Investigating the efficiency of financial stock markets with high frequency data

Dominik Rösch

University at Buffalo

drosch@buffalo.edu

October 10, 2017

This work is supported in part by NSF ACI-1541215.
Arbitrageurs enforce the law of one price. But how does arbitrage affect liquidity?
Arbitrageurs enforce the law of one price. But how does arbitrage affect liquidity?

If arbitrage opportunities arise as a result of demand shocks, then “arbitrageurs provide liquidity” [survey Gromb and Vayanos (2010)]
Arbitrageurs enforce the law of one price. But how does arbitrage affect liquidity?

If arbitrage opportunities arise as a result of demand shocks

then “arbitrageurs provide liquidity” [survey Gromb and Vayanos (2010)]

But if arbitrage opportunities arise as a result of differences in information

then “with arbitrage present, the adverse selection costs of domestic dealers increase, so that ... liquidity falls” [Domowitz, Glen, and Madhavan (1998)]
Arbitrageurs enforce the law of one price. But how does arbitrage affect liquidity?

If arbitrage opportunities arise as a result of demand shocks then “arbitrageurs provide liquidity” [survey Gromb and Vayanos (2010)]

But if arbitrage opportunities arise as a result of differences in information then “with arbitrage present, the adverse selection costs of domestic dealers increase, so that ... liquidity falls” [Domowitz, Glen, and Madhavan (1998)]

Impact of arbitrage on liquidity depends on reasons why arbitrage opportunities arise
The questions
The questions

- Part 1: How do arbitrage opportunities arise?
The questions

- Part 1: How do arbitrage opportunities arise?
- Part 2: How does arbitrage affect market liquidity?
The questions

- Part 1: How do arbitrage opportunities arise?
- Part 2: How does arbitrage affect market liquidity?
- For example, if arbitrage opportunities arise because of demand pressure, arbitrageurs might trade against net market order imbalance:

  [Chordia, Roll, Subrahmanyam (2002)]
  [O'Hara and Oldfield (1986) and Comerton-Forde, Hendershott, Jones, Moulton, Seasholes (2010)]
The questions

- Part 1: How do arbitrage opportunities arise?
- Part 2: How does arbitrage affect market liquidity?
- for example, if arbitrage opportunities arise because of demand pressure arbitrageurs might trade against net market order imbalance:
  - which would improve liquidity contemporaneous [Chordia, Roll, Subrahmanym (2002)]
The questions

- **Part 1:** How do arbitrage opportunities arise?
- **Part 2:** How does arbitrage affect market liquidity?
- for example, if arbitrage opportunities arise because of demand pressure arbitrageurs might trade against net market order imbalance:
  - which would improve liquidity contemporaneous [Chordia, Roll, Subrahmanyam (2002)]
  - and improve future liquidity [O’Hara and Oldfield (1986) and Comerton-Forde, Hendershott, Jones, Moulton, Seasholes (2010)]
Why care?

- Liquidity and efficiency are crucial for the well functioning of financial markets. Deviations from the law of one price are inefficiencies.
Why care?

- Liquidity and efficiency are crucial for the well functioning of financial markets. Deviations from the law of one price are inefficiencies.

- limits-of-arbitrage:
  
  liquidity decreases $\implies$ arbitrage activity decreases
Why care?

- Liquidity and efficiency are crucial for the well functioning of financial markets. Deviations from the law of one price are inefficiencies.

- limits-of-arbitrage:
  
  liquidity decreases $\implies$ arbitrage activity decreases

- here: arbitrage activity decreases $\implies$ liquidity decreases
Why care?

- Liquidity and efficiency are crucial for the well functioning of financial markets. Deviations from the law of one price are inefficiencies.

- limits-of-arbitrage:
  
  liquidity decreases $\implies$ arbitrage activity decreases

- here: arbitrage activity decreases $\implies$ liquidity decreases

for example, several frictions affect arbitrage activity

- Short sell constrains; Transaction tax (2018, in the EU);
  Margin requirements

Do these frictions only harm the efficiency of the market or also its liquidity?
The setting: Data (American Depositary Receipts)

- American Depositary Receipts (ADRs) which are cross-listed securities, and should trade at the same price as home market stock, because
  - give same cash flow as home market stock and can be converted to each other, which minimizes risk in arbitrage
- this makes ADR especially suitable to study arbitrage [Gagnon and Karolyi (2010)]
- Standard sources to create sample: Datastream, adrbnyn mellon.com and adr.db.com
- 5 different home markets, NYSE and Forex (72 stock pairs)
The setting: Data (American Depositary Receipts)

