Multiscale Modeling of Order Book Dynamics

Jose Blanchet (with Xinyun Chen and Yanan Pei)

Columbia University

May, 2015

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

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

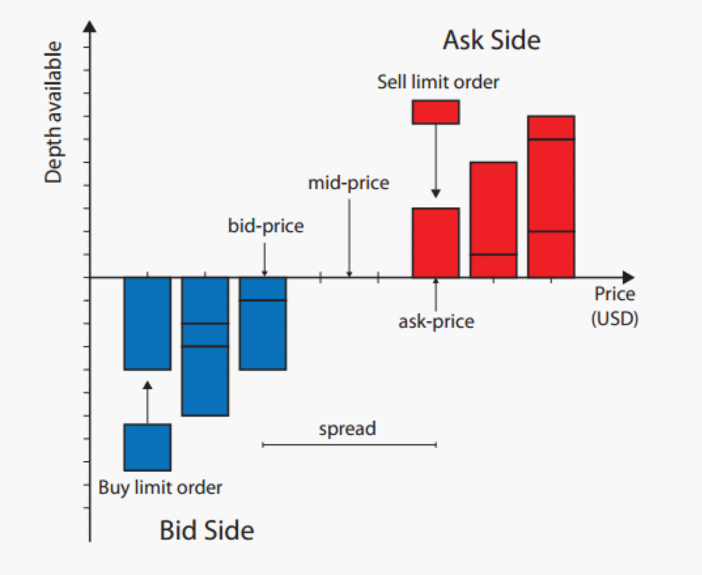


FIG. 1 Schematic of a LOB

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

- $S(t)$ **= size of bid-ask spread & $M(t)$ = mid-price.**

$$\begin{aligned} dS(t) &= W_{\mu,\sigma}(t) + S(t_-)\,dJ_+(t) + S(t_-)\,dJ_-(t) + dL(t), \\ dM(t) &= \bar{W}_{\bar{\mu},\bar{\sigma}}(t) + S(t_-)\,dJ_+(t) - S(t_-)\,dJ_-(t), \\ S(t)\,dL(t) &= 0, \\ dL(t) &\geq 0. \end{aligned}$$

What's the Final Model

- $S(t)$ **= size of bid-ask spread** **&** $M(t)$ **= mid-price.**

$$\begin{aligned} dS(t) &= W_{\mu,\sigma}(t) + S(t_-)\, dJ_+(t) + S(t_-)\, dJ_-(t) + dL(t), \\ dM(t) &= \bar{W}_{\bar{\mu},\bar{\sigma}}(t) + S(t_-)\, dJ_+(t) - S(t_-)\, dJ_-(t), \\ S(t)\, dL(t) &= 0, \\ dL(t) &\geq 0. \end{aligned}$$

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

What's the Final Model

- $S(t)$ **= size of bid-ask spread &** $M(t)$ **= mid-price.**

$$\begin{aligned} dS(t) &= W_{\mu,\sigma}(t) + S(t_-)\,dJ_+(t) + S(t_-)\,dJ_-(t) + dL(t), \\ dM(t) &= \bar{W}_{\bar{\mu},\bar{\sigma}}(t) + S(t_-)\,dJ_+(t) - S(t_-)\,dJ_-(t), \\ S(t)\,dL(t) &= 0, \\ dL(t) &\geq 0. \end{aligned}$$

- $W_{\mu,\sigma}$ and $\bar{W}_{\bar{\mu},\bar{\sigma}}$ 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?

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?

- $A(t)$ = Ask Price, $B(t)$ = Bid Price.

How Does the Order Book Enters?

- $A(t) =$ Ask Price, $B(t) =$ Bid Price.
- $\bar{\Pi}_{BUY}(x; A(t), B(t)) =$ "Buy" orders at price $< B(t) - x$.

How Does the Order Book Enters?

- $A(t) =$ Ask Price, $B(t) =$ Bid Price.
- $\bar{\Pi}_{BUY}(x; A(t), B(t)) =$ "Buy" orders at price $< B(t) - x$.
- $\bar{\Pi}_{SELL}(x; A(t), B(t)) =$ "Sell" orders at price $> A(t) + x$.

How Does the Order Book Enters?

- $A(t)$ = Ask Price, $B(t)$ = Bid Price.
- $\bar{\Pi}_{BUY}(x; A(t), B(t))$ = "Buy" orders at price $< B(t) - x$.
- $\bar{\Pi}_{SELL}(x; A(t), B(t))$ = "Sell" orders at price $> A(t) + x$.
- γ_{BUY} = patience *ratio* for "buy" orders ($\gamma_{BUY} \approx 0$ little patience = high cancellation).

How Does the Order Book Enters?

- $A(t)$ = Ask Price, $B(t)$ = Bid Price.
- $\bar{\Pi}_{BUY}(x; A(t), B(t))$ = "Buy" orders at price $< B(t) - x$.
- $\bar{\Pi}_{SELL}(x; A(t), B(t))$ = "Sell" orders at price $> A(t) + x$.
- γ_{BUY} = patience *ratio* for "buy" orders ($\gamma_{BUY} \approx 0$ little patience = high cancellation).
- γ_{SELL} = patience *ratio* for "sell" orders.

