Multiscale Modeling of Order Book Dynamics

Jose Blanchet (with Xinyun Chen and Yanan Pei)

Columbia University

May, 2015
1. Goal of the Talk

2. Our Model: Definition and Empirical Validation

3. Price Formation via Queueing Microstructure

4. Cancellation Policy and Continuous Time Dynamics

5. Conclusions
Goal of the Talk

Goal: Present and discuss a model for price and the bid-ask spread which:

a) Is informed by the full order book dynamics,
b) It captures key stylized features observed empirically,
c) Useful in intra-day trading (many minutes / few hours).
A Picture of a Limit Order Book

FIG. 1. Schematic of a LOB

- Ask Side
  - Sell limit order
  - Mid-price
  - Ask price

- Bid Side
  - Buy limit order
  - Spread
Outline

1. Goal of the Talk

2. Our Model: Definition and Empirical Validation

3. Price Formation via Queueing Microstructure

4. Cancellation Policy and Continuous Time Dynamics

5. Conclusions
What’s the Final Model

- $S(t) = \text{size of bid-ask spread}$ \& $M(t) = \text{mid-price}$.

\[
\begin{align*}
    dS(t) & = \mathcal{W}_{\mu,\sigma}(t) + S(t^-)dJ_+(t) + S(t^-)dJ_-(t) + dL(t), \\
    dM(t) & = \bar{\mathcal{W}}_{\bar{\mu},\bar{\sigma}}(t) + S(t^-)dJ_+(t) - S(t^-)dJ_-(t), \\
    S(t)dL(t) & = 0, \\
    dL(t) & \geq 0.
\end{align*}
\]
What’s the Final Model

- $S(t) = \text{size of bid-ask spread}$ \& $M(t) = \text{mid-price}.$

\[
    dS(t) = \mathcal{W}_{\mu,\sigma}(t) + S(t^-) dJ_+(t) + S(t^-) dJ_-(t) + dL(t),
\]
\[
    dM(t) = \tilde{\mathcal{W}}_{\bar{\mu},\bar{\sigma}}(t) + S(t^-) dJ_+(t) - S(t^-) dJ_-(t),
\]
\[
    S(t) dL(t) = 0,
\]
\[
    dL(t) \geq 0.
\]

- $\mathcal{W}_{\mu,\sigma}$ and $\tilde{\mathcal{W}}_{\bar{\mu},\bar{\sigma}}$ are independent Brownian motions.
What’s the Final Model

- $S(t) = \text{size of bid-ask spread}$ \& $M(t) = \text{mid-price}$.

\[
\begin{align*}
    dS(t) &= \mathcal{W}_{\mu,\sigma}(t) + S(t_-) dJ_+(t) + S(t_-) dJ_-(t) + dL(t), \\
    dM(t) &= \tilde{\mathcal{W}}_{\tilde{\mu},\tilde{\sigma}}(t) + S(t_-) dJ_+(t) - S(t_-) dJ_-(t), \\
    S(t) dL(t) &= 0, \\
    dL(t) &\geq 0.
\end{align*}
\]

- $\mathcal{W}_{\mu,\sigma}$ and $\tilde{\mathcal{W}}_{\tilde{\mu},\tilde{\sigma}}$ are independent Brownian motions.

- $J_-(\cdot)$ and $J_+(\cdot)$ independent compound Poisson processes with jumps $V_-$ and $V_+$
How is this Model Related to the Order Book?

- When do we expect this model to perform well for prices?

Answer: When applied to assets with relatively large spread sizes variation (relative to the volatility) over medium time horizon (several minutes, maybe up to a few hours).
How is this Model Related to the Order Book?

When do we expect this model to perform well for prices?

Answer: When applied to assets with relatively large spread sizes variation (relative to the volatility) over medium time horizon (several minutes, maybe up to a few hours).
How is this Model Related to the Order Book?

- When do we expect this model to perform well for prices?
- Answer: *When applied to assets with relatively large spread sizes variation (relative to the volatility) over medium time horizon (several minutes, maybe up to a few hours).*
- How is this model informed by the order book?
How Does the Order Book Enters?

- $A(t) = \text{Ask Price}$, $B(t) = \text{Bid Price}$.
How Does the Order Book Enters?