- American Depositary Receipts (ADRs) which are cross-listed securities, and should trade at the same price as home market stock, because
  - give same cash flow as home market stock and can be converted to each other, which minimizes risk in arbitrage
- this makes ADR especially suitable to study arbitrage [Gagnon and Karolyi (2010)]
- Standard sources to create sample: Datastream, adrbnymellon.com and adr.db.com
- 5 different home markets, NYSE and Forex (72 stock pairs)
- Tick-by-tick data from 1996 to 2013 on (almost 9 billion) quotes and (almost 1 billion) trades: TRTH
The setting: The clock (in UTC) 2008-10-15
The mechanics of arbitrage in the ADR market

- VOD.L
- VOD.N
Table of content

Introduction

Part 0: How to work with terabytes of financial data?

Part 1: How do price deviations arise?

Part 2: How does arbitrage affect market liquidity?
   Exogenous variation in impediments to arbitrage within the day
   Exogenous variation in impediments to arbitrage across days
   Impulse Response Functions

Transaction taxes
US tick-by-tick data (TAQ)
My first attempt: Mysql framework

- I started writing my stored procedure:
  - `get_ohlc_for(ticker, between, bucket)`

- when I coded:

  ```sql
  SUBSTRING_INDEX(
      MAX(CONCAT(time, ' ', price)), ' ', -1
  ) AS 'close'
  ```

- I realized MySQL might not be the best DB for financial data.
My second attempt: OneTick framework
Step 0: Setting up reference data

- Exchange trading times (e.g., LSE: 08:00-16:30 GMT)
- Price adjustments (e.g., VOD.L in pence)
- Price currency (e.g., VOD.L in GBP)
- Symbology mappings
- Corporate actions
Step 1: Convert prices into common currency
Step 2: Calculate price parity deviations
Market (in)efficiency: deviations from the fair price
Step 3 Run on server (using GNU Parallel)

database = ${1}
query = ${2}

for date in dates.txt ; do
    sem -k --id exp -P ot.cpus
        export.pl ${database} ${date} ${query}
    done

sem --wait --id exp
**Table of content**

**Introduction**

**Part 0: How to work with terabytes of financial data?**

**Part 1: How do price deviations arise?**

**Part 2: How does arbitrage affect market liquidity?**

- Exogenous variation in impediments to arbitrage within the day
- Exogenous variation in impediments to arbitrage across days
- Impulse Response Functions

**Transaction taxes**
Example Vodafone on 1999-06-15

![Graph showing VOD.L and VOD.N stock prices]

- **VOD.L**
- **VOD.N**

Time Range: 13:30 to 15:30
### How do price deviations arise? (part of Table 2)

Following Schultz and Shive (2010)

<table>
<thead>
<tr>
<th></th>
<th># [MM]</th>
<th>%price pressure</th>
<th>%Home</th>
<th>%Host</th>
<th>%Both</th>
<th>%Forex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Home</td>
<td>3.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Host</td>
<td>4.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Both</td>
<td>2.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forex</td>
<td>1.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
How do price deviations arise? (part of Table 2)

Following Schultz and Shive (2010)

<table>
<thead>
<tr>
<th></th>
<th># [MM]</th>
<th>%price pressure</th>
<th>%Home</th>
<th>%Host</th>
<th>%Both</th>
<th>%Forex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Home</td>
<td>3.74</td>
<td>0.70***</td>
<td>0.45</td>
<td>0.27</td>
<td>0.17</td>
<td>0.11</td>
</tr>
<tr>
<td>Host</td>
<td>4.64</td>
<td>0.78***</td>
<td>0.19</td>
<td>0.52</td>
<td>0.17</td>
<td>0.11</td>
</tr>
<tr>
<td>Both</td>
<td>2.29</td>
<td></td>
<td>0.24</td>
<td>0.32</td>
<td>0.36</td>
<td>0.09</td>
</tr>
<tr>
<td>Forex</td>
<td>1.95</td>
<td></td>
<td>0.20</td>
<td>0.26</td>
<td>0.12</td>
<td>0.43</td>
</tr>
</tbody>
</table>
Table of content

Introduction

Part 0: How to work with terabytes of financial data?

Part 1: How do price deviations arise?

Part 2: How does arbitrage affect market liquidity?
   Exogenous variation in impediments to arbitrage within the day
   Exogenous variation in impediments to arbitrage across days
   Impulse Response Functions

Transaction taxes
Two main challenges:

1) How to identify impediments to arbitrage?
Two main challenges:

1) How to identify impediments to arbitrage?

