THE RISKS OF HIGH-FREQUENCY TRADING

Irene Aldridge

October 24, 2013

5th Annual Conference in HF Modeling in Finance
Present situation

Investors concerned about HFT

HFT has been fingered in
- Market volatility
- Market crashes = Black swans (fat tails of returns)
- Market liquidity
- Market fairness
  - The key activity that has been identified by regulators as market-manipulating is spoofing

Huge market fragmentation
- About 18 exchanges just in equities, multiple dark pools, etc.
- Fragmentation driven by the need of differentiation

Question of this study: are some exchanges HFT-riskier?
What about other issues?

• Flash Crash, 2010
• BATS Crash, 2012
• Knight Crash, 2012
• Twitter Hack Attack, 2013

Volatility
Crashes
Liquidity
Fairness
Present U.S. Equity Exchanges

- Fragmentation driven by the need of differentiation

<table>
<thead>
<tr>
<th>All Present U.S. Equity Exchanges</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Price-time priority</strong></td>
</tr>
<tr>
<td>Normal</td>
</tr>
<tr>
<td>- pay for market orders,</td>
</tr>
<tr>
<td>rebates for limit orders</td>
</tr>
<tr>
<td>Inverted</td>
</tr>
<tr>
<td>- pay for limit orders,</td>
</tr>
<tr>
<td>rebates for market orders</td>
</tr>
<tr>
<td><strong>Price-size priority (pro-rata)</strong></td>
</tr>
</tbody>
</table>

© Copyright Irene Aldridge 2013
All Present Exchange Models

All Present Exchanges

Price-time priority
- Normal
  - pay for market orders, rebates for limit orders
- Inverted
  - pay for limit orders, rebates for market orders

Price-size priority (pro-rata)

No-cancel Range

© Copyright Irene Aldridge 2013
Key Question

• Which market structure is best?
ialdridge@AbleAlpha.com

- Active HFT research since 2007 (market making, other strategies): AbleAlpha.com
- Sell real-time feeds to billion-dollar hedge funds (patents):
  - Identifying % of aggressive HFTs in a given market right now
  - Identifying probability of a crash within the next few hours
  - Accurately predicting market impact should you execute right now
- Also sell: real-time trade sentiment feeds (trade secret)
- HFT Training: HFTTraining.com Optimal execution course in NYC coming up in Nov.
- Big Data Finance – sponsorship

- New Book: HFT Second Edition
  ISBN: 978-1118343500

© Copyright Irene Aldridge 2013
### Key Conclusions

- Different order matching principles result in theoretically and empirically different market characteristics:

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Normal</th>
<th>Inverted</th>
<th>Pro-rata</th>
<th>No-cancel range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order cancellations at top of the book</td>
<td>High</td>
<td>Medium</td>
<td>Low</td>
<td>Zero</td>
</tr>
<tr>
<td>Top-of-the-book liquidity “quality”</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
<td>Highest</td>
</tr>
<tr>
<td>Ability to detect spoofing market manipulation</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
<td>Highest</td>
</tr>
</tbody>
</table>

© Copyright Irene Aldridge 2013
Key Conclusions (cont.)

• Different order matching principles result in theoretically and empirically different market characteristics:

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Vanilla Price-Time Priority</th>
<th>No-Cancel-Range Either Priority</th>
<th>Vanila Pro-Rata Matching</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market volatility</td>
<td>High</td>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>Fat tails of returns</td>
<td>High</td>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>Market liquidity</td>
<td>High</td>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>Ability to detect spoofing market manipulation</td>
<td>Low</td>
<td>High</td>
<td>High</td>
</tr>
</tbody>
</table>
Outline

1. Modern market microstructure
2. Analytical results
3. Empirical results
4. The problem of market manipulation
5. Conclusions
1. What is market microstructure?

- It is the study of how price formation happens
- Also known as price tattonnement (bargaining) in economics
- The reason for “micro:” this is the study of what happens under the microscope:
  - How prices evolve in the very short term
  - In “micro”, we do not particularly care about “macro” fundamentals that drive the markets in the long term
1. Why do we care about microstructure?