How Does the Order Book Enters?

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

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

and

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

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

Empirical Validation Procedure

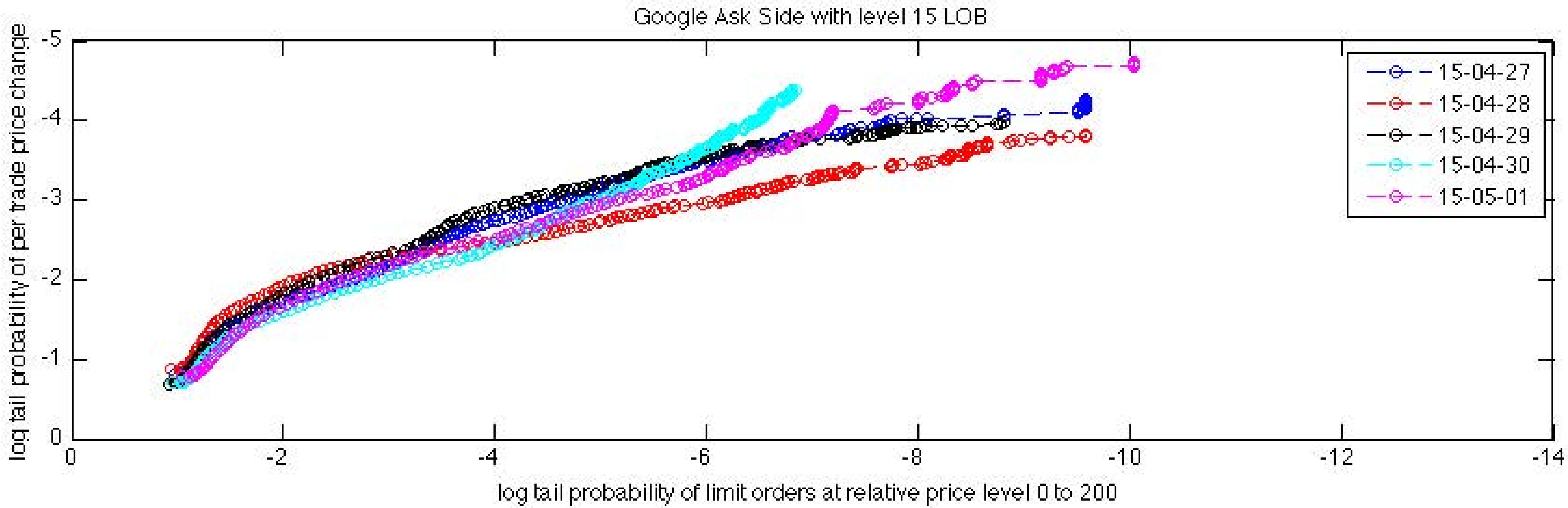
- **Step 1:** Fit model to **MID PRICE AND SPREAD SIZES** $<-$ **USING PRICE SERIES ONLY.**
- **Step 2:** Estimate $\bar{\Pi}_{BUY}(\cdot)$**.and** $\bar{\Pi}_{SELL}(\cdot)$ from **ORDER BOOK** $<-$ **USING ORDER BOOK ONLY.**
- **Step 3**: **Plot** $\log \bar{F}_{V_+}(x/S(t))$ **vs** $\log \Pi_{SELL}(x; A(t), B(t))$.& same for BUY side.

