- The characteristic function of the mixture is

$$\varphi_{r_{t:t+\tau}}^{b,s}(u) = \frac{e^{iu\delta_{t}^{b,s}}}{1 - i\frac{\mu_{t}^{s}}{\lambda_{t}} u + \frac{(\sigma_{t}^{s})^{2}}{2\lambda_{t}} u^{2}}$$

corresponding to the characteristic function of an asymmetric Laplace distribution (Kotz, Kozubowski, Podgorski, 2001) with

\[
\mathbb{E}_t (r_{t:t+\tau}^{b,s}) = \delta_{t}^{b,s} + \frac{\mu_{t}^{s}}{\lambda_{t}}
\]

\[
\nabla_t (r_{t:t+\tau}^{b,s}) = \frac{1}{\lambda_{t}} \left( (\mu_{t}^{s})^{2} + (\sigma_{t}^{s})^{2} \right).
\]
Distribution of returns under exponential latency and negative drift
⇒ Asymmetric Laplace
Example: Inverse gamma distribution

- P.d.f. of inverse Gamma latency:

\[
\pi_t(\tau) = \frac{\beta_t^\alpha}{\Gamma(\alpha)} \tau^{-(\alpha+1)} \exp\left(- \frac{\beta_t}{\tau}\right),
\]

with \( \mathbb{E}_t(\tau) = \frac{\beta_t}{\alpha-1} \) and \( \mathbb{V}_t(\tau) = \frac{\beta_t^2}{(\alpha-1)^2(\alpha-2)} \), where \( \Gamma(\alpha) \) denotes the Gamma function.

- Then, the conditional distribution of the returns of the arbitrageur’s future returns \( r_{(t:t+\tau)}^{b,s} \) corresponds to a non-standard Student’s \( t \) distribution with p.d.f. given by

\[
\pi_t\left(r_{(t:t+\tau)}^{b,s}\right) = \frac{\Gamma\left(\frac{2\alpha+1}{2}\right)}{\Gamma(\alpha) \sqrt{2\pi(\sigma_t^s)^2}\beta_t} \left(1 + \frac{\left(r_{(t:t+\tau)}^{b,s} - \delta_t^{b,s}\right)^2}{\alpha(\sigma_t^s)^2\beta_t}\right)^{-\frac{2\alpha+1}{2}},
\]

with \( \mathbb{E}_t\left(r_{(t:t+\tau)}^{b,s}\right) = \delta_t^{b,s} \) and \( \mathbb{V}_t\left(r_{(t:t+\tau)}^{b,s}\right) = \frac{(\sigma_t^s)^2\beta_t}{\alpha-1} \).
2. Stochastic Latency and Limits to Arbitrage

Risk Aversion ...
2. Stochastic Latency and Limits to Arbitrage

**Assumption 3.** The arbitrageur has an utility function $U_\gamma(r)$ with risk aversion parameter $\gamma > 1$, and $U'_\gamma(r) > 0$ and $U''_\gamma(r) < 0$.

Arbitrageur exploits price differences if and only if the certainty equivalent of trading,

$$\mathbb{E}\left( U_\gamma \left( r_{(t:t+\tau)}^{b,s} \right) \right) = U_\gamma (CE),$$

is positive.

**Theorem.** Under Assumptions 1-3, the certainty equivalent (CE) is given by

$$CE = \delta_t^{b,s} + \mathbb{E}_t(\tau)\mu_t^s \quad + \quad \sum_{k=2}^{\infty} \left( \frac{U^{(k)}_\gamma \left( \delta_t^{b,s} + \mathbb{E}_t(\tau)\mu_t^s \right)}{k!U'_\gamma \left( \delta_t^{b,s} + \mathbb{E}_t(\tau)\mu_t^s \right)} \mathbb{E}\left( \left( r_{(t:t+\tau)}^{b,s} - \delta_t^{b,s} - \mathbb{E}_t(\tau)\mu_t^s \right)^k \right) \right),$$

where $U^{(k)}_\gamma (\mu_r) := \frac{\partial^k}{\partial \mu_r^k} U_\gamma (\mu_r)$.
Arbitrage bound

The arbitrage boundary $d_t^s$ is defined as the minimum price difference necessary such that the arbitrageur prefers to trade if $\delta_{t}^{b,s} > d_t^s$.