- $A(t) = \text{Ask Price}$, $B(t) = \text{Bid Price}$.
- $\bar{\Pi}_{BUY}(x; A(t), B(t)) = \text{"Buy" orders at price } < B(t) - x.$
How Does the Order Book Enters?

- $A(t) = \text{Ask Price, } B(t) = \text{Bid Price.}$
- $\Pi_{BUY}(x; A(t), B(t)) = \text{"Buy" orders at price } < B(t) - x.$
- $\Pi_{SELL}(x; A(t), B(t)) = \text{"Sell" orders at price } > A(t) + x.$
How Does the Order Book Enters?

- \( A(t) = \) Ask Price, \( B(t) = \) Bid Price.
- \( \tilde{\Pi}_{BUY}(x; A(t), B(t)) = \) "Buy" orders at price \(< B(t) - x\).
- \( \tilde{\Pi}_{SELL}(x; A(t), B(t)) = \) "Sell" orders at price \(> A(t) + x\).
- \( \gamma_{BUY} = \) patience ratio for "buy" orders (\( \gamma_{BUY} \approx 0 \) little patience \( \approx \) high cancellation).
How Does the Order Book Enters?

- $A(t) = \text{Ask Price}, \quad B(t) = \text{Bid Price}$.
- $\tilde{\Pi}_{BUY}(x; A(t), B(t)) = \text{"Buy" orders at price } < B(t) - x$.
- $\tilde{\Pi}_{SELL}(x; A(t), B(t)) = \text{"Sell" orders at price } > A(t) + x$.
- $\gamma_{BUY} = \text{patience ratio for } \text{"buy" orders} \ (\gamma_{BUY} \approx 0 \text{ little patience = high cancellation})$.
- $\gamma_{SELL} = \text{patience ratio for } \text{"sell" orders}$.
How Does the Order Book Enters?

Then for $x \geq \varepsilon_0 \geq 0$

$$\Pi_{SELL}(x; A(t), B(t))^{\gamma_{SELL}} = P(V_+ > x/S(t)).$$

and

$$\Pi_{BUY}(x; A(t), B(t))^{\gamma_{BUY}} = P(V_- > x/S(t)).$$
Empirical Validation Procedure

- **Step 1:** Fit model to **MID PRICE AND SPREAD SIZES** ← **USING PRICE SERIES ONLY.**
Empirical Validation Procedure

- **Step 1:** Fit model to MID PRICE AND SPREAD SIZES \( \leftarrow \) USING PRICE SERIES ONLY.
- **Step 2:** Estimate \( \tilde{\Pi}_{BUY}(\cdot) \) and \( \tilde{\Pi}_{SELL}(\cdot) \) from ORDER BOOK \( \leftarrow \) USING ORDER BOOK ONLY.
Empirical Validation Procedure

- **Step 1:** Fit model to MID PRICE AND SPREAD SIZES \(<\) USING PRICE SERIES ONLY.

- **Step 2:** Estimate $\bar{\Pi}_{\text{BUY}}(\cdot)$ and $\bar{\Pi}_{\text{SELL}}(\cdot)$ from ORDER BOOK \(<\) USING ORDER BOOK ONLY.

- **Step 3:** Plot $\log \bar{F}_{V+}(x/S(t))$ vs $\log \Pi_{\text{SELL}}(x; A(t), B(t))$. & same for BUY side.
Empirical Validation Procedure

- **Step 1:** Fit model to MID PRICE AND SPREAD SIZES ← USING PRICE SERIES ONLY.
- **Step 2:** Estimate $\bar{\Pi}_{BUY}(\cdot)$ and $\bar{\Pi}_{SELL}(\cdot)$ from ORDER BOOK ← USING ORDER BOOK ONLY.
- **Step 3:** Plot $\log \bar{F}_v^+(x/S(t))$ vs $\log \Pi_{SELL}(x; A(t), B(t))$. & same for BUY side.
- **Outcome should look like a straight line!**
Some More Empirical Analysis

Pictures here...
Warning: Don’t Use Model for Small Spread-Size Stocks

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Google</td>
<td>28.69</td>
<td>27.07</td>
<td>34.11</td>
<td>30.39</td>
<td>27.23</td>
</tr>
<tr>
<td>Facebook</td>
<td>1.43</td>
<td>1.41</td>
<td>1.68</td>
<td>1.42</td>
<td>1.45</td>
</tr>
<tr>
<td>Amazon</td>
<td>16.22</td>
<td>14.86</td>
<td>21.95</td>
<td>20.58</td>
<td>17.23</td>
</tr>
</tbody>
</table>
Important Implications

Bid-Ask processes only encode lots of info from full order book!