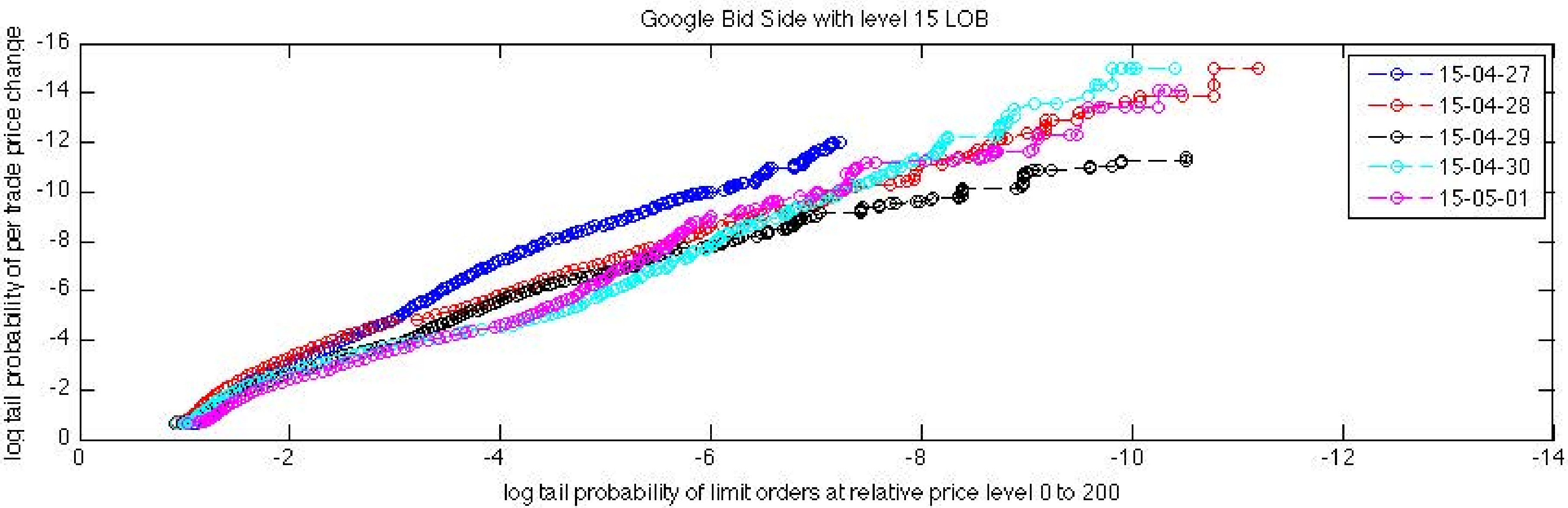
Empirical Validation Procedure

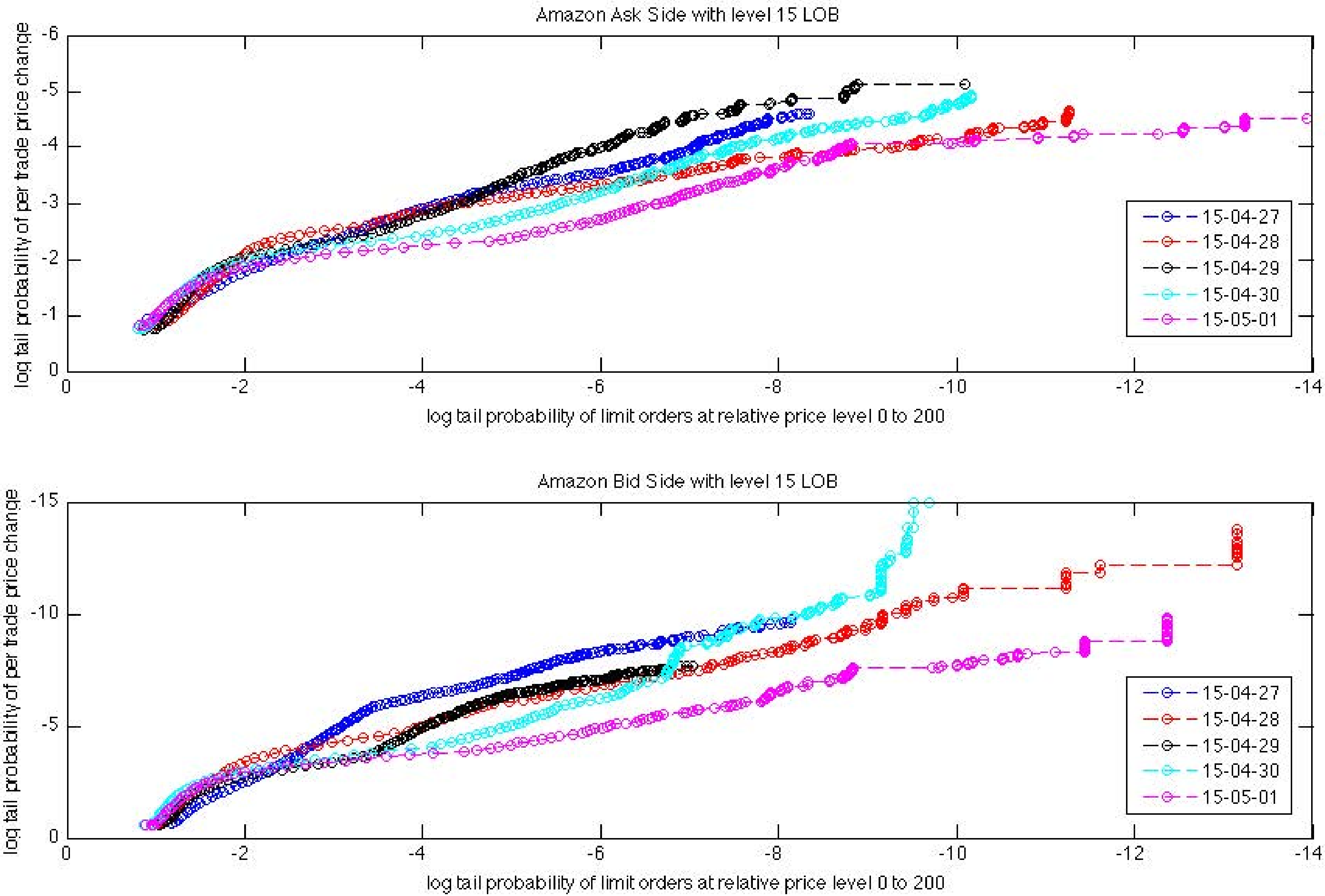
- **Step 1:** Fit model to **MID PRICE AND SPREAD SIZES** $<-$ **USING PRICE SERIES ONLY.**
- **Step 2:** Estimate $\bar{\Pi}_{BUY}(\cdot)$**.and** $\bar{\Pi}_{SELL}(\cdot)$ from **ORDER BOOK** $<-$ **USING ORDER BOOK ONLY.**
- **Step 3**: **Plot** $\log \bar{F}_{V_+}(x/S(t))$ **vs** $\log \Pi_{SELL}(x;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