**Definition.** $d_t^s$ is the maximum of zero and the (unique) root of

$$F(d) = d + \mathbb{E}_t(\tau)\mu_t^s$$

$$+ \sum_{k=2}^{\infty} \frac{U_{\gamma}^{(k)}}{k!U_{\gamma}'} \left( d + \mathbb{E}_t(\tau)\mu_t^s \right) \mathbb{E}_t \left( \left( r_{(t:t+\tau)}^{b,s} - d - \mathbb{E}_t(\tau)\mu_t^s \right)^k \right)$$

Price differences below the arbitrage boundary might persist as the arbitrageur prefers not to trade in such a scenario.
Constant absolute risk aversion (CARA)

Assumption. The arbitrageur is equipped with constant absolute risk aversion and exponential utility function $U_\gamma(r) := \frac{1-e^{-\gamma r}}{\gamma}$ with risk aversion parameter $\gamma > 1$.

Lemma. In case of an exponential utility function and the assumptions above, price differences are exploited if

$$CE > 0 \iff \delta_{t}^{b,s} > d_{t}^{s},$$

with

$$d_{t}^{s} := -\mathbb{E}_t (\tau) \mu_{t}^{s} + \frac{\gamma}{2} \left( \mathbb{V}_t (\tau) \left( \mu_{t}^{s} \right)^2 + (\sigma_{t}^{s})^2 \mathbb{E}_t (\tau) \right)$$

$$- \frac{\gamma^2}{6} \left( 3\mu_{t}^{s} (\sigma_{t}^{s})^2 \mathbb{V}_t (\tau) + (\mu_{t}^{s})^3 \mathbb{E}_t ((\tau - \mathbb{E}_t (\tau))^3) \right)$$

$$+ \frac{\gamma^3}{24} \left( (\mu_{t}^{s})^4 \mathbb{E}_t ((\tau - \mathbb{E}_t (\tau))^4) + 6 (\sigma_{t}^{s})^2 (\mu_{t}^{s})^2 (\mathbb{E} (\tau)^3 + \mathbb{E}_t (\tau^3) - 2\mathbb{E}_t (\tau) \mathbb{E}_t (\tau)) \right)$$

$$+ \frac{\gamma^3}{24} \left( 3\mathbb{E}_t (\tau^2) (\sigma_{t}^{s})^4 \right)$$
Lemma. Under the above assumptions with $\mu_s^t = 0$, exponential utility and any well defined latency distribution, the arbitrage bound is given by

$$d_t^s = \frac{1}{2} \gamma (\sigma_s^t)^2 \mathbb{E}_t(\tau) + \frac{1}{8} \gamma^3 (\sigma_s^t)^4 \mathbb{E}_t(\tau^2).$$

Stochastic latency implies limits to arbitrage, which increase if

- spot volatility is high
- expected latency is large
- latency uncertainty is high
- risk aversion is high
Constant relative risk aversion (CRRA)

- Arbitrageur has power utility function

\[ U_\gamma(r) := \frac{r^{1-\gamma} - 1}{1 - \gamma}, \]

with relative risk aversion parameter \( \gamma > 1 \).

- Then, the arbitrage bound for \( \mu_t^s = 0 \) is given by

\[ d_t^s = \frac{1}{2} \sigma_t^s \sqrt{\gamma \mathbb{E}_t(\tau) + \sqrt{\gamma^2 \mathbb{E}_t(\tau)^2 + 2\gamma(\gamma + 1)(\gamma + 2) \mathbb{E}_t(\tau^2)}}. \]
3. Transaction Costs & Settlement Fees
3. Transaction Costs & Settlement Fees

Transaction costs

▶ Consider proportional transaction costs of the form

\[ B_i^t(q) = B_i^t \left( 1 - \rho_i^{i,B}(q) \right) \]
\[ A_i^t(q) = A_i^t \left( 1 + \rho_i^{i,A}(q) \right), \]

with \( \rho_i^{i,B}(q) > 0 \) and \( \rho_i^{i,A}(q) > 0 \) denoting transaction cost functions in dependence of the trading volume \( q > 0 \).