Full order book directly feeds the dynamics of Bid-Ask processes!

This whole encoding obeys relatively simple statistical rules!
(Proportional hazards.)
Outline

1. Goal of the Talk
2. Our Model: Definition and Empirical Validation
3. Price Formation via Queueing Microstructure
4. Cancellation Policy and Continuous Time Dynamics
5. Conclusions
Objective: Use queueing theory to explain price formation.
Objective: Use queueing theory to explain price formation.

Motivation:
Objective: Use queueing theory to explain price formation.

Motivation:

1. Understand multiscale nature of the problem.
Objective: Use queueing theory to explain price formation.

Motivation:

1. Understand multiscale nature of the problem.
2. Understand role of cancellation.
Objective: Use queueing theory to explain price formation.

Motivation:

1. Understand multiscale nature of the problem.
2. Understand role of cancellation.
3. Understand why so much information can be decoded from prices only?
Assumptions

1. Arrival Limit Orders \( \lambda_n \gg \mu_n \) = Arrival Market Orders.
1. Arrival Limit Orders $= \lambda_n \gg \mu_n = \text{Arrival Market Orders.}$
2. Times of market orders $\{t_k\}$ (Poisson $\Rightarrow$ can be relaxed).
Assumptions

1. Arrival Limit Orders = $\lambda_n \gg \mu_n$ = Arrival Market Orders.
2. Times of market orders $\{t_k\}$ (Poisson $\rightarrow$ can be relaxed).
3. Limit Orders between market orders at relative price $i\delta$ with $p_{BUY}(i\delta, \tilde{A}(t_k), \tilde{B}(t_k)) \cdot \delta$ (same for SELL).
Assumptions

1. Arrival Limit Orders $= \lambda n >> \mu n = $ Arrival Market Orders.
2. Times of market orders $\{ t_k \}$ (Poisson $\rightarrow$ can be relaxed).
3. Limit Orders between market orders at relative price $i\delta$ with $p_{BUY}(i\delta, \bar{A}(t_k), \bar{B}(t_k)) \cdot \delta$ (same for SELL).
4. No cross over of limit orders to opposite side of mid-price.
Assumptions

1. Arrival Limit Orders $= \lambda_n \gg \mu_n = \text{Arrival Market Orders.}$
2. Times of market orders $\{t_k\}$ (Poisson $\rightarrow$ can be relaxed).
3. Limit Orders between market orders at relative price $i\delta$ with $p_{\text{BUY}}(i\delta, \bar{A}(t_k), \bar{B}(t_k)) \cdot \delta$ (same for SELL).
4. No cross over of limit orders to opposite side of mid-price.
5. Each order at relative price $i\delta$ cancels at rate $\alpha_{\text{BUY}}(i\delta, \bar{A}(t_k), \bar{B}(t_k))$ or $\alpha_{\text{SELL}}(i\delta, \bar{A}(t_k), \bar{B}(t_k))$. 
Arrival Limit Orders $= \lambda_n \gg \mu_n = \text{Arrival Market Orders}$:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Google</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Limit Orders</td>
<td>80,537</td>
<td>78,944</td>
<td>77,215</td>
<td>100,798</td>
<td>66,238</td>
</tr>
<tr>
<td>Total Market Orders</td>
<td>9,412</td>
<td>8,016</td>
<td>5,868</td>
<td>8,505</td>
<td>7,038</td>
</tr>
<tr>
<td>Facebook</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Limit Orders</td>
<td>442,425</td>
<td>483,338</td>
<td>489,886</td>
<td>472,251</td>
<td>363,833</td>
</tr>
<tr>
<td>Total Market Orders</td>
<td>31,973</td>
<td>37,378</td>
<td>29,456</td>
<td>36,558</td>
<td>30,455</td>
</tr>
<tr>
<td>Amazon</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Limit Orders</td>
<td>100,263</td>
<td>131,648</td>
<td>125,555</td>
<td>162,561</td>
<td>123,127</td>
</tr>
<tr>
<td>Total Market Orders</td>
<td>13,148</td>
<td>14,804</td>
<td>9,225</td>
<td>11,344</td>
<td>10,094</td>
</tr>
</tbody>
</table>
**A2:** Can use Hawkes instead of Poisson market orders (we’ll see why).
Discussion About Assumptions