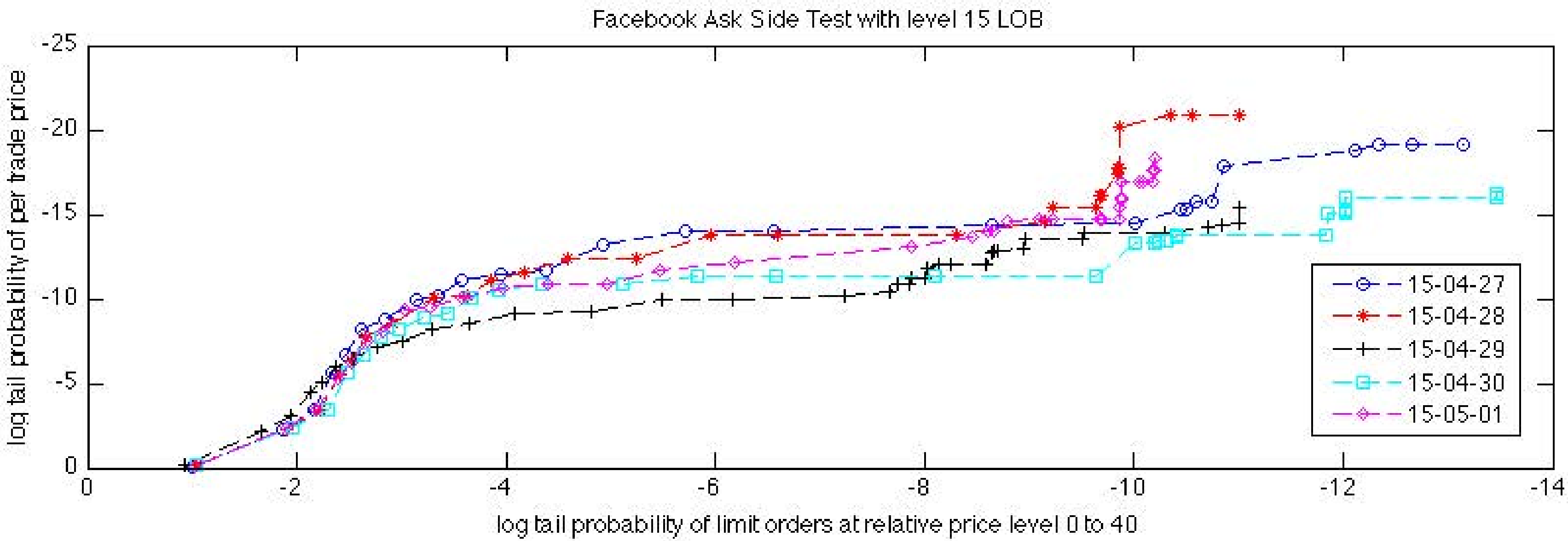
Some More Empirical Analysis

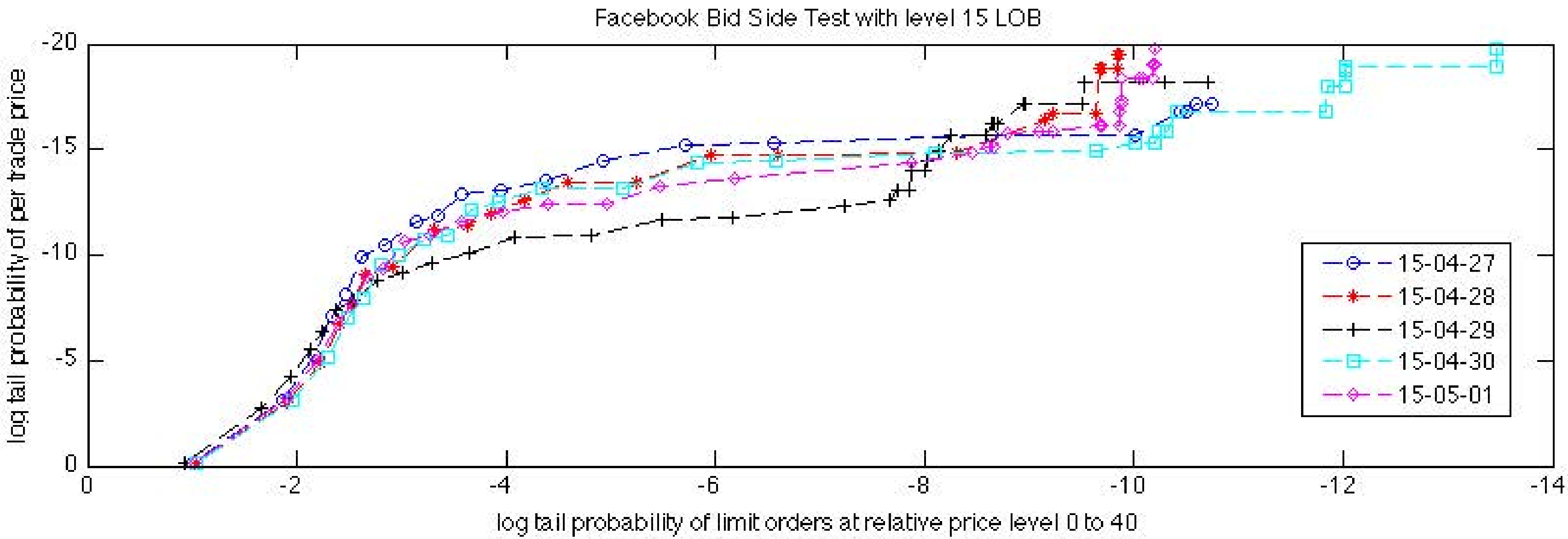
Pictures here...