▶ Log return of arbitrage strategy changes to

\[ \tilde{r}_b^{b,s}(t:t+\tilde{\tau}) = r_b^{b,s}(t:t+\tilde{\tau}) - \ln \left( \frac{1 + \rho_b^{b,A}(q)}{1 - \rho_s^{s,B}(q)} \right). \]

▶ **Lemma.** Arbitrageur exploits price differences if

\[ \delta_t^{b,s} - \ln \left( \frac{1 + \rho_b^{b,A}(q)}{1 - \rho_s^{s,B}(q)} \right) > d_t^s \]

for given quantity \( q > 0 \).
Latency-reducing settlement fees

- **Assumption.** A settlement fee $f > 0$ implies a latency distribution $\pi_t(\tau|f)$ that can be ordered in the sense that for $\tilde{f} > f$, $\pi_t(\tau|f)$ first-order stochastically dominates $\pi_t(\tau|\tilde{f})$, i.e.,

$$\Pr(\tau \leq x|\tilde{f}) > \Pr(\tau \leq x|f) \text{ for all } x \in \mathbb{R}_+.$$ 

- Denote by $d_{s}^{s}(f)$ the arbitrage boundary associated with the latency distribution $\pi_t(\tau|f)$. Then, $d_{s}^{s}(f) > d_{s}^{s}(\tilde{f})$.

- Arbitrageur’s trading decision features trade-off between $q$ and $f$ with endogenous arbitrage boundaries.

- **Lemma.** Under the given assumptions the arbitrageur prefers to trade a quantity $q > 0$ and pay a settlement fee $f > 0$ over staying idle if

$$\delta_{t}^{b,s} - \ln \left( \frac{1 + \rho^{b,A}(q + f)}{1 - \rho^{s,B}(q)} \right) > d_{s}^{s}(f).$$
3. Transaction Costs & Settlement Fees

Arbitrageur maximizes

$$\max_{\{q,f\}\in \mathbb{R}^2_+} B^s_t \left(1 - \rho^{s,B}(q)\right) q - A^b_t (1 + \rho^{b,A}(q + f))(q + f)$$

subject to

$$\delta^{b,s}_t - \ln \left(\frac{1 + \rho^{b,A}(q + f)}{1 - \rho^{s,B}(q)}\right) \geq d^s_t(f).$$

**Lemma.** A total return maximizing arbitrageur chooses trading quantities $q^* > 0$ and settlement fees $f^* \geq 0$ such that

$$\delta^{b,s}_t - \ln \left(\frac{1 + \rho^{b,A}(q^* + f^*)}{1 - \rho^{s,B}(q^*)}\right) = d^s_t(f^*).$$

He pays $f^* > 0$ to trade $q^* > 0$ if the following necessary conditions are met:

$$\frac{1 - \rho^{s,B}(q^*)}{q^*} > \rho^{s,B'}(q^*)$$

$$-\frac{\partial}{\partial f} d^s_t(f^*) > \frac{\rho^{s,B'}(q^*)}{1 + \rho^{s,B}(q^*)}.$$

Otherwise, the arbitrageur optimally sets $f^* = 0$. 
4. Data
4. Data

Bitcoin orderbook data

Minute-level Bitcoin/Dollar orderbooks from 16 large exchanges (≈ 95% of trading volume)

Buy and sell orders for the first 25 levels since April 2018.