- We want to know how prices evolve in the short term so we can:
  - Estimate where the prices are going (up or down)
  - Evaluate whether the volatility is increasing
  - Estimate the probability of important news hitting the market and the resulting implication, if any
  - Detect market manipulation
  - …
1. A limit order book, snapshot
1. When orders arrive...

- A large market sell order sweeps the bids in the book, realizes lower aggregate price, changes the best bid.
- A new limit buy order arrives at a price lower than the best bid but higher than the best ask.
- A new limit buy order arrives at a price lower than the best bid.
- A new limit buy order arrives at a price higher than the best bid but lower than the best ask.
1. LOB with Price-Time Priority

A market sell order arrives

The “oldest” limit buy order placed at the best bid earliest is executed first; if the market sell is not filled completely with the oldest limit order, the market sell is next matched with the second oldest limit buy order placed at the best bid. The process continues until the market sell order is matched in full or the best bid queue is exhausted.
An incoming market sell order is matched with a constant proportion of every limit buy order placed at the best bid price, independently of the arrival times of the orders.
Outline

1. Modern market microstructure
2. Analytical results
3. Empirical results
4. The problem of market manipulation
5. Conclusions
2. Modeling Limit Order Books

© Copyright Irene Aldridge 2013
2. LOB = Queuing System!

Queuing network, bid side

Queuing network, offer side
2. LOB Queues

- Order cancellations are disallowed:

- Order cancellations are allowed (vanilla matching):

- Top-of-the-book cancellations not allowed, but okay elsewhere
2. Related literature

- Cont, Stoikov and Talreja (OR, 2010)
- Cont and Kukanov (2012)
- Hasbrouck (2012, 2013)
- Easley, Lopez de Prado, and O’Hara (2011, 2012)
- Whitt (2000 – present)
- Many others
2. Modeling queues

Simplest case: M/M/1
- Simple memoryless queues (1 bid queue and 1 ask queue)
- Limit order arrivals with rates $\lambda_1, \lambda_2, \ldots \lambda_k$
- Market order “arrivals” with rate $\mu$
- Limit order cancellations with rates $\delta_1, \delta_2, \ldots \delta_k$

Queue disciplines:
- Price-time priority = FIFO (or First Come First Served)
- Pro-rata matching = Processor Sharing (PS)
- Order cancellations are reneging customers

Can analytically estimate:
- Steady-state number of orders in each queue
- Steady-state time until execution
2. Modeling results 1

Proposition 1 (Aldridge 2013):
Steady-state probability of observing exactly \( z \) orders in the \( k \)th queue when order cancellations are allowed (\( \delta_k \neq 0 \)):

\[
(I_{\infty,k} = z) = \left( 1 + \sum_{z=1}^{\infty} \frac{\lambda_1^z}{\prod_{j=1}^{z}(\mu + \delta_1)} \right)^{-1} \left( 1 \right)
\]

\[
+ \sum_{z=1}^{\infty} \frac{\lambda_2^z}{\prod_{j=1}^{z}(\mu + \delta_2)} \left( 1 \right)
\]

\[
+ \sum_{z=1}^{\infty} \frac{\lambda_k^z}{\prod_{j=1}^{z}(\mu + \delta_k)} \frac{\lambda_k^z}{\prod_{j=1}^{z}(\mu + \delta_k)}
\]

© Copyright Irene Aldridge 2013
2. Modeling results 2

• Proposition 2: (Aldridge, 2013)

Wait time to execution when cancellations are allowed

\[ P(W_{\infty,k} \leq t) = \left[ \prod_{i=1}^{k}(1 - (\mu + \delta_i - \lambda_i) \exp(-(\mu + \delta_i - \lambda_i)t)) \right] \times \exp \left( 1 - \sum_{i=1}^{k-1} \lambda_i t \right) \]
2. Sketch of a Proof

- See Aldridge (2013) for detailed proof (SSRN)
- Any priority (FIFO or PS – same steady-state probabilities discussed here), no cancellations
- Assumptions: steady-state, exponentially-distributed limit and market order arrivals, where limit order cancellations are not allowed (order cancellation rate $\delta = 0$)

\[ P(I_{\infty,k} = z) = \left( (1 - \rho_1)(1 - \rho_2) \ldots (1 - \rho_k) \right) \rho_k^z \]

where $\rho_k \equiv \frac{\lambda_k}{\mu}$ is traffic intensity in the kth queue, and $\lambda_k$ is the arrival rate of limit orders into queue $k$. 
2. Time to execution in queue k