Google Ask Side with level 15 LOB
log tail probability of per trade price change
log tail probability of limit orders at relative price level 0 to 200
-5
-4
-3
-2
-1
0
0
-2
-4
-6
-8
-10
-12
-14
15-04-27
15-04-28
15-04-29
15-04-30
15-05-01


Google Bid Side with level 15 LOB
log tail probability of per trade price change
log tail probability of limit orders at relative price level 0 to 200
-16
-14
-12
-10
-8
-6
-4
-2
0
0
-2
-4
-6
-8
-10
-12
-14
15-04-27
15-04-28
15-04-29
15-04-30
15-05-01


Amazon Ask Side with level 15 LOB
log tail probability of per trade price change
log tail probability of limit orders at relative price level 0 to 200
15-04-27
15-04-28
15-04-29
15-04-30
15-05-01
Amazon Bid Side with level 15 LOB
log tail probability of per trade price change
log tail probability of limit orders at relative price level 0 to 200
15-04-27
15-04-28
15-04-29
15-04-30
15-05-01


Facebook Ask Side Test with level 15 LOB
log tail probability of per trade price
log tail probability of limit orders at relative price level 0 to 40
-25
-20
-15
-10
-5
0
0
-2
-4
-6
-8
-10
-12
-14
15-04-27
15-04-28
15-04-29
15-04-30
15-05-01


Facebook Bid Side Test with level 15 LOB
log tail probability of per trade price
log tail probability of limit orders at relative price level 0 to 40
-20
-15
-10
-5
0
0
-2
-4
-6
-8
-10
-12
-14
15-04-27
15-04-28
15-04-29
15-04-30
15-05-01

Warning: Don't Use Model for Small Spread-Size Stocks

Avg. Spread Size (in $ cents)	4/27	4/28	4/29	4/30	5/1
Google	28.69	27.07	34.11	30.39	27.23
Facebook	1.43	1.41	1.68	1.42	1.45
Amazon	16.22	14.86	21.95	20.58	17.23

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

Price Formation via Queueing Microstructure

- Objective: Use queueing theory to explain price formation.

Price Formation via Queueing Microstructure

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

Price Formation via Queueing Microstructure

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

Price Formation via Queueing Microstructure

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

Price Formation via Queueing Microstructure

- 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 $= \lambda_n >> \mu_n =$ Arrival Market Orders.

Assumptions

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

Assumptions

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

Assumptions

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

Assumptions

1. Arrival Limit Orders $= \lambda_n >> \mu_n =$ Arrival Market Orders.
2. Times of market orders $\{t_k\}$ (Poisson –> can be relaxed).
3. Limit Orders between market orders at relative price $i\delta$ with $p_{BUY}\left(i\delta, \bar{A}\left(t_k\right), \bar{B}\left(t_k\right)\right) \cdot \delta$ (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 $\alpha_{BUY}\left(i\delta, \bar{A}\left(t_k\right), \bar{B}\left(t_k\right)\right)$ or $\alpha_{SELL}\left(i\delta, \bar{A}\left(t_k\right), \bar{B}\left(t_k\right)\right)$.

Discussion About Assumptions

1. Arrival Limit Orders $= \lambda_n >> \mu_n =$ Arrival Market Orders:

Google	4/27/2015	4/28/2015	4/29/2015	4/30/2015	5/1/2015
Total Limit Orders	80,537	78,944	77,215	100,798	66,238
Total Market Orders	9,412	8,016	5,868	8,505	7,038
Facebook					
Total Limit Orders	442,425	483,338	489,886	472,251	363,833
Total Market Orders	31,973	37,378	29,456	36,558	30,455
Amazon					
Total Limit Orders	100,263	131,648	125,555	162,561	123,127
Total Market Orders	13,148	14,804	9,225	11,344	10,094

Discussion About Assumptions

- **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:** $p_{BUY}\left(i\delta, \bar{A}\left(t_k\right), \bar{B}\left(t_k\right)\right) \cdot \delta$ 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:** $p_{BUY}\left(i\delta, \bar{A}\left(t_k\right), \bar{B}\left(t_k\right)\right) \cdot \delta$ prob. of placing order at $i\delta$, standard in literature
- **A4:** No cross over of limit orders to opposite side of mid-price.

