

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.

“The impact of arbitrage on market liquidity”, WP 2017

- ▶ Arbitrageurs enforce the law of one price. But how does arbitrage affect liquidity?

“The impact of arbitrage on market liquidity”, WP 2017

- ▶ 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)]

“The impact of arbitrage on market liquidity”, WP 2017

- ▶ 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)]

“The impact of arbitrage on market liquidity”, WP 2017

- ▶ 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:

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)]

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?

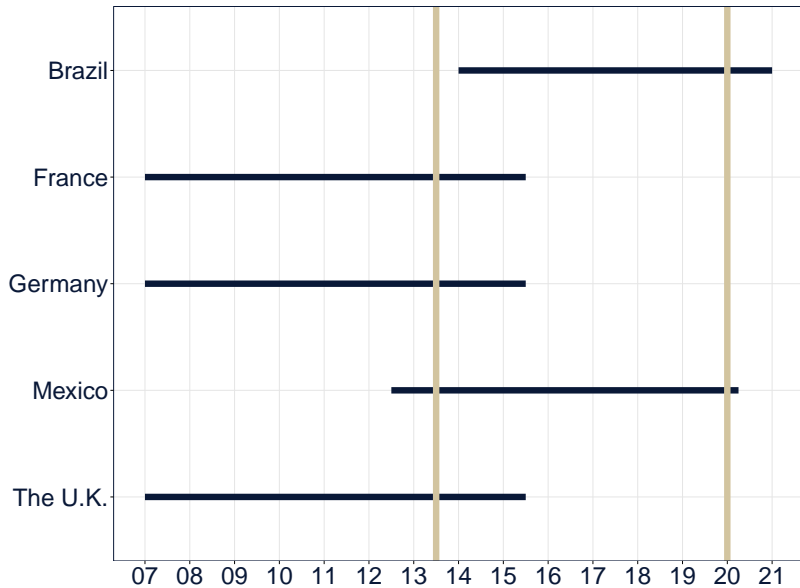
The setting: Data (American Depositary Receipts)

- ▶ 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, adrbnymellon.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

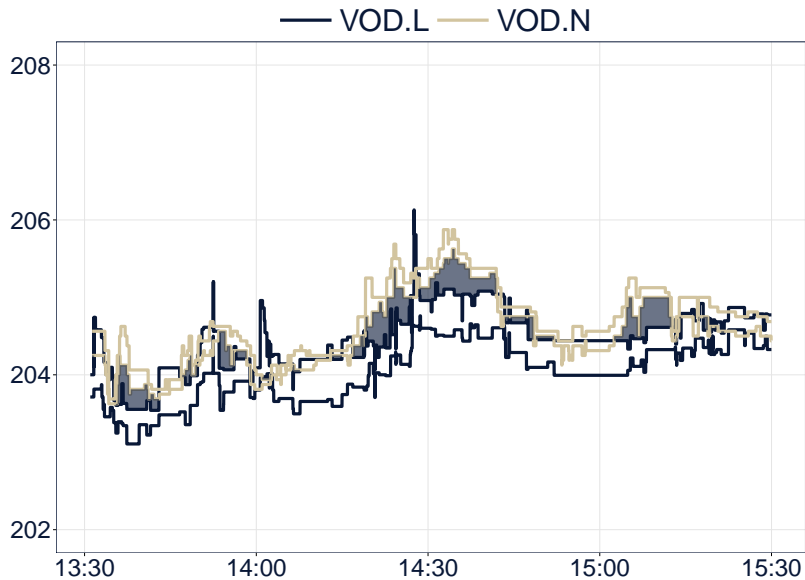


Table of content

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?