<table>
<thead>
<tr>
<th>Orderbooks</th>
<th>Spread (USD)</th>
<th>Spread (bp)</th>
<th>Taker Fee</th>
<th>With. Fee</th>
<th>Conf.</th>
<th>Margin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binance</td>
<td>677,930</td>
<td>2.88</td>
<td>3.77</td>
<td>0.10</td>
<td>0.10</td>
<td>2</td>
</tr>
<tr>
<td>Bitfinex</td>
<td>676,019</td>
<td>0.52</td>
<td>0.68</td>
<td>0.20</td>
<td>0.08</td>
<td>3</td>
</tr>
<tr>
<td>bitFlyer</td>
<td>656,244</td>
<td>12.80</td>
<td>19.95</td>
<td>0.15</td>
<td>0.08</td>
<td>TRUE</td>
</tr>
<tr>
<td>Bitstamp</td>
<td>675,256</td>
<td>4.53</td>
<td>6.09</td>
<td>0.25</td>
<td>0.00</td>
<td>3</td>
</tr>
<tr>
<td>Bittrex</td>
<td>677,298</td>
<td>10.13</td>
<td>15.63</td>
<td>0.25</td>
<td>0.00</td>
<td>2</td>
</tr>
<tr>
<td>CEX.IO</td>
<td>674,595</td>
<td>10.37</td>
<td>14.56</td>
<td>0.25</td>
<td>0.10</td>
<td>3</td>
</tr>
<tr>
<td>Gate</td>
<td>644,379</td>
<td>108.80</td>
<td>122.00</td>
<td>0.20</td>
<td>0.20</td>
<td>2</td>
</tr>
<tr>
<td>Coinbase Pro</td>
<td>678,216</td>
<td>0.14</td>
<td>0.24</td>
<td>0.30</td>
<td>0.00</td>
<td>3</td>
</tr>
<tr>
<td>Gemini</td>
<td>651,425</td>
<td>2.00</td>
<td>3.08</td>
<td>1.00</td>
<td>0.20</td>
<td>3</td>
</tr>
<tr>
<td>HitBTC</td>
<td>656,195</td>
<td>3.19</td>
<td>4.20</td>
<td>0.10</td>
<td>0.08</td>
<td>2</td>
</tr>
<tr>
<td>Kraken</td>
<td>673,730</td>
<td>2.74</td>
<td>3.56</td>
<td>0.26</td>
<td>0.10</td>
<td>6</td>
</tr>
<tr>
<td>Liqui</td>
<td>491,516</td>
<td>30.15</td>
<td>45.13</td>
<td>0.25</td>
<td></td>
<td>TRUE</td>
</tr>
<tr>
<td>Lykke</td>
<td>655,407</td>
<td>34.03</td>
<td>51.37</td>
<td>0.00</td>
<td>0.05</td>
<td>3</td>
</tr>
<tr>
<td>Poloniex</td>
<td>654,104</td>
<td>5.46</td>
<td>8.25</td>
<td>0.20</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>xBTCe</td>
<td>623,912</td>
<td>8.36</td>
<td>13.67</td>
<td>0.25</td>
<td>0.30</td>
<td>3</td>
</tr>
</tbody>
</table>
Avg. price differences (after transaction costs)
Average cross-market price differences
### Bitcoin blockchain mempool

- Data from [www.blockchain.com](http://www.blockchain.com)
- All confirmed blocks from January 2018 until April 2019
- 92,626,780 transactions verified in 71,992 blocks
- For each transaction: unique ID, size, fee, latency

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>5 %</th>
<th>25 %</th>
<th>Median</th>
<th>75 %</th>
<th>95 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fee per Byte (Satoshi)</td>
<td>48.79</td>
<td>219.48</td>
<td>1.53</td>
<td>4.59</td>
<td>10.75</td>
<td>32.00</td>
<td>282.49</td>
</tr>
<tr>
<td>Fee per Transaction (USD)</td>
<td>2.09</td>
<td>29.13</td>
<td>0.02</td>
<td>0.07</td>
<td>0.18</td>
<td>0.62</td>
<td>10.84</td>
</tr>
<tr>
<td>Latency</td>
<td>38.71</td>
<td>335.02</td>
<td>0.73</td>
<td>3.55</td>
<td>8.75</td>
<td>20.10</td>
<td>92.42</td>
</tr>
<tr>
<td>Transaction Size</td>
<td>527.57</td>
<td>2274.81</td>
<td>192.00</td>
<td>225.00</td>
<td>247.00</td>
<td>372.00</td>
<td>963.00</td>
</tr>
<tr>
<td>Mempool Size</td>
<td>8437.26</td>
<td>14438.03</td>
<td>324.00</td>
<td>1336.00</td>
<td>3429.50</td>
<td>8064.50</td>
<td>39415.00</td>
</tr>
<tr>
<td>Block Validation Time</td>
<td>9.70</td>
<td>9.62</td>
<td>0.55</td>
<td>2.85</td>
<td>6.77</td>
<td>13.42</td>
<td>28.95</td>
</tr>
</tbody>
</table>
5. Quantifying Arbitrage Bounds
5. Quantifying Arbitrage Bounds