Assume no cancellations:

- Steady-state, under no cancellations rule, probability of executing a limit order placed \( k \) ticks away from the current market price within time \( t \) equals

\[
P(W_{\infty,k} \leq t | \delta = 0) = 
\left[ \prod_{i=1}^{k} (1 - (\mu - \lambda_i) \exp(-(\mu - \lambda_i)t)) \right] \times \exp(1 - \sum_{i=1}^{k-1} \lambda_i t)
\]
2. Time to execution, sketch of a proof

• The probability of executing within time \( t \) a limit order placed in queue 1, therefore, is
\[
P(W_{\infty,1} \leq t) = 1 - (\mu - \lambda_1) \exp(-(\mu - \lambda_1)t)
\]
• The probability of executing within time \( t \) a limit order placed in queue 2 when queue 1 is empty is then
\[
P(W_{\infty,2} \leq t | I_{\infty,k=1} = 0) = (1 - (\mu - \lambda_2) \exp(-(\mu - \lambda_2)t))P(no\ arrivals\ in\ k = 1\ until\ t)
\]
• \( P(W_{\infty,2} \leq t | I_{\infty,k=1} = 0) = (1 - (\mu - \lambda_2) \exp(-(\mu - \lambda_2)t)) \exp(1 - \lambda_1 t) \)
• ...

© Copyright Irene Aldridge 2013
2. So what?

- Can deduce return characteristics

\[
\mathbb{E}[\Delta S_t] = \sum_{k>0} \Delta S_{k,t} \times \mathbb{P}[\Delta S_{k,t}] = \sum_{k>0} \Delta S_{k,t} \times P(W_{\infty,k} \leq t)
\]

\[
\mathbb{V}[\Delta S_t] = \sum_{k>0} (\Delta S_{k,t} - \mathbb{E}[\Delta S_t])^2 \times \mathbb{P}[\Delta S_{k,t}] = \sum_{k>0} (\Delta S_{k,t} - \mathbb{E}[\Delta S_t])^2 \times P(W_{\infty,k} \leq t)
\]

\[
\mathbb{K}[\Delta S_t] = \left(\sum_{k>0} (\Delta S_{k,t} - \mathbb{E}[\Delta S_t])^2 \times \mathbb{P}[\Delta S_{k,t}]\right)^{-2} \left(\sum_{k>0} (\Delta S_{k,t} - \mathbb{E}[\Delta S_t])^4 \times \mathbb{P}[\Delta S_{k,t}]\right)
\]

\[
= \left(\sum_{k>0} (\Delta S_{k,t} - \mathbb{E}[\Delta S_t])^2 \times P(W_{\infty,k} \leq t)\right)^{-2} \left(\sum_{k>0} (\Delta S_{k,t} - \mathbb{E}[\Delta S_t])^4 \times P(W_{\infty,k} \leq t)\right)
\]

Example:

\[
P(W_{\infty,k} \leq t) = \left[\prod_{i=1}^{k}(1 - (\mu + \delta_i - \lambda_i) \exp(-((\mu + \delta_i - \lambda_i)t)))\right] \times \exp(1 - \sum_{i=1}^{k-1} \lambda_i t)
\]

© Copyright Irene Aldridge 2013
2. And liquidity characteristics

- \( E[z_k] = \sum_{n>0} nP(z_k = n) \)
- \( \mathbb{V}[z_k] = \sum_{n>0} (n - \mathbb{E}[z_k])^2 P(z_k = n) \)
- \( \mathbb{K}[z_k] = (\sum_{n>0} (n - \mathbb{E}[z_k])^2 P(z_k = n))^2 \left( \sum_{n>0} (n - \mathbb{E}[z_k])^4 P(z_k = n) \right) \)

where

\[
P(I_{\infty,k} = z) = \left( 1 + \sum_{z=1}^{\infty} \frac{\lambda_1^z}{\prod_{j=1}^{z}(\mu + \delta_1)} \right)^{-1} \left( 1 + \sum_{z=1}^{\infty} \frac{\lambda_2^z}{\prod_{j=1}^{z}(\mu + \delta_2)} \right)^{-1} \ldots \left( 1 + \sum_{z=1}^{\infty} \frac{\lambda_k^z}{\prod_{j=1}^{z}(\mu + \delta_k)} \right)^{-1} \frac{\lambda_k^z}{\prod_{j=1}^{z}(\mu + \delta_k)}
\]

Liquidity, by definition, is “available immediacy of trading” (Demsetz, 1968), i.e., the number of limit orders available to match incoming market orders.