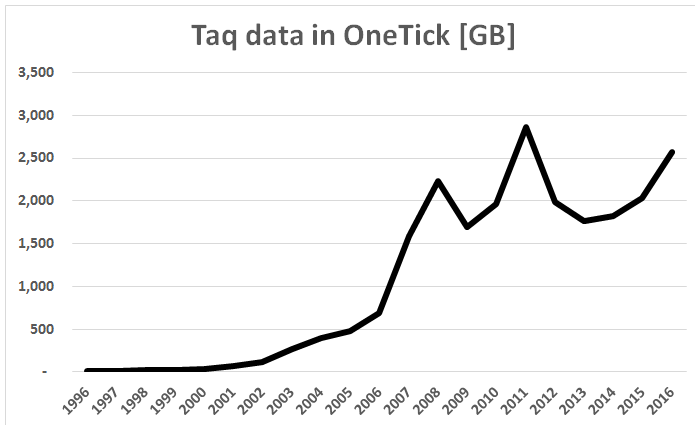
Exogenous variation in impediments to arbitrage within the day

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:

```
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

The screenshot shows the OneTick Display application interface. The main window displays a table of market data for TAQ-IBM. Below the table are controls for date and time filtering, and a table of active queries.

Index	Symbol	Time	PRICE	SIZE	TRD_EX	COND	CORR	GL27	OMDSEQ	BuySellFlag
1	TAQ-IBM	2010/01/20 09:30:00.000	130.4600000	100	Z	F	0	0	0	-1.0000000
2	TAQ-IBM	2010/01/20 09:30:00.000	130.4400000	100	P	Q	0	0	1	1.0000000
3	TAQ-IBM	2010/01/20 09:30:00.000	130.4400000	100	P	F	0	0	2	1.0000000
4	TAQ-IBM	2010/01/20 09:30:00.000	130.4400000	100	P	F	0	0	3	1.0000000
5	TAQ-IBM	2010/01/20 09:30:01.000	130.4700000							
6	TAQ-IBM	2010/01/20 09:30:01.000	130.4600000							
7	TAQ-IBM	2010/01/20 09:30:01.000	130.4700000							
8	TAQ-IBM	2010/01/20 09:30:01.000	130.4800000							
9	TAQ-IBM	2010/01/20 09:30:01.000	130.4900000							
10	TAQ-IBM	2010/01/20 09:30:01.000	130.4800000							
11	TAQ-IBM	2010/01/20 09:30:01.000	130.5000000							
12	TAQ-IBM	2010/01/20 09:30:01.000	130.5200000							
13	TAQ-IBM	2010/01/20 09:30:01.000	130.5200000							
14	TAQ-IBM	2010/01/20 09:30:01.000	130.5300000							
15	TAQ-IBM	2010/01/20 09:30:02.000	130.4800000							
16	TAQ-IBM	2010/01/20 09:30:02.000	130.4800000							
17	TAQ-IBM	2010/01/20 09:30:03.000	130.4600000							
18	TAQ-IBM	2010/01/20 09:30:03.000	130.5300000							
19	TAQ-IBM	2010/01/20 09:30:03.000	130.5300000							
20	TAQ-IBM	2010/01/20 09:30:03.000	130.5300000							
21	TAQ-IBM	2010/01/20 09:30:03.000	130.5000000							
22	TAQ-IBM	2010/01/20 09:30:03.000	130.4900000							
23	TAQ-IBM	2010/01/20 09:30:03.000	130.5300000							

Below the table, the 'Date and time' section shows: Start: 01/20/2010 09:30:00.000, End: 01/20/2010 10:00:00.000, Timezone: New York. The 'Apply times daily' checkbox is checked.

The 'Query' table at the bottom shows:

Name	Security	Query
sign trades	TAQ-IBM	LEE_AND_READY(0)

The 'Graph Editor' window is open, showing a query named 'sign_trades' for security 'ONE MARKET DATA'. It displays a graph with three nodes: 'QTE PASSTHROUGH QTE', 'TRD PASSTHROUGH TRD', and 'LEE_AND_READY'. Arrows point from the top two nodes to the bottom node.

OneTick "code"

The screenshot shows a GitHub repository page for the file `one_tick/queries/report_new_el`. The page displays a commit history table on the left and the corresponding code on the right.

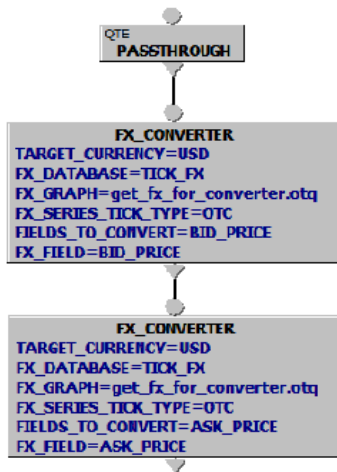
Commit Message	Time Ago	Line
Initial commit	4 years ago	1
rewrote get_exchange_efficiency to han...	3 months ago	2
adjust for new data till 2013	3 years ago	3
rewrote get_exchange_efficiency to han...	3 months ago	4
adjust for new data till 2013	3 years ago	5
minor changes	2 years ago	6
adjust for new data till 2013	3 years ago	7
rewrote get_exchange_efficiency to han...	3 months ago	8
adjust for new data till 2013	3 years ago	9
		10
		11
		12
		13
		14
		15
		16
		17
		18
		19
		20
		21
		22
		23
		24
		25