Proportion of Cross Limit Orders	**4/27**	**4/28**	**4/29**	**4/30**	**5/1**
Google	2.12%	1.77%	2.31%	1.79%	1.84%
Facebook	2.82%	3.80%	4.77%	3.40%	3.01%
Amazon	2.55%	2.47%	2.45%	1.57%	1.94%

Discussion About Assumptions

- **A5:** $\alpha_{BUY}\left(i\delta, \bar{A}\left(t_k\right), \bar{B}\left(t_k\right)\right)$ 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

$$\rho_n(i) = \lambda_n \frac{p(i\delta, A(t_k), B(t_k))}{\alpha(i\delta, A(t_k), B(t_k))} \cdot \delta.$$

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

How to Pick a Cancellation Policy that Explains Empirical Findings?

- **Assumption (C):** For $x > x_0$

$$\alpha(x, A, B) \approx \gamma \times \left(1 - \bar{\Pi}(x, A, B)\right).$$

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**

$$\Pi\left(x; A, B\right)^{\gamma} \approx \theta\left(x, A, B\right),$$

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**

$$\Pi\left(x; A, B\right)^{\gamma} \approx \theta\left(x, A, B\right),$$

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 correspondance 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 correspondance between increment distribution & Limit Order Book structure.
- So, **auxiliary increment distribution:=** $\Delta_{k+1}(S(t_k))$ depending on spread:

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

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

$$\begin{aligned} \Delta_{k+1}(S(t_k)) &= (-1)^{R_{k+1}}(1 - I_{k+1})\,\delta_n \left[U_{k+1}/\left(n^{1/2}\delta_n\right)\right] \\ &\quad + I_{k+1}[S(t_k)\,V_k/\delta_n]\delta_n. \end{aligned}$$

Continuous Time Dynamics

- Given Assumption (C) there is a one to one correspondance 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

$$\begin{aligned}\Delta_{k+1}(S(t_k)) &= (-1)^{R_{k+1}}(1 - I_{k+1})\delta_n\left[U_{k+1}/\left(n^{1/2}\delta_n\right)\right] \\ &\quad + I_{k+1}[S(t_k)V_k/\delta_n]\delta_n.\end{aligned}$$

- Second piece captures the structural part, namely, the fact that for $x > x_0$

$$\Pi(x; A, B)^{\gamma} \approx \theta(x, A, B).$$

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 "$-$".

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

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

  similar for BUY with "$-$".