- start with Fama (1991): a market is efficient, if “prices reflect information to the point where the marginal benefits of acting ... do not exceed the marginal costs”

- Limits-of-arbitrage tells us what these costs are: e.g. risk, illiquidity, and capital constrains
Two main challenges:

1) How to identify impediments to arbitrage?

- start with Fama (1991): a market is efficient, if “prices reflect information to the point where the marginal benefits of acting ... do not exceed the marginal costs”

- Limits-of-arbitrage tells us what these costs are: e.g. risk, illiquidity, and capital constrains

- e.g., price deviations are “a symptom of a market in severe shortage of arbitrage capital” (Hu, Pan, Wang 2013)
Two main challenges:

1) How to identify impediments to arbitrage?

- start with Fama (1991): a market is efficient, if “prices reflect information to the point where the marginal benefits of acting ... do not exceed the marginal costs”

- Limits-of-arbitrage tells us what these costs are: e.g. risk, illiquidity, and capital constrains

- e.g., price deviations are “a symptom of a market in severe shortage of arbitrage capital” (Hu, Pan, Wang 2013)

- thus price deviations are a proxy for the impediments to arbitrage
The setting: Daily proxies

For impediments to arbitrage:

- $INARB_d$: seconds it takes for a price deviation to vanish
- $\Delta QTE_d$: difference in best bid and best ask price across the ADR and home market share
- $\Delta TRD_d$: difference in prices of simultaneous trades
The setting: Daily proxies

For impediments to arbitrage:

- $INARB_d$: seconds it takes for a price deviation to vanish
- $\Delta QTE_d$: difference in best bid and best ask price across the ADR and home market share
- $\Delta TRD_d$: difference in prices of simultaneous trades

For market quality:

- quoted spreads
The setting: Daily proxies

For impediments to arbitrage:

- $INARB_d$: seconds it takes for a price deviation to vanish
- $\Delta QTE_d$: difference in best bid and best ask price across the ADR and home market share
- $\Delta TRD_d$: difference in prices of simultaneous trades

For market quality:

- quoted spreads
- effective spreads (in paper)
Two main challenges:

2) How to address reverse causality between illiquidity and impediments to arbitrage?
Table of content

Introduction

Part 0: How to work with terabytes of financial data?

Part 1: How do price deviations arise?

Part 2: How does arbitrage affect market liquidity?
  
  Exogenous variation in impediments to arbitrage within the day
  
  Exogenous variation in impediments to arbitrage across days
  
  Impulse Response Functions

Transaction taxes
Omitted variables?

- Control directly for other important variables that explain illiquidity, e.g. volatility (in paper)
- Use a panel regression to control for time- and stock-invariant heterogeneity (in paper)
- Use a difference approach...
Omitted variables?

- Control directly for other important variables that explain illiquidity, e.g. volatility (in paper)
Omitted variables?

- Control directly for other important variables that explain illiquidity, e.g. volatility (in paper)

- Use a panel regression to control for time- and stock-invariant heterogeneity (in paper)
Omitted variables?

- Control directly for other important variables that explain illiquidity, e.g. volatility (in paper)
- Use a panel regression to control for time- and stock-invariant heterogeneity (in paper)
- Use a difference approach...
Difference in illiquidity during and outside overlapping trading times

The U.K.
Mexico
Germany
France
Brazil

07 08 09 10 11 12 13 14 15 16 17 18 19 20 21
Table of content

Introduction

Part 0: How to work with terabytes of financial data?

Part 1: How do price deviations arise?

Part 2: How does arbitrage affect market liquidity?

Exogenous variation in impediments to arbitrage within the day

Exogenous variation in impediments to arbitrage across days

Impulse Response Functions

Transaction taxes
The mechanics of arbitrage in the ADR market
Summary statistics (part of Table 1)

<table>
<thead>
<tr>
<th></th>
<th>avg</th>
<th>stddev</th>
<th>min</th>
<th>median</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{avg}(\Delta TRD)$</td>
<td>0.45</td>
<td>0.33</td>
<td>0.06</td>
<td>0.40</td>
<td>2.41</td>
</tr>
<tr>
<td>$\text{avg}(\Delta QTE)$</td>
<td>0.25</td>
<td>0.20</td>
<td>0.00</td>
<td>0.23</td>
<td>0.88</td>
</tr>
<tr>
<td>$\text{max}(\Delta QTE)$</td>
<td>0.73</td>
<td>0.46</td>
<td>0.14</td>
<td>0.66</td>
<td>3.07</td>
</tr>
<tr>
<td>duration [sec]</td>
<td>414</td>
<td>429</td>
<td>4</td>
<td>315</td>
<td>2,479</td>
</tr>
</tbody>
</table>
Summary statistics (part of Table 1)