© Copyright Irene Aldridge 2013
2. Simulation 1: Top of the book probability of execution within a given time $t$

- $P(W_{\infty,1} \leq t) = (1 - (\mu + \delta_1 - \lambda_1) \exp(-(\mu + \delta_1 - \lambda_1)t))$

- Fix $\mu=5$, $t=1$ unit of time, vary $\delta_1$ and $\lambda_1$

<table>
<thead>
<tr>
<th>$\lambda_1 \setminus \delta_1$</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
<th>35</th>
<th>40</th>
<th>45</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.96631</td>
<td>0.999546</td>
<td>0.999995</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>0.96631</td>
<td>0.999546</td>
<td>0.999995</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>0.96631</td>
<td>0.999546</td>
<td>0.999995</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>1</td>
<td>0.96631</td>
<td>0.999546</td>
<td>0.999995</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>25</td>
<td>1</td>
<td>0.96631</td>
<td>0.999546</td>
<td>0.999995</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>1</td>
<td>0.96631</td>
<td>0.999546</td>
<td>0.999995</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>1</td>
<td>0.96631</td>
<td>0.999546</td>
<td>0.999995</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>1</td>
<td>0.96631</td>
<td>0.999546</td>
<td>0.999995</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>0.96631</td>
</tr>
</tbody>
</table>
2. Simulation 1: Top of the book probability of execution within a given time $t$

- Conclusions:
  1. Higher cancellation rate delta = higher probability of execution within a given wait time, but
  2. It is really

Higher cancellation / limit order arrival = higher probability of execution within a given wait time

In other words, when $\mu$ is fixed:

$$\text{lower } \frac{\lambda_1}{\delta_1} \Rightarrow \text{higher } P(W_{\infty,1} \leq t)$$

Furthermore, when $\mu$ is allowed to vary,

$$\text{lower } \frac{\lambda_1}{\delta_1 + \mu} \Rightarrow \text{higher } P(W_{\infty,1} \leq t)$$
2. Simulation 1: Top of the book probability of execution within a given time $t$

- Conclusions:

\[
\text{lower } \frac{\lambda_1}{\delta_1 + \mu} \Rightarrow \text{higher } P(W_{\omega,1} \leq t)
\]

Since

- $E[\Delta S_t] = \sum_{k>0} \Delta S_{k,t} \times P[\Delta S_{k,t}] = \sum_{k>0} \Delta S_{k,t} \times P(W_{\omega,k} \leq t)$
- $V[\Delta S_t] = \sum_{k>0} (\Delta S_{k,t} - E[\Delta S_t])^2 \times P[\Delta S_{k,t}] = \sum_{k>0} (\Delta S_{k,t} - E[\Delta S_t])^2 \times P(W_{\omega,k} \leq t)$
- $K[\Delta S_t] = \left(\sum_{k>0} (\Delta S_{k,t} - E[\Delta S_t])^2 \times P[\Delta S_{k,t}]\right)^{-2} \left(\sum_{k>0} (\Delta S_{k,t} - E[\Delta S_t])^4 \times P[\Delta S_{k,t}]\right)$
- $= \left(\sum_{k>0} (\Delta S_{k,t} - E[\Delta S_t])^2 \times P(W_{\omega,k} \leq t)\right)^{-2} \left(\sum_{k>0} (\Delta S_{k,t} - E[\Delta S_t])^4 \times P(W_{\omega,k} \leq t)\right)$

\[
\text{lower } \frac{\lambda_1}{\delta_1 + \mu} =>
\]
- Bigger price changes
- Higher volatility
- Higher probability of extreme events (black swans)
2. Simulation 2: Probability of finding $z$ orders at the top of the book