Limits to arbitrage for CRRA case are given by

$$\hat{d}_t^s = \frac{1}{2} \hat{\sigma}_t^s \sqrt{\gamma \hat{E}_t(\tau) + \sqrt{\gamma^2 \hat{E}_t(\tau)^2 + 2\gamma(\gamma + 1)(\gamma + 2)\hat{E}_t(\tau^2)}}.$$ 

**Ingredients:**

1. Spot volatility $(\hat{\sigma}_t^s)^2$
2. Estimates of $\hat{E}_t(\tau)$ and $\hat{E}_t(\tau)^2$.
3. Estimates of $\rho^{b,A}(q)$ and $\rho^{s,B}(q)$ for optimal $q$. 
Estimating spot volatilities $\sigma_t^s$

- Current volatility affects price risk of arbitrageur
  \[ db_t^s = \sigma_t^s dW_t^s \]

- Nonparametric filtering of the realized spot volatility (Kristensen, 2010)

- For each market $n$ and time $t$, we estimate $(\sigma_t^n)^2$ by
  \[ (\hat{\sigma}_t^n)^2 (h) = \sum_{s=1}^{t} K (s - t, h) (b_n^s - b_n^{s-1})^2, \]
  where $K (s - t, h)$ denotes a one-sided Gaussian kernel smoother with bandwidth $h$.

- Bandwidth $h$ chosen by minimizing the Integrated Squared Error (ISE)
  \[ \hat{\text{ISE}}_{T-1} (h) = \sum_{i=1}^{I} \left[ (b_i^n - b_{i-1}^n)^2 - (\hat{\sigma}_i^n)^2 (h) \right]^2, \]
  where $i = 1, \ldots, I$ refers to the observations on day $T - 1$ and $(\hat{\sigma}_t^n)^2 (h)$ is the spot volatility estimator based on bandwidth $h$. 
5. Quantifying Arbitrage Bounds

Daily cross-exchange averages of spot volatilities

Spot Volatility (in %)
Parameterizing waiting times $\pi(\tau_t | \mathcal{I}_t)$

- **Exponential model:**

\[
\pi_t(\tau_i) = \lambda_i \exp(-\lambda_i \tau_i),
\]
\[
\lambda_i = \exp(-x'_i \theta_t),
\]

- **Gamma model:**

\[
\pi_t(\tau_i) = \frac{\beta_i \alpha}{\Gamma(\alpha)} \tau^{\alpha-1} e^{-\beta_i \tau_i}
\]
\[
\beta_i = \exp(-x'_i \theta_t), \alpha > 0
\]

- **Covariates** $x_i$: number of unconfirmed transactions in mempool; network fee per byte
### Parameter estimates

<table>
<thead>
<tr>
<th></th>
<th>Exponential</th>
<th></th>
<th>Gamma</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>W/o Covariates</td>
<td>W/ Covariates</td>
<td>W/o Covariates</td>
<td>W/ Covariates</td>
</tr>
<tr>
<td>Intercept</td>
<td>3.02 [2.442, 3.687]</td>
<td>1.13 [-0.043, 2.549]</td>
<td>3.4 [2.514, 4.379]</td>
<td>1.15 [-0.021, 2.406]</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.7 [0.487, 0.95]</td>
<td>0.72</td>
<td>0.7 [0.487, 0.95]</td>
<td>0.72 [0.505, 0.945]</td>
</tr>
<tr>
<td>Fee per Byte</td>
<td>-0.04 [-0.082, -0.005]</td>
<td></td>
<td>-0.02 [-0.082, -0.005]</td>
<td>2.43 [0.535, 4.099]</td>
</tr>
<tr>
<td>Mempool Size</td>
<td>2.04 [0.221, 3.528]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LR (Covariates)</td>
<td>93.61</td>
<td></td>
<td>84.33</td>
<td></td>
</tr>
<tr>
<td>LR (Gamma vs. Exponential)</td>
<td>92.58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSPE (Out of Sample)</td>
<td>37.77</td>
<td>37.44</td>
<td>37.77</td>
<td>37.58</td>
</tr>
<tr>
<td>MSPE (In Sample)</td>
<td>35.82</td>
<td>35.22</td>
<td>35.82</td>
<td>35.29</td>
</tr>
</tbody>
</table>