```
[get_exchange_efficiency]
COMMENT = RUN in New York time zone. Otherwise cannot join by time, p
CPU_NUMBER = 1
DB_HIHT_FOR_PROCESSING_HOST =
graph_reuse = 0
NODE_10 = COMPUTE(COMPUTE="HIGH(INPUT_FIELD_NAME=HIGH.PROFIT_PCT,OUTPUT
AVERAGE(INPUT_FIELD_NAME=VWAP.PROFIT_PCT,OUTPUT_FIELD_NAME=PROFIT_PCT)
HIGH(INPUT_FIELD_NAME=HIGH.PROFIT,OUTPUT_FIELD_NAME=PROFIT)",BUCKET_IN
NODE_10_SOURCE = NODE_24..IF
NODE_10_X = 1316
NODE_10_Y = 476
NODE_11 = NESTED_OTQ add_field.otq::add_days_between_corp_act
NODE_11_MINIMIZED = 1
NODE_11_PARAMETER = MAX_DAYS_BETWEEN_CORPS 21
NODE_11_SOURCE = NODE_6.NODE_13.
NODE_11_SOURCE_DESCRIPTION = NODE_6.NODE_13. IN.
NODE_11_X = 1822
NODE_11_Y = 1216
NODE_14 = MERGE
NODE_14_BIND_SECURITY = eval("get_symbols.otq::get_symbols","DB="TICK_
NODE_14_BIND_SECURITY = DAILY_AGGREGATED_INPUT_EX::VOD.L 19950101 No
NODE_14_SOURCE = NODE_32
NODE_14_X = 1856
NODE_14_Y = 2128
```

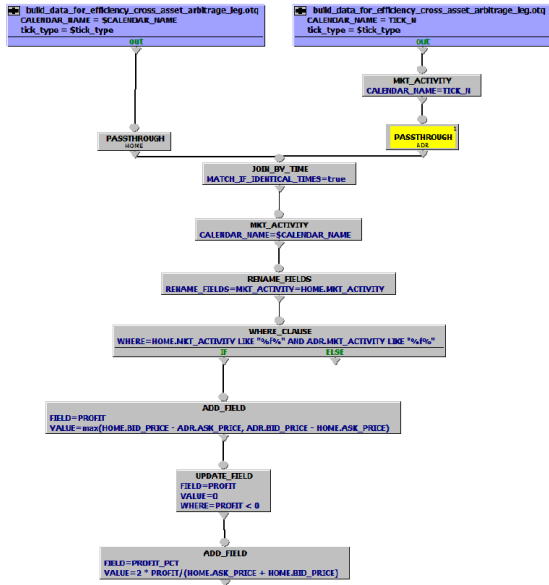

Step 0: Setting up reference data

- ▶ 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

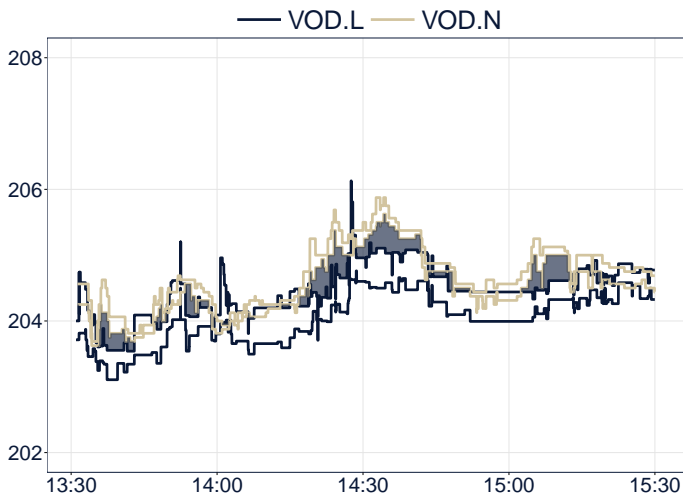
Step 1: Convert prices into common currency



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?

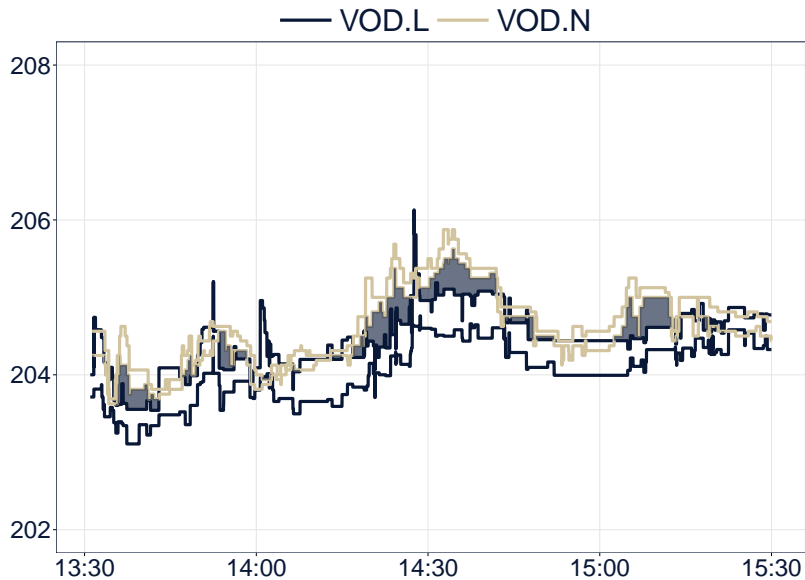
Exogenous variation in impediments to arbitrage within the day