- $P(I_{\infty,1} = z) = \frac{\lambda_1^z}{\Pi_{j=1}^z (\mu + \delta_1)} \left( 1 + \sum_{z=1}^\infty \frac{\lambda_1^z}{\Pi_{j=1}^z (\mu + \delta_1)} \right)^{-1}$

- Fix $\mu=5$, $z=1$, vary $\delta_1$ and $\lambda_1$

<table>
<thead>
<tr>
<th>$\lambda_1 \backslash \delta_1$</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
<th>35</th>
<th>40</th>
<th>45</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.333333</td>
<td>0.25</td>
<td>0.2</td>
<td>0.166667</td>
<td>0.142857</td>
<td>0.125</td>
<td>0.111111</td>
<td>0.1</td>
<td>0.090909</td>
<td>0.083333</td>
</tr>
<tr>
<td>10</td>
<td>0.5</td>
<td>0.4</td>
<td>0.333333</td>
<td>0.285714</td>
<td>0.25</td>
<td>0.222222</td>
<td>0.2</td>
<td>0.181818</td>
<td>0.166667</td>
<td>0.153846</td>
</tr>
<tr>
<td>15</td>
<td>0.5</td>
<td>0.428571</td>
<td>0.375</td>
<td>0.333333</td>
<td>0.3</td>
<td>0.272727</td>
<td>0.25</td>
<td>0.230769</td>
<td>0.214286</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.5</td>
<td>0.444444</td>
<td>0.4</td>
<td>0.363636</td>
<td>0.333333</td>
<td>0.307692</td>
<td>0.285714</td>
<td>0.266667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>0.5</td>
<td>0.454545</td>
<td>0.416667</td>
<td>0.384615</td>
<td>0.357143</td>
<td>0.333333</td>
<td>0.3125</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0.5</td>
<td>0.461538</td>
<td>0.428571</td>
<td>0.4</td>
<td>0.375</td>
<td>0.352941</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>0.5</td>
<td>0.466667</td>
<td>0.4375</td>
<td>0.411765</td>
<td>0.388889</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>0.5</td>
<td>0.470588</td>
<td>0.444444</td>
<td>0.421053</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>0.5</td>
<td>0.473684</td>
<td>0.45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.5</td>
<td>0.47619</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

© Copyright Irene Aldridge 2013
2. Simulation 2: Probability of finding \( z \) orders at the top of the book

- Conclusions:
  1. Higher \( z \) => lower probability of finding \( z \) orders
  2. Higher cancellation rate \( \delta \) = lower probability of finding \( z \) orders, but
  3. It is really

Higher cancellation / limit order arrival = lower probability of finding \( z \) orders

In other words, for fixed \( \mu \):

\[
\text{lower } \frac{\lambda_1}{\delta_1} \implies \text{lower } P(I_{\infty,1} = z)
\]

When \( \mu \) is allowed to vary:

\[
\text{lower } \frac{\lambda_1}{\delta_1 + \mu} \implies \text{lower } P(I_{\infty,1} = z)
\]
2. Simulation 2: Probability of finding $z$ orders at the top of the book

- Conclusions:
  \[
  \text{lower} \quad \frac{\lambda_1}{\delta_1 + \mu} \quad \Rightarrow \quad \text{lower} \quad P(I_{\infty,1} = z)
  \]

Since Liquidity moments

- $\mathbb{E}[z_k] = \sum_{n>0} n P(z_k = n)$
- $\mathbb{V}[z_k] = \sum_{n>0} (n - \mathbb{E}[z_k])^2 P(z_k = n)$
- $\mathbb{K}[z_k] = \left( \sum_{n>0} (n - \mathbb{E}[z_k])^2 P(z_k = n) \right)^{-2} \left( \sum_{n>0} (n - \mathbb{E}[z_k])^4 P(z_k = n) \right)$

\[
\text{lower} \quad \frac{\lambda_1}{\delta_1 + \mu} \Rightarrow
\]
- Lower liquidity
- Lower liquidity volatility
- Lower instances of extreme liquidity (positive or negative)
2. Can think of market quality

\[ MQ = \frac{\lambda_1}{\delta_1 + \mu} \]