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

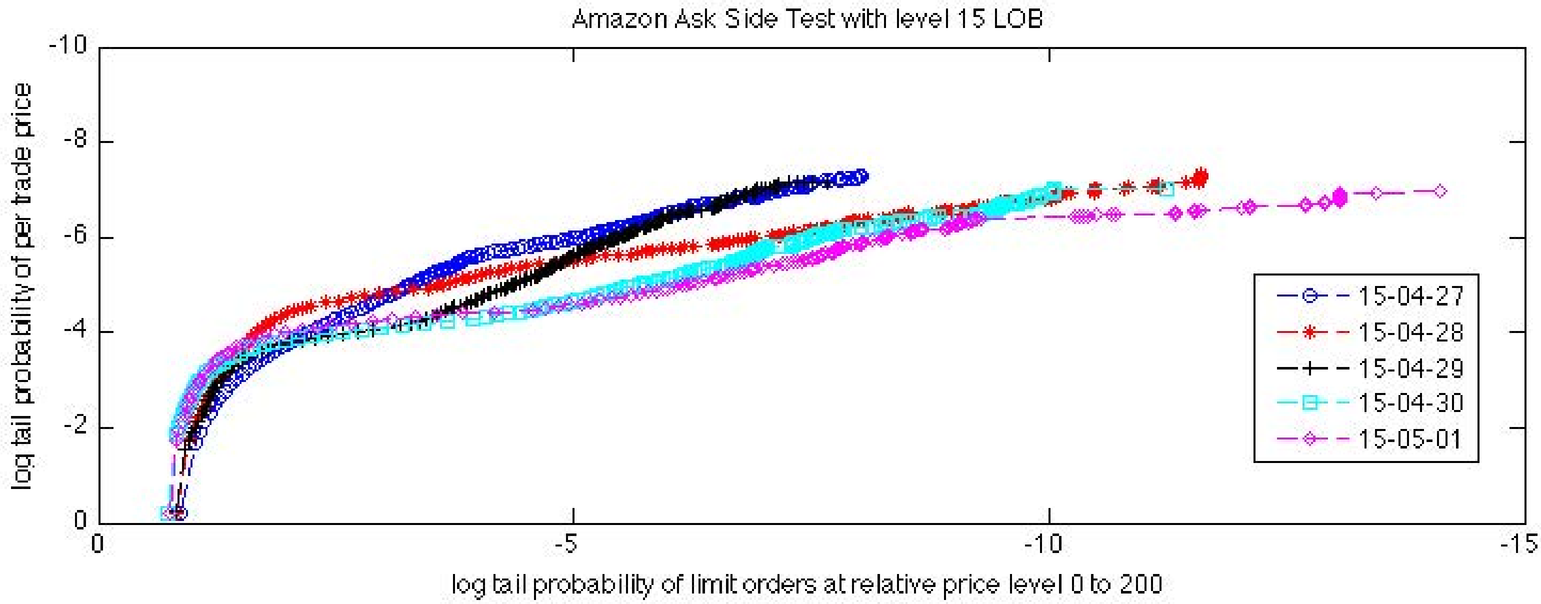
5 Conclusions

Conclusions

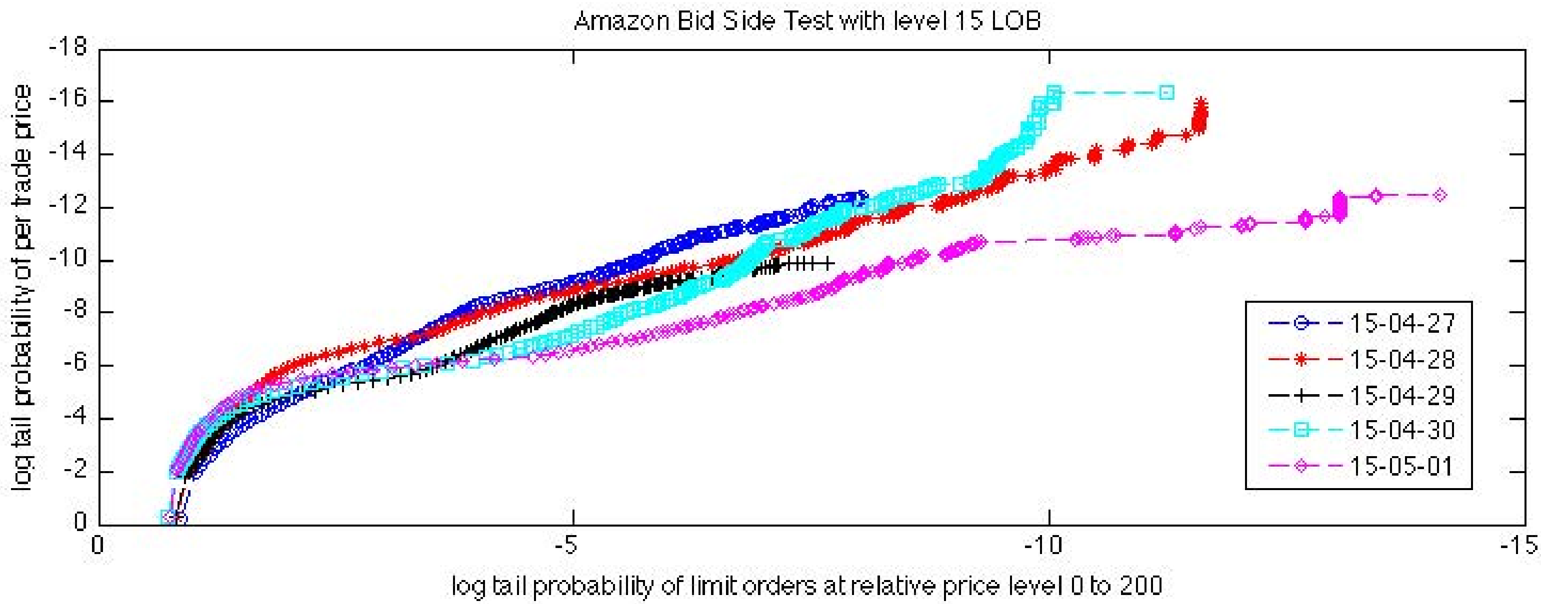
- **Presented and discussed a model for price and the bid-ask spread which:**
  a) Is informed by the full order book dynamics,
  b) It captures key stylized features observed empirically,
  c) Useful in intra-day trading (many minutes / few hours):

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

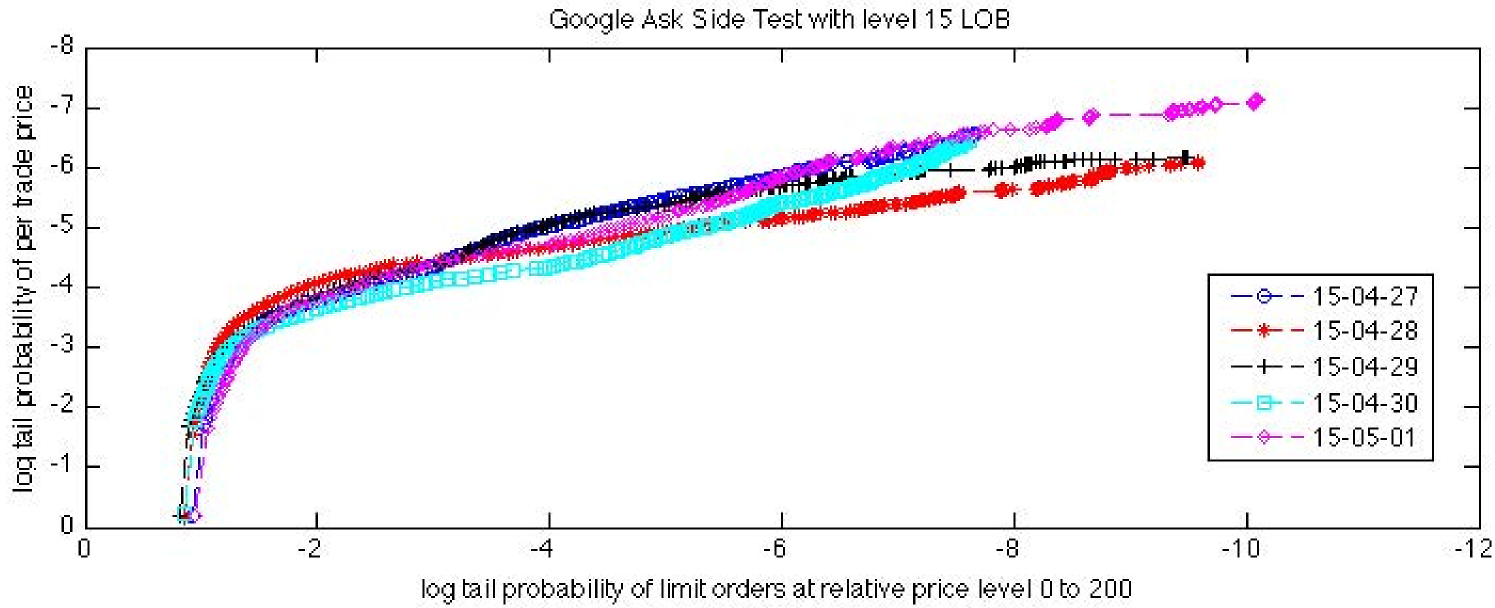
     Full order book directly feeds the dynamics of Bid-Ask processes.