Exogenous variation in impediments to arbitrage across days

Impulse Response Functions

Transaction taxes

Example Vodafone on 1999-06-15



How do price deviations arise? (part of Table 2)

Following Schultz and Shive (2010)

	# [MM]	%price pressure	%Home	%Host	%Both	%Forex
<i>Home</i>	3.3					
<i>Host</i>	4.2					
<i>Both</i>	2.1					
<i>Forex</i>	1.8					

How do price deviations arise? (part of Table 2)

Following Schultz and Shive (2010)

	# [MM]	%price pressure	% <i>Home</i>	% <i>Host</i>	% <i>Both</i>	% <i>Forex</i>
<i>Home</i>	3.74	0.70***	0.45	0.27	0.17	0.11
<i>Host</i>	4.64	0.78***	0.19	0.52	0.17	0.11
<i>Both</i>	2.29		0.24	0.32	0.36	0.09
<i>Forex</i>	1.95		0.20	0.26	0.12	0.43

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 constraints

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 constraints
- ▶ 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 constraints
- ▶ 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
- ▶ ΔQTE_d : difference in best bid and best ask price across the ADR and home market share
- ▶ ΔTRD_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
- ▶ ΔQTE_d : difference in best bid and best ask price across the ADR and home market share
- ▶ ΔTRD_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
- ▶ ΔQTE_d : difference in best bid and best ask price across the ADR and home market share