<table>
<thead>
<tr>
<th></th>
<th>avg</th>
<th>stddev</th>
<th>min</th>
<th>median</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Price deviations outside days between corporate actions:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>avg($\Delta TRD$)</td>
<td>0.45</td>
<td>0.33</td>
<td>0.06</td>
<td>0.40</td>
<td>2.41</td>
</tr>
<tr>
<td>avg($\Delta QTE$)</td>
<td>0.25</td>
<td>0.20</td>
<td>0.00</td>
<td>0.23</td>
<td>0.88</td>
</tr>
<tr>
<td>max($\Delta QTE$)</td>
<td>0.73</td>
<td>0.46</td>
<td>0.14</td>
<td>0.66</td>
<td>3.07</td>
</tr>
<tr>
<td>duration [sec]</td>
<td>414</td>
<td>429</td>
<td>4</td>
<td>315</td>
<td>2,479</td>
</tr>
<tr>
<td>Panel B: Price deviations during days between corporate actions:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># days</td>
<td>18</td>
<td>21</td>
<td>0</td>
<td>12</td>
<td>102</td>
</tr>
<tr>
<td>avg($\Delta QTE$)</td>
<td>1.46</td>
<td>2.57</td>
<td>0.07</td>
<td>0.85</td>
<td>13.10</td>
</tr>
<tr>
<td>avg($\Delta AQTE$)</td>
<td>0.62</td>
<td>0.84</td>
<td>0.00</td>
<td>0.30</td>
<td>4.82</td>
</tr>
<tr>
<td>max($\Delta QTE$)</td>
<td>2.48</td>
<td>2.55</td>
<td>0.25</td>
<td>2.00</td>
<td>13.80</td>
</tr>
<tr>
<td>max($\Delta AQTE$)</td>
<td>1.27</td>
<td>0.98</td>
<td>0.14</td>
<td>0.89</td>
<td>5.17</td>
</tr>
</tbody>
</table>
How does arbitrage affect liquidity: Panel instrumental variable regression.

(part of Table 6)

\[
PQSPR_{i,d} = \alpha + \beta \times \Delta Price_{i,d} + \beta \times Controls_{i,d} + \epsilon_{i,d}
\]

<table>
<thead>
<tr>
<th></th>
<th>INARB</th>
<th>avg((\Delta QTE_d))</th>
<th>max((\Delta QTE_d))</th>
<th>max((\Delta AQTE_d))</th>
<th>avg((\Delta TRD_d))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{INARB})</td>
<td>0.01**</td>
<td>0.02**</td>
<td>0.01**</td>
<td>0.04*</td>
<td>0.02**</td>
</tr>
<tr>
<td>(0.03)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.09)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>avg((\Delta QTE_d))</td>
<td>0.02**</td>
<td></td>
<td></td>
<td></td>
<td>0.02**</td>
</tr>
<tr>
<td>(0.06)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.05)</td>
</tr>
<tr>
<td>max((\Delta QTE_d))</td>
<td></td>
<td></td>
<td>0.01**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.05)</td>
<td></td>
<td></td>
<td>(0.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>max((\Delta AQTE_d))</td>
<td></td>
<td>0.01*</td>
<td></td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>(0.09)</td>
<td></td>
<td>(0.02)</td>
<td></td>
<td>(0.11)</td>
<td></td>
</tr>
<tr>
<td>avg((\Delta TRD_d))</td>
<td></td>
<td>0.02**</td>
<td></td>
<td></td>
<td>0.01**</td>
</tr>
<tr>
<td>(0.03)</td>
<td></td>
<td>(0.03)</td>
<td></td>
<td></td>
<td>(0.05)</td>
</tr>
</tbody>
</table>

Controls: Yes
StockFE: Yes
DayFE: Yes
StockDays: 136,363
How does arbitrage affect liquidity: Panel instrumental variable regression.