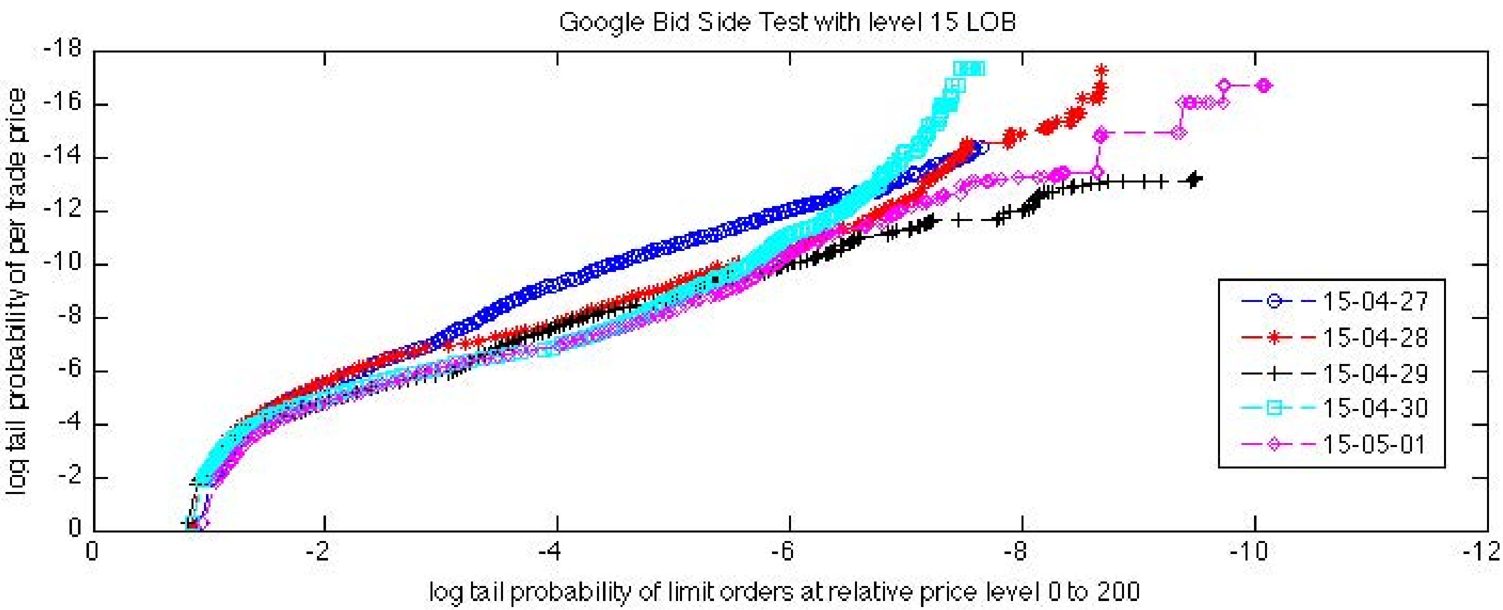


Amazon Ask Side Test with level 15 LOB
log tail probability of per trade price
log tail probability of limit orders at relative price level 0 to 200
-10
-8
-6
-4
-2
0
0
-5
-10
-15
15-04-27
15-04-28
15-04-29
15-04-30
15-05-01



Amazon Bid Side Test with level 15 LOB
log tail probability of per trade price
log tail probability of limit orders at relative price level 0 to 200
-18
-16
-14
-12
-10
-8
-6
-4
-2
0
0
-5
-10
-15
15-04-27
15-04-28
15-04-29
15-04-30
15-05-01


Google Ask Side Test with level 15 LOB
log tail probability of per trade price
log tail probability of limit orders at relative price level 0 to 200
15-04-27
15-04-28
15-04-29
15-04-30
15-05-01


Google Bid Side Test with level 15 LOB
log tail probability of per trade price
log tail probability of limit orders at relative price level 0 to 200
15-04-27
15-04-28
15-04-29
15-04-30
15-05-01