- **A2:** Can use Hawkes instead of Poisson market orders (we’ll see why).
- **A3:** $p_{BUY}(i\delta, \bar{A}(t_k), \bar{B}(t_k)) \cdot \delta$ prob. of placing order at $i\delta$, standard in literature
Discussion About Assumptions

- **A2:** Can use Hawkes instead of Poisson market orders (we’ll see why).
- **A3:** $p_{BUY}(i\delta, \bar{A}(t_k), \bar{B}(t_k)) \cdot \delta$ prob. of placing order at $i\delta$, standard in literature
- **A4:** No cross over of limit orders to opposite side of mid-price.

<table>
<thead>
<tr>
<th>Proportion of Cross Limit Orders</th>
<th>4/27</th>
<th>4/28</th>
<th>4/29</th>
<th>4/30</th>
<th>5/1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google</td>
<td>2.12%</td>
<td>1.77%</td>
<td>2.31%</td>
<td>1.79%</td>
<td>1.84%</td>
</tr>
<tr>
<td>Facebook</td>
<td>2.82%</td>
<td>3.80%</td>
<td>4.77%</td>
<td>3.40%</td>
<td>3.01%</td>
</tr>
<tr>
<td>Amazon</td>
<td>2.55%</td>
<td>2.47%</td>
<td>2.45%</td>
<td>1.57%</td>
<td>1.94%</td>
</tr>
</tbody>
</table>
• **A5:** $\alpha_{BUY}(i\delta, \bar{A}(t_k), \bar{B}(t_k))$ cancellation rate $i\delta$ PER order not standard in literature BUT this makes sense...

  Whole system is a coupled multiclass two server queuing network.
Averaging Principle Between Market Order Arrivals

Theorem

(B., Chen & Pei) At arrival of \((k + 1)\)-th market order the order book follows the distribution of independent \(M/M/\infty\) queueing systems the \(i\)-th with parameter

\[
\rho_n(i) = \lambda_n \frac{\rho(i\delta, A(t_k), B(t_k))}{\alpha(i\delta, A(t_k), B(t_k))} \cdot \delta.
\]

Recall: Steady-state number in system of \(M/M/\infty\) is Poisson with parameter \(\rho_n(i)\).

Proof.

Averaging principle for Martingale problems (see Kurtz (1992)).
Outline

1. Goal of the Talk

2. Our Model: Definition and Empirical Validation

3. Price Formation via Queueing Microstructure

4. Cancellation Policy and Continuous Time Dynamics

5. Conclusions
**Assumption (C):** For $x > x_0$

$$\alpha (x, A, B) \approx \gamma \times (1 - \Pi (x, A, B)) .$$
Discussion of Assumption (C)

- Mainly introduced because of mathematical tractability.
Discussion of Assumption (C)

- Mainly introduced because of mathematical tractability.
- BUT allows us to establish a ONE to ONE correspondance between price increment & ORDER BOOK of the form

\[ \Pi (x; A, B)^\gamma \approx \theta (x, A, B), \]

for \( x > x_0 \).
Discussion of Assumption (C)

- *Mainly introduced because of mathematical tractability.*

- **BUT** allows us to establish a ONE to ONE correspondance between price increment & ORDER BOOK of the form

\[
\Pi (x; A, B)^\gamma \approx \theta (x, A, B),
\]

for \( x > x_0 \).

- Tends to increase as depth increases (very reasonable).
Discussion of Assumption (C)

- **Mainly introduced because of mathematical tractability.**
- **BUT allows us to establish a ONE to ONE correspondance between price increment & ORDER BOOK of the form**

\[
\Pi (x; A, B)^\gamma \approx \theta (x, A, B),
\]

for \( x > x_0 \).
- Tends to increase as depth increases (very reasonable).
- If \( \gamma \approx 1 \) one can argue that, the equilibrium rate of execution, is

\[
\frac{\mu \theta (i \delta, A, B)}{\mu \theta (i \delta, A, B) + \alpha (i \delta, A, B)} \approx \frac{\mu}{1 + \mu}
\]

is constant at any level.
Continuous Time Dynamics

- Given Assumption (C) there is a one to one correspondance between increment distribution & Limit Order Book structure.
Continuous Time Dynamics

- Given Assumption (C) there is a one to one correspondence between increment distribution & Limit Order Book structure.