Higher \( MQ \) =>

- Smaller price changes
- Lower volatility
- Lower probability of extreme events (black swans)
- Higher liquidity
- Higher liquidity volatility
- Higher instances of extreme liquidity (positive or negative)
Outline

1. Modern market microstructure
2. Analytical results
3. Empirical results
4. The problem of market manipulation
5. Conclusions
3. Methodology

Setup

- Use Reuters tick data for December 2012
  - Level I data:
    - Security identification
    - Microsecond timestamp
    - Best bid price and size
    - Best ask (offer) price and size
  - Last trade price and size
- 3,249 securities traded on PSX in December 2012

Exchanges

<table>
<thead>
<tr>
<th>Exchange</th>
<th>Pricing</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>BX</td>
<td>Inverted</td>
<td>Price-time</td>
</tr>
<tr>
<td>EDGEA</td>
<td>Inverted</td>
<td>Price-time</td>
</tr>
<tr>
<td>EDGEX</td>
<td>Normal</td>
<td>Price-time</td>
</tr>
<tr>
<td>BATS</td>
<td>Normal</td>
<td>Price-time</td>
</tr>
<tr>
<td>BATS Y</td>
<td>Inverted</td>
<td>Price-time</td>
</tr>
<tr>
<td>PSX</td>
<td>Normal</td>
<td>Pro-rata</td>
</tr>
</tbody>
</table>
3. Methodology

Data Organization

- Divide all data into 1, 10, 30, 60, 300 second “bars”
- Use volume clock for orders
  - A 100-share order is considered to be 100 independent instantaneously-arriving 1-share orders
  - First proposed by Easley, Lopez de Prado, O’Hara (2011)

Variables

- Count within each bar:
  - Number of limit orders arriving at the top of the book:
    - $\lambda_{1,t}^b$ for limit bids
    - $\lambda_{1,t}^a$ for limit asks (offers)
  - Number of market orders arriving:
    - $\mu_t^b$ for market buy orders
    - $\mu_t^s$ for market sell orders
  - Number of limit order cancellations at the top of the book:
    - $\delta_{1,t}^b$ for limit bid cancellations
    - $\delta_{1,t}^a$ for limit ask cancellations
### 3. Counting order arrivals

**Example, tick data**

<table>
<thead>
<tr>
<th></th>
<th>1.08</th>
<th>50</th>
<th>1.10</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q</td>
<td>1.08</td>
<td>20</td>
<td>1.10</td>
<td>100</td>
</tr>
<tr>
<td>T</td>
<td>1.08</td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.08</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>1.07</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q</td>
<td>1.07</td>
<td>50</td>
<td>1.09</td>
<td>80</td>
</tr>
<tr>
<td>Q</td>
<td>1.08</td>
<td>100</td>
<td>1.09</td>
<td>80</td>
</tr>
<tr>
<td>Q</td>
<td>1.08</td>
<td>200</td>
<td>1.09</td>
<td>80</td>
</tr>
<tr>
<td>Q</td>
<td>1.08</td>
<td>200</td>
<td>1.09</td>
<td>50</td>
</tr>
<tr>
<td>T</td>
<td>1.09</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.08</td>
<td>200</td>
<td>1.09</td>
<td>50</td>
</tr>
</tbody>
</table>

- Market sell 30
- Market sell 20
- Market sell 100
- Limit sell 80
- Limit buy 50
- Limit buy 100
- Cancel sell 30
- Market buy 50
3. Counting order arrivals

Did a trade print just go through?  

Yes

Was the trade print closer to BB than to BO?  

Yes

The trade was initiated by a market sell:  
\[ \mu_t^s = \mu_t^s + \text{trade size} \]

No

Was the trade print closer to BO than to BB?  