- Higher fees reduce latency (Easley et al, 2019)
- Blockchain congestion increases latency
Estimation of arbitrage bounds

- Limits to arbitrage for CRRA case are given by

\[
d_t^s = \frac{1}{2} \hat{\sigma}_t^s \sqrt{\gamma c_1 + \sqrt{\gamma^2 c_1^2 + 2\gamma(\gamma + 1)(\gamma + 2)c_2}},
\]

\[
c_1 = \hat{E}_t(\tau) + \hat{E}(\tau_B) \cdot (B^s - 1)
\]

\[
c_2 = \hat{V}_t(\tau) + \hat{V}(\tau_B) \cdot (B^s - 1)^2 + \left(\hat{E}(\tau_B) \cdot (B^s - 1) + \hat{E}_t(\tau)\right)^2,
\]

where \( B^s \) refers to the number blocks that the sell-side exchange \( s \) requires to consider incoming transactions as valid.

- \( \hat{E}_t(\tau_i) \) and \( \hat{V}_t(\tau_i) \) computed based on (rolling window) day \( T - 1 \) parameter estimates.
Estimated arbitrage bounds over time ($\gamma = 2$)
Estimated arbitrage bounds (in bps, $\gamma = 2$)

- Average boundary about 96bps
- Latency uncertainty accounts for 9% on average

<table>
<thead>
<tr>
<th>Exchange</th>
<th>Mean</th>
<th>SD</th>
<th>5%</th>
<th>25%</th>
<th>Median</th>
<th>75%</th>
<th>95%</th>
<th>Uncertainty</th>
<th>Security</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binance</td>
<td>77.09</td>
<td>62.84</td>
<td>20.41</td>
<td>35.47</td>
<td>56.77</td>
<td>97.75</td>
<td>199.73</td>
<td>9.00</td>
<td>18.98</td>
</tr>
<tr>
<td>Bitfinex</td>
<td>84.01</td>
<td>69.10</td>
<td>15.38</td>
<td>35.37</td>
<td>62.42</td>
<td>113.17</td>
<td>222.94</td>
<td>8.46</td>
<td>31.49</td>
</tr>
<tr>
<td>bitFlyer</td>
<td>92.98</td>
<td>65.72</td>
<td>28.72</td>
<td>50.39</td>
<td>73.49</td>
<td>113.91</td>
<td>223.44</td>
<td>8.46</td>
<td>31.53</td>
</tr>
<tr>
<td>Bitstamp</td>
<td>91.55</td>
<td>69.67</td>
<td>24.35</td>
<td>44.05</td>
<td>69.46</td>
<td>118.17</td>
<td>230.76</td>
<td>8.56</td>
<td>31.14</td>
</tr>
<tr>
<td>Bittrex</td>
<td>98.30</td>
<td>63.45</td>
<td>28.79</td>
<td>54.26</td>
<td>83.23</td>
<td>124.56</td>
<td>219.35</td>
<td>8.57</td>
<td>20.11</td>
</tr>
<tr>