- ▶ ΔTRD_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?

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

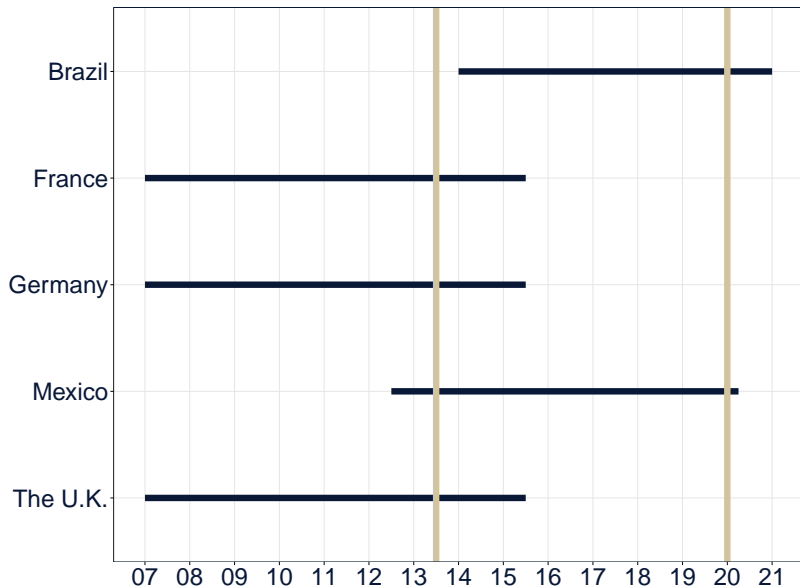


Table of content

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?

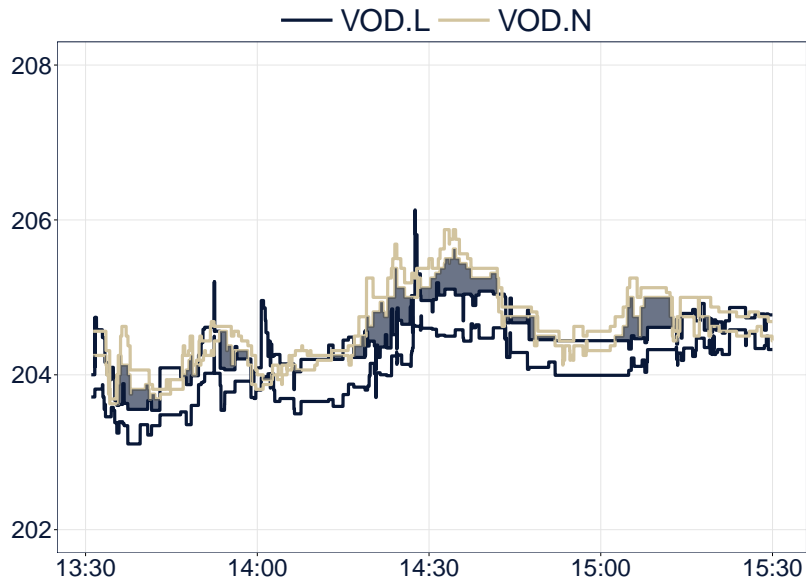
Exogenous variation in impediments to arbitrage within the day

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)

	avg	stddev	min	median	max
$avg(\Delta TRD)$	0.45	0.33	0.06	0.40	2.41
$avg(\Delta QTE)$	0.25	0.20	0.00	0.23	0.88
$max(\Delta QTE)$	0.73	0.46	0.14	0.66	3.07
duration [sec]	414	429	4	315	2,479

Summary statistics (part of Table 1)

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

How does arbitrage affect liquidity: Panel instrumental variable regression.

(part of Table 6)

$$PQSPR_{i,d} = \alpha + \beta \times \widehat{\Delta Price}_{i,d} + \beta \times \mathbf{Controls}_{i,d} + \epsilon_{i,d}$$

*Host*_{*i,d*} *Host*_{*i,d*} *Host*_{*i,d*} *Host*_{*i,d*} *Host*_{*i,d*} *Home*_{*i,d*} *Home*_{*i,d*} *Home*_{*i,d*} *Home*_{*i,d*} *Home*_{*i,d*}

\widehat{INARB}_d	0.01**		0.01*	
(0.03)			(0.06)	
$\overline{avg(\Delta QTE_d)}$	0.02**		0.02**	
(0.02)			(0.05)	
$\overline{max(\Delta QTE_d)}$	0.01**		0.01*	
	(0.02)		(0.05)	
$\overline{max(\Delta AQTE_d)}$		0.04*		0.03
		(0.09)		(0.11)