Yes

The trade was initiated by a market buy:  
\[ \mu_t^b = \mu_t^b + \text{trade size} \]

No

No

Go to next page

Use tick rule  
(per Lee Ready, 1991)
3. Counting order arrivals

Did a trade print just go through?
Yes → Go to previous page
No → Did the best bid price change?
Yes → Were the best bid price or size increased?
Yes → New best bid arrived
\[ \lambda_{1,t}^b = \lambda_{1,t}^b + \Delta|\text{bid size}| \]
No → Repeat same for ask (offer) side
No → Were the best bid price or size decreased?
Yes → New cancel bid arrived
\[ \delta_{1,t}^b = \delta_{1,t}^b + |\Delta\text{bid size}| \]
No → Did the best bid price change?
Yes → Go to previous page
No → Repeat same for ask (offer) side
3. Key results

Limit Buy order arrivals  
\[ \text{Adj. } R^2 > 95\% \]

- Limit buy order arrival rate  
  \[ \lambda_{1,t}^b \]

  - Depends on
    - Limit buy cancellation rate  
      \[ \delta_{1,t}^b \]  
      (coefficient ranges from 0.2 to 1, depends on security, t-stat > 30)
    - Market sell arrival rate  
      \[ \mu_t^s \]  
      (coefficient ranges from 0.001 to 0.01, t-stat > 5)
    - Contemporaneous return  
      (positive coefficient, t-stat > 200)

  - Does not depend on: market type

Limit Sell order arrivals  
\[ \text{Adj. } R^2 > 95\% \]

- Limit sell order arrival rate  
  \[ \lambda_{1,t}^s \]

  - Depends on
    - Limit buy cancellation rate  
      \[ \delta_{1,t}^s \]  
      (coefficient ranges from 0.2 to 1, depends on security, t-stat > 30)
    - Market Buy arrival rate  
      \[ \mu_t^b \]  
      (coefficient ranges from 0.001 to 0.01, t-stat > 5)
    - Contemporaneous return  
      (negative coefficient, t-stat > 200)

  - Does not depend on: market type
3. Key results

Said earlier: it is really the $MQ = \frac{\lambda_1}{\delta_1 + \mu}$ that matters.

Higher $MQ =>$
- Smaller price changes
- Lower volatility
- Lower probability of extreme events (black swans)
- Higher liquidity
- Higher liquidity volatility
- Higher instances of extreme liquidity (positive or negative)
3. Market Quality $MQ = \frac{\lambda_1}{\delta_1 + \mu}$ Buy Analysis

<table>
<thead>
<tr>
<th>MQ Buy</th>
<th>Intercept</th>
<th>Return_300</th>
<th>Inverted</th>
<th>Pro-Rata</th>
<th>Adj. R Square</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.25543</td>
<td>22.94372</td>
<td>0.282047</td>
<td>1.963681</td>
<td>0.280422</td>
<td>8099</td>
</tr>
<tr>
<td></td>
<td>T-stat</td>
<td>13.36384</td>
<td>2.945253</td>
<td>10.04814</td>
<td>55.07404</td>
<td></td>
</tr>
<tr>
<td>AB</td>
<td>0.488891</td>
<td>8.139966</td>
<td>0.781161</td>
<td>0.085976</td>
<td>0.062137</td>
<td>6127</td>
</tr>
<tr>
<td></td>
<td>T-stat</td>
<td>17.82281</td>
<td>1.040882</td>
<td>19.76173</td>
<td>1.188815</td>
<td></td>
</tr>
<tr>
<td>ABB</td>
<td>0.572639</td>
<td>-66.1061</td>
<td>0.046547</td>
<td>0.379842</td>
<td>0.038345</td>
<td>8785</td>
</tr>
<tr>
<td></td>
<td>T-stat</td>
<td>47.30008</td>
<td>-8.38283</td>
<td>2.670016</td>
<td>16.46376</td>
<td></td>
</tr>
</tbody>
</table>

Conclusions:
- Inverted markets have higher buy MQ than do normal markets
- Pro-rata markets have higher buy MQ than do normal and inverted markets

© Copyright Irene Aldridge 2013
3. Market Quality $MQ = \frac{\lambda_1}{\delta_1 + \mu}$