<td>CEX.IO</td>
<td>88.90</td>
<td>61.98</td>
<td>25.32</td>
<td>45.58</td>
<td>72.37</td>
<td>112.11</td>
<td>213.13</td>
<td>8.42</td>
<td>31.73</td>
</tr>
<tr>
<td>Gate</td>
<td>73.45</td>
<td>54.07</td>
<td>19.89</td>
<td>37.20</td>
<td>58.19</td>
<td>92.44</td>
<td>178.68</td>
<td>8.77</td>
<td>19.63</td>
</tr>
<tr>
<td>Gatecoin</td>
<td>185.74</td>
<td>211.13</td>
<td>2.70</td>
<td>44.90</td>
<td>112.77</td>
<td>253.63</td>
<td>606.39</td>
<td>9.14</td>
<td>50.98</td>
</tr>
<tr>
<td>Coinbase Pro</td>
<td>78.83</td>
<td>66.43</td>
<td>14.51</td>
<td>32.86</td>
<td>58.54</td>
<td>103.35</td>
<td>213.89</td>
<td>8.56</td>
<td>31.99</td>
</tr>
<tr>
<td>Gemini</td>
<td>80.81</td>
<td>66.12</td>
<td>17.48</td>
<td>35.99</td>
<td>60.55</td>
<td>103.73</td>
<td>216.48</td>
<td>8.57</td>
<td>31.68</td>
</tr>
<tr>
<td>HitBTC</td>
<td>66.70</td>
<td>56.45</td>
<td>15.39</td>
<td>30.05</td>
<td>49.73</td>
<td>83.98</td>
<td>175.08</td>
<td>8.84</td>
<td>19.69</td>
</tr>
<tr>
<td>Kraken</td>
<td>107.08</td>
<td>84.59</td>
<td>21.26</td>
<td>46.79</td>
<td>80.83</td>
<td>143.14</td>
<td>282.51</td>
<td>9.15</td>
<td>49.91</td>
</tr>
<tr>
<td>Liqui</td>
<td>81.99</td>
<td>48.51</td>
<td>22.02</td>
<td>47.69</td>
<td>73.41</td>
<td>105.93</td>
<td>170.37</td>
<td>8.27</td>
<td>33.40</td>
</tr>
<tr>
<td>Lykke</td>
<td>92.38</td>
<td>86.87</td>
<td>13.83</td>
<td>35.99</td>
<td>65.35</td>
<td>116.95</td>
<td>265.90</td>
<td>8.53</td>
<td>31.52</td>
</tr>
<tr>
<td>Poloniex</td>
<td>59.68</td>
<td>51.21</td>
<td>14.21</td>
<td>27.04</td>
<td>43.77</td>
<td>74.66</td>
<td>158.05</td>
<td>13.76</td>
<td>0.00</td>
</tr>
<tr>
<td>xBTCE</td>
<td>82.70</td>
<td>69.44</td>
<td>15.21</td>
<td>34.73</td>
<td>61.46</td>
<td>110.61</td>
<td>219.68</td>
<td>8.56</td>
<td>31.46</td>
</tr>
</tbody>
</table>
5. Quantifying Arbitrage Bounds

![Price Differences within Boundaries (in %)](chart)

- **Jan 2018**
- **Apr 2018**
- **Jul 2018**
- **Oct 2018**
- **Jan 2019**
- **Apr 2019**