$\overline{avg(\Delta TRD_d)}$			0.02**	0.01**
			(0.03)	(0.05)
Controls			Yes	
StockFE			Yes	
DayFE			Yes	
StockDays			136,363	

How does arbitrage affect liquidity: Panel instrumental variable regression.
(part of Table 6)

$$\delta PQSPR_{i,d} = \alpha + \beta \times \widehat{\Delta Price}_{i,d} + \beta \times \mathbf{Controls}_{i,d} + \epsilon_{i,d}$$

*Host*_{*i,d*} *Host*_{*i,d*} *Host*_{*i,d*} *Host*_{*i,d*} *Host*_{*i,d*} *Home*_{*i,d*} *Home*_{*i,d*} *Home*_{*i,d*} *Home*_{*i,d*} *Home*_{*i,d*}

$\widehat{INARBQ}_{i,d}$	0.008***		0.010*	
	(0.00)		(0.07)	
$\widehat{avg}(\Delta QTE_{i,d})$	0.018***		0.023*	
	(0.00)		(0.09)	
$\widehat{max}(\Delta QTE_{i,d})$	0.013***		0.017*	
	(0.00)		(0.07)	
$\widehat{max}(\Delta AQTE_{i,d})$		0.060***		0.075*
		(0.00)		(0.07)
$\widehat{avg}(\Delta TRD_{i,d})$		0.017***		0.022*
		(0.00)		(0.09)
Controls		Yes		
StockFE		Yes		
DayFE		Yes		
StockDays		136,363		

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?

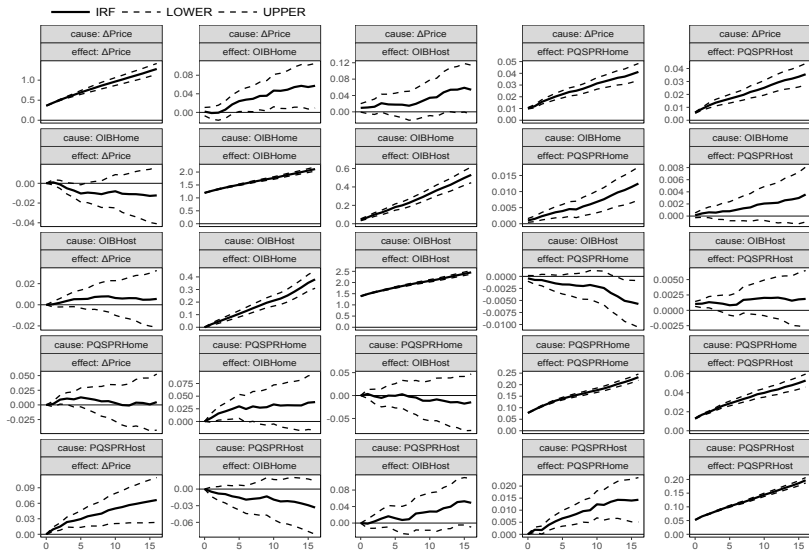
Exogenous variation in impediments to arbitrage within the day

Exogenous variation in impediments to arbitrage across days

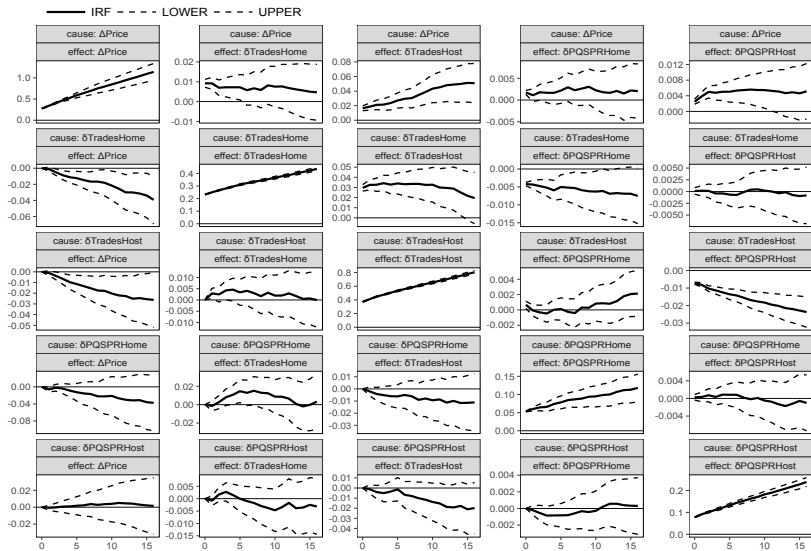
Impulse Response Functions

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)

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