<table>
<thead>
<tr>
<th>MQ Sell</th>
<th>Intercept</th>
<th>Return_300</th>
<th>Inverted</th>
<th>Pro-Rata</th>
<th>Adj. R Square</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.242525</td>
<td>-32.6824</td>
<td>0.276434</td>
<td>1.937734</td>
<td>0.263442</td>
<td>8099</td>
</tr>
<tr>
<td></td>
<td>T-stat</td>
<td>12.2999</td>
<td>-4.06685</td>
<td>9.54644</td>
<td>52.6812</td>
<td></td>
</tr>
<tr>
<td>AB</td>
<td>0.539573</td>
<td>-19.5682</td>
<td>0.662541</td>
<td>-0.09036</td>
<td>0.054618</td>
<td>6127</td>
</tr>
<tr>
<td></td>
<td>T-stat</td>
<td>20.72015</td>
<td>-2.63578</td>
<td>17.65531</td>
<td>-1.31616</td>
<td></td>
</tr>
<tr>
<td>ABB</td>
<td>0.572639</td>
<td>-66.1061</td>
<td>0.046547</td>
<td>0.379842</td>
<td>0.038345</td>
<td>8785</td>
</tr>
<tr>
<td></td>
<td>T-stat</td>
<td>47.30008</td>
<td>-8.38283</td>
<td>2.670016</td>
<td>16.46376</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

- Conclusions:
  - Inverted markets have higher sell MQ than do normal markets
  - Pro-rata markets have higher sell MQ than do normal and inverted markets
3. Further results

- The model predicts well market characteristics out-of-sample
- Results available in the forthcoming paper
Outline

1. Modern market microstructure
2. Analytical results
3. Empirical results
4. The problem of market manipulation
5. Conclusions
4. Market manipulation

- Regulators recognize one HFT-related form of manipulation: spoofing
- Spoofing is illegal and has been prosecuted
- How does spoofing work?
  - Traders and algos make short-term inferences from the shape of the limit order book
  - Spoofers distort the limit order book to mislead others, main point: cause ST market move, but cancel limit orders prior to execution
  - Stylized example:
4. Detection is complicated

• In price-time priority markets, many broker-dealers earn their living by securing time priority for their clients.
• But, most of those broker-dealers would place orders on both sides of the market.
• Still, some brokers may anticipate clients’ demands and place orders on one side only.
4. Spoofing ability depends on microstructure

- Proposition 5 (Aldridge 2013):
  - Spoofing is most likely in vanilla price-time priority markets
  - Spoofing is least likely and easiest to identify in markets with wide no-cancel range

- Sketch of a proof: spoofers never place fake orders in no-cancel range
4. Key observation

- Under price-time priority, broker-dealers place and cancel lots and lots of limit orders
- Under pro-rata, broker-dealers get filled only partially, have to place larger orders, cancel some excess orders
  - Still, pro-rata cancellation rates are lower than price-time priority
- Under no-cancel range, order cancellations are disallowed altogether
Outline

1. Modern market microstructure
2. Analytical results
3. Empirical results
4. The problem of market manipulation
5. Conclusions
Key Conclusions

- Different order matching principles result in theoretically and empirically different market characteristics:

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Normal</th>
<th>Inverted</th>
<th>Pro-rata</th>
<th>No-cancel range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order cancellations at top of the book</td>
<td>High</td>
<td>Medium</td>
<td>Low</td>
<td>Zero</td>
</tr>
<tr>
<td>Top-of-the-book liquidity “quality”</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
<td>Highest</td>
</tr>
<tr>
<td>Ability to detect spoofing market manipulation</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
<td>Highest</td>
</tr>
</tbody>
</table>
Key Conclusions (cont.)

- Different order matching principles result in theoretically and empirically different market characteristics:

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Vanilla Price-Time Priority</th>
<th>No-Cancel-Range Either Priority</th>
<th>Vanilla Pro-Rata Matching</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market volatility</td>
<td>High</td>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>Fat tails of returns</td>
<td>High</td>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>Market liquidity</td>
<td>High</td>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>Ability to detect spoofing market manipulation</td>
<td>Low</td>
<td>High</td>
<td>High</td>
</tr>
</tbody>
</table>
Implications

• Exchanges minimizing adverse HFT activity exist
  • For those concerned about spoofing:
    • Pro-rata exchanges are better than price-time priority
    • No cancel ranges are better than vanilla exchanges
  • In futures, CME E-minis are traded with pro-rata, as do many other instruments
  • No such exchanges in equities at the moment
    • PSX was converted to price-time priority in May 2013
    • Plenty of room for new entrants