**Legend:**
- **1**
- **2**
- **4**
Implied Relative Risk Aversion
## 5. Quantifying Arbitrage Bounds

### Price Differences (in%)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volatility</td>
<td>6.733***</td>
<td>6.749***</td>
<td>6.333***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(17.13)</td>
<td>(17.17)</td>
<td>(16.09)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latency (Median)</td>
<td>0.006***</td>
<td>0.006***</td>
<td>0.007***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5.21)</td>
<td>(5.23)</td>
<td>(5.24)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latency (SD)</td>
<td></td>
<td></td>
<td></td>
<td>0.001***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(5.72)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arbitrage Boundary</td>
<td></td>
<td></td>
<td></td>
<td>0.968***</td>
<td>1.354***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(26.80)</td>
<td>(27.74)</td>
<td></td>
</tr>
<tr>
<td>Boundary × Margin Trading</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.750***</td>
<td>(-10.51)</td>
</tr>
<tr>
<td>Margin Trading</td>
<td>1.585***</td>
<td>1.763***</td>
<td>1.587***</td>
<td>1.551***</td>
<td>1.432***</td>
<td>2.129***</td>
</tr>
<tr>
<td></td>
<td>(23.67)</td>
<td>(26.84)</td>
<td>(23.69)</td>
<td>(23.00)</td>
<td>(21.44)</td>
<td>(23.43)</td>
</tr>
<tr>
<td>Spread</td>
<td>0.281***</td>
<td>0.403***</td>
<td>0.282***</td>
<td>0.280***</td>
<td>0.239***</td>
<td>0.196***</td>
</tr>
<tr>
<td></td>
<td>(5.83)</td>
<td>(8.56)</td>
<td>(5.86)</td>
<td>(5.80)</td>
<td>(5.00)</td>
<td>(4.05)</td>
</tr>
<tr>
<td>Number of Confirmations</td>
<td>1.164***</td>
<td>1.133***</td>
<td>1.166***</td>
<td>1.189***</td>
<td>0.969***</td>
<td>0.889***</td>
</tr>
<tr>
<td></td>
<td>(53.05)</td>
<td>(52.22)</td>
<td>(53.09)</td>
<td>(53.04)</td>
<td>(44.24)</td>
<td>(40.28)</td>
</tr>
<tr>
<td>Tether</td>
<td>0.617***</td>
<td>0.483***</td>
<td>0.614***</td>
<td>0.670***</td>
<td>0.357***</td>
<td>0.201***</td>
</tr>
<tr>
<td></td>
<td>(8.21)</td>
<td>(6.44)</td>
<td>(8.17)</td>
<td>(8.80)</td>
<td>(4.85)</td>
<td>(2.69)</td>
</tr>
<tr>
<td>Business Account</td>
<td>-0.791***</td>
<td>-0.761***</td>
<td>-0.796***</td>
<td>-0.829***</td>
<td>-0.911***</td>
<td>-0.957***</td>
</tr>
<tr>
<td></td>
<td>(-9.18)</td>
<td>(-8.94)</td>
<td>(-9.25)</td>
<td>(-9.59)</td>
<td>(-10.63)</td>
<td>(-11.17)</td>
</tr>
<tr>
<td>Exchange Fixed Effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>.23</td>
<td>.22</td>
<td>.23</td>
<td>.23</td>
<td>.23</td>
<td>.23</td>
</tr>
</tbody>
</table>
6. Limits to Arbitrage and Cross-Exchange Flows
Flow dataset

- We identify 62 million wallets associated to 15 exchanges in our sample
- 3.7 million transactions with 54 million USD average daily volume
Exchange flows and arbitrage opportunities

- We estimate the simple linear model

\[ y_{i,t} = \alpha_i + \beta x_{i,t} + \varepsilon_{i,t}, \]

where \( y_{i,t} \) are (hourly) flows between exchanges and \( x_{i,t} \) are different measures of price differences.

<table>
<thead>
<tr>
<th>Exchange Flows (in USD)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price Differences</td>
<td>182.263</td>
<td>(0.00)</td>
<td></td>
</tr>
<tr>
<td>Price Differences Adjusted for Transaction Costs</td>
<td></td>
<td>84.945***</td>
<td>(6.65)</td>
</tr>
<tr>
<td>Price Differences Adjusted for Transaction Costs in Excess of Arbitrage Boundaries</td>
<td></td>
<td>111.010***</td>
<td>(4.57)</td>
</tr>
<tr>
<td>Exchange Pair Fixed Effect</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>39,246,152</td>
<td>39,246,152</td>
<td>38,806,690</td>
</tr>
</tbody>
</table>
Conclusions

Stochastic latency imposes limits to arbitrage

Key friction of blockchain-based settlement systems

Quantitatively important friction in Bitcoin markets
  ▶ Arbitrage bounds range around 90bp
  ▶ On average, 75% of all cross-market price differences are explained by stochastic latency solely.
  ▶ Additionally accounting for trading costs, 95% are captured.

Far reaching implications:
  ▶ Reduction of price efficiency; hindering price discovery
  ▶ Pricing of securities difficult; risk-neutral probabilities not unique
  ▶ Market makers have more room for quoting