(part of Table 6)

\[
\delta P_{QSPR_{i,d}} = \alpha + \beta \times \Delta \text{Price}_{i,d} + \beta \times \text{Controls}_{i,d} + \epsilon_{i,d}
\]

Host_{i,d} \quad Host_{i,d} \quad Host_{i,d} \quad Host_{i,d} \quad Host_{i,d} \quad Home_{i,d} \quad Home_{i,d} \quad Home_{i,d} \quad Home_{i,d} \quad Home_{i,d}

\begin{align*}
\text{INARB}_{i,d} & \quad 0.008*** \\
& \quad (0.00) \quad 0.010* \\
\text{avg}(\Delta QTE_{i,d}) & \quad 0.018*** \\
& \quad (0.00) \quad 0.023* \\
\text{max}(\Delta QTE_{i,d}) & \quad 0.013*** \\
& \quad (0.00) \quad 0.017* \\
\text{max}(\Delta AQTE_{i,d}) & \quad 0.060*** \\
& \quad (0.00) \quad 0.075* \\
\text{avg}(\Delta TRD_{i,d}) & \quad 0.017*** \\
& \quad (0.00) \quad 0.022* \\
\text{Controls} & \quad Yes \\
\text{StockFE} & \quad Yes \\
\text{DayFE} & \quad Yes \\
\text{StockDays} & \quad 136,363
\end{align*}
Table of content

Introduction

Part 0: How to work with terabytes of financial data?

Part 1: How do price deviations arise?

Part 2: How does arbitrage affect market liquidity?
  Exogenous variation in impediments to arbitrage within the day
  Exogenous variation in impediments to arbitrage across days
  Impulse Response Functions

Transaction taxes
How does arbitrage affect liquidity: Impulse response functions (Figure 1)
How does arbitrage affect liquidity: Impulse response functions (Figure 3)
The impact of transaction taxes on impediments to arbitrage. (Table 8)

\[ LHS_{i,d} = FE + \beta_0 \times FFTT_{i,d} + \beta_1 \times AfterFFTT_{home,i,d} + \beta_2 \times AfterFFTT_{host,i,d} \]

<table>
<thead>
<tr>
<th></th>
<th>INARB(_{i,d})</th>
<th>(a(\Delta Q_{i,d}))</th>
<th>(m(\Delta Q_{i,d}))</th>
<th>(m(\Delta AQ_{i,d}))</th>
<th>(a(\Delta T_{i,d}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(FFTT_{i,d})</td>
<td>-0.726</td>
<td>-0.069</td>
<td>0.274</td>
<td>-0.061</td>
<td>0.083</td>
</tr>
<tr>
<td></td>
<td>(0.48)</td>
<td>(0.33)</td>
<td>(0.30)</td>
<td>(0.39)</td>
<td>(0.53)</td>
</tr>
<tr>
<td>(AfterFFTT_{Home,i,d})</td>
<td>1.299***</td>
<td>0.089***</td>
<td>0.128***</td>
<td>0.091***</td>
<td>0.125***</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>(AfterFFTT_{Host,i,d})</td>
<td>0.415</td>
<td>0.054</td>
<td>0.031</td>
<td>0.007</td>
<td>0.049</td>
</tr>
<tr>
<td></td>
<td>(0.49)</td>
<td>(0.48)</td>
<td>(0.74)</td>
<td>(0.90)</td>
<td>(0.55)</td>
</tr>
<tr>
<td>StockFE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>DayFE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>StockDays</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>17,358</td>
</tr>
</tbody>
</table>
The impact of transaction taxes on illiquidity. (Table 8)

\[ LHS_{i,d} = FE + \beta_0 \times FFTT_{i,d} + \beta_1 \times AfterFFTT_{home,i,d} + \beta_2 \times AfterFFTT_{host,i,d} + \epsilon_{i,d} \]

<table>
<thead>
<tr>
<th></th>
<th>Panel B: PQSPR</th>
<th>Panel C: PESPR</th>
<th>Panel D: ( \delta PQSPR )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Host_{i,d}</td>
<td>Home_{i,d}</td>
<td>Host_{i,d}</td>
</tr>
<tr>
<td>( FFTT_{i,d} )</td>
<td>0.066</td>
<td>0.012</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>(0.27)</td>
<td>(0.12)</td>
<td>(0.39)</td>
</tr>
<tr>
<td>( AfterFFTT_{Home,i,d} )</td>
<td>0.046***</td>
<td>0.005**</td>
<td>0.037*</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.04)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>( AfterFFTT_{Host,i,d} )</td>
<td>0.013</td>
<td>0.001</td>
<td>-0.021</td>
</tr>
<tr>
<td></td>
<td>(0.18)</td>
<td>(0.62)</td>
<td>(0.27)</td>
</tr>
</tbody>
</table>

StockFE: Yes
DayFE: Yes
StockDays: 136,363
Summary

- Arbitrage opportunities mainly arise due to demand pressure
- An increase in the impediments to arbitrage deteriorates liquidity
  - contemporaneously
  - and over the coming days
- In particular, transaction taxes lower liquidity and thereby increase the cost of capital for firms