	$INARB_{i,d}$	$a(\Delta Q_{i,d})$	$m(\Delta Q_{i,d})$	$m(\Delta AQ_{i,d})$	$a(\Delta T_{i,d})$
$FFTT_{i,d}$	-0.726 (0.48)	-0.069 (0.33)	0.274 (0.30)	-0.061 (0.39)	0.083 (0.53)
$AfterFFTT_{Home,i,d}$	1.299*** (0.00)	0.089*** (0.00)	0.128*** (0.00)	0.091*** (0.00)	0.125*** (0.00)
$AfterFFTT_{Host,i,d}$	0.415 (0.49)	0.054 (0.48)	0.031 (0.74)	0.007 (0.90)	0.049 (0.55)
StockFE			Yes		
DayFE			Yes		
StockDays			17,358		

The impact of transaction taxes on illiquidity. (Table 8)

$$LHS_{i,d} = FE + \beta_0 \times FTT_{i,d} + \beta_1 \times AfterFTT_{home,i,d} + \beta_2 \times AfterFTT_{host,i,d}$$

	Panel B: <i>PQSPR</i>		Panel C: <i>PESPR</i>		Panel D: $\delta PQSPR$	
	<i>Host</i> _{<i>i,d</i>}	<i>Home</i> _{<i>i,d</i>}	<i>Host</i> _{<i>i,d</i>}	<i>Home</i> _{<i>i,d</i>}	<i>Host</i> _{<i>i,d</i>}	<i>Home</i> _{<i>i,d</i>}
<i>FTT</i> _{<i>i,d</i>}	0.066 (0.27)	0.012 (0.12)	0.023 (0.39)	0.011 (0.26)	0.044** (0.04)	0.001 (0.84)
<i>AfterFTT</i> _{<i>Home</i>_{<i>i,d</i>}}	0.046*** (0.00)	0.005** (0.04)	0.037* (0.07)	0.005* (0.09)	-0.011 (0.39)	0.000 (0.87)
<i>AfterFTT</i> _{<i>Host</i>_{<i>i,d</i>}}	-0.013 (0.18)	0.001 (0.62)	-0.021 (0.27)	-0.001 (0.63)	-0.007 (0.16)	0.002 (0.11)
<i>StockFE</i>				Yes		
<i>DayFE</i>				Yes		
<i>StockDays</i>				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