- So, **auxiliary increment distribution:** \( \Delta_{k+1}(S(t_k)) \) depending on spread:

\[
\theta(x, A, B) = P(\max(\Delta_{k+1}(S(t_k)), -[S(t_k) / (2\delta)]\delta) > x).
\]
Continuous Time Dynamics

- Given Assumption (C) there is a one to one correspondence between increment distribution & Limit Order Book structure.
- So, auxiliary increment distribution: $\Delta_{k+1}(S(t_k))$ depending on spread:
  \[
  \theta(x, A, B) = P(\max(\Delta_{k+1}(S(t_k)), -[S(t_k)/(2\delta)]\delta) > x).
  \]
- Formula for $\Delta_{k+1}(S(t_k))$ ugly, but intuition simple... explain in words
  \[
  \Delta_{k+1}(S(t_k)) = (-1)^{R_{k+1}}(1 - I_{k+1})\delta_n\left[U_{k+1}/\left(n^{1/2}\delta_n\right)\right] + I_{k+1}[S(t_k)V_k/\delta_n]\delta_n.
  \]
Given Assumption (C) there is a one to one correspondence between increment distribution & Limit Order Book structure.

So, auxiliary increment distribution: \( \Delta_{k+1} (S(t_k)) \) depending on spread:

\[
\theta (x, A, B) = P (\max(\Delta_{k+1} (S(t_k))), -[S(t_k)/(2\delta)]\delta > x).
\]

Formula for \( \Delta_{k+1} (S(t_k)) \) ugly, but intuition simple... explain in words

\[
\Delta_{k+1} (S(t_k)) = (-1)^{R_k+1} (1 - I_{k+1}) \delta_n \left[ U_{k+1} / \left( n^{1/2} \delta_n \right) \right] \\
+ I_{k+1} [S(t_k) V_k / \delta_n] \delta_n.
\]

Second piece captures the structural part, namely, the fact that for \( x > x_0 \)

\[
\Pi (x; A, B)^{\gamma} \approx \theta (x, A, B).
\]
It turns out that in terms of auxiliary increment distribution $\Delta_{k+1}(S(t_k))$, one gets

$$A(t_{k+1}) = A(t_k) + \max(\Delta_{k+1}(S(t_k)), -\frac{S(t_k)(t_k)}{(2\delta)\delta})$$

similar for \textit{BUY} with "−".
Why Local Times in Limiting Process?

- It turns out that in terms of auxiliary increment distribution $\Delta_{k+1} (S (t_k))$, one gets
  
  $$A(t_{k+1}) = A(t_k) + (\max(\Delta_{k+1} (S (t_k)), -[S(t_k)/(2\delta)]) \delta),$$

  similar for $BUY$ with "-".

- The "max" operator "similar" to Skorokhod map in queueing theory (but NOT exactly).
Why Local Times in Limiting Process?

- It turns out that in terms of auxiliary increment distribution $\Delta_{k+1}(S(t_k))$, one gets

$$A(t_{k+1}) = A(t_k) + \left( \max(\Delta_{k+1}(S(t_k)), -\left[ S(t_k) / (2\delta) \right]\delta) \right),$$

similar for $BUY$ with "-".

- The "max" operator "similar" to Skorokhod map in queueing theory (but NOT exactly).

- We use techniques from diffusion approximations and continuity of the so-called Skorokhod map to establish the result.
Outline

1. Goal of the Talk

2. Our Model: Definition and Empirical Validation

3. Price Formation via Queueing Microstructure

4. Cancellation Policy and Continuous Time Dynamics

5. Conclusions
Conclusions

Presented and discussed a model for price and the bid-ask spread which:

a) Is informed by the full order book dynamics,

b) It captures key stylized features observed empirically,

c) Useful in intra-day trading (many minutes / few hours):
   
   Bid-Ask processes only encode lots of info from full order book.
   
   Full order book directly feeds the dynamics of Bid-Ask processes.