Catchy Title: Prices, Predictions, and Other Stuff

Boring Title: Stylized Facts and Models of the Limit Order Book

Ryan Donnelly

Department of Mathematics, King's College London

based on joint work with

Álvaro Cartea, University of Oxford
Sebastian Jaimungal, University of Toronto

April 2023

General Outline

- What is the Limit Order Book (LOB)?
- How do actions affect the LOB?
- Some statistical fun with order book events
- Model of order book signals and optimal order placement

What is a price?

- How much is your mobile phone worth to you?

What is a price?

- How much is your mobile phone worth to you?
- The buy value of an asset may be different from the sell value (subject to a successful experiment)

What is a price?

- How much is your mobile phone worth to you?
- The buy value of an asset may be different from the sell value (subject to a successful experiment)
- Buy and sell values are subjective, may be different for different individuals

What is a price?

- How much is your mobile phone worth to you?
- The buy value of an asset may be different from the sell value (subject to a successful experiment)
- Buy and sell values are subjective, may be different for different individuals
- If buy and sell values between two different agents coincide, ideally they would make a transaction

The Limit Order Book

- The LOB is a record of collective interest to buy or sell an asset
- A collection of standing orders to buy or sell certain amounts at certain prices
- Think of it as a collection of advertisements that are open for the taking
- When considering times and sizes at small enough scale, prices are not uniquely determined, and may not even exist at all

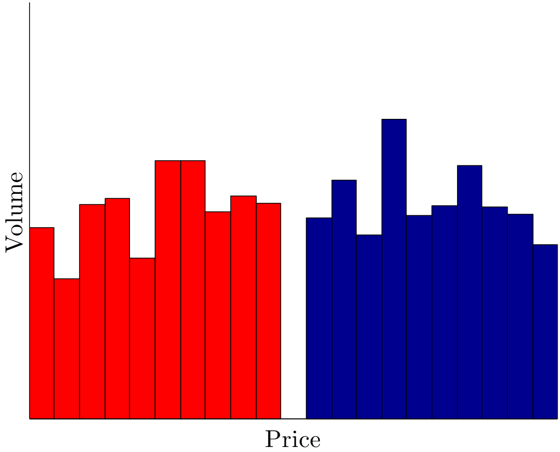
The Limit Order Book: fictitious example

Buy Orders		Sell Orders	
Price	Volume	Price	Volume
60.00	80	60.03	75
59.98	100	60.04	75
59.97	90	60.05	50
59.95	82	60.09	55
59.91	200	60.11	100
59.86	12	60.12	144
59.85	50	60.13	70
59.84	25	60.16	100

- Outstanding orders which compose the LOB are called **limit orders**
- Other market participants decide to match a standing limit order by submitting a **market order**

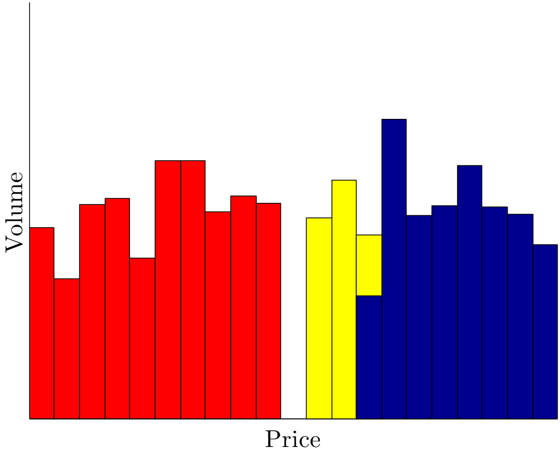
Graphical Representation of the LOB

- Each price has a volume of available shares



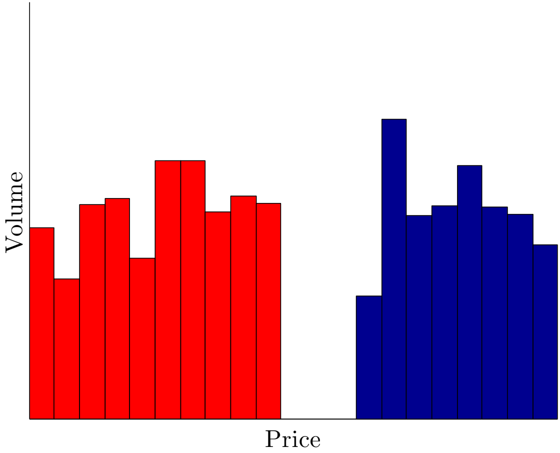
Arrival of a Market Order

- An incoming market order matches with limit orders



State of LOB after Market Order

- The market order lifts limit orders from the book



Queuing in the LOB

- The volume available at a single price could be a combination of several LOs

Queuing in the LOB

- The volume available at a single price could be a combination of several LOs
- This collection of orders form a queue, with those at the front transacting first with an arriving MO

Queuing in the LOB

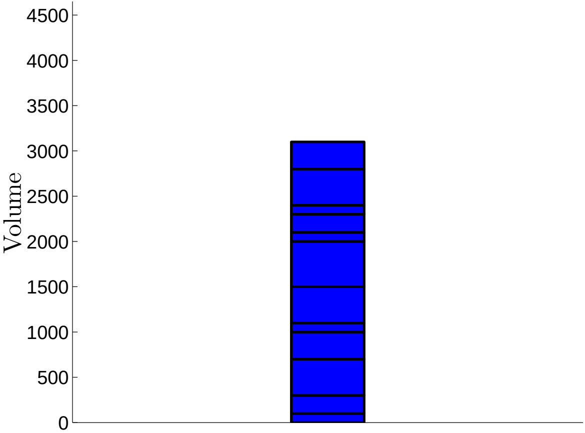
- The volume available at a single price could be a combination of several LOs
- This collection of orders form a queue, with those at the front transacting first with an arriving MO
- A newly placed order is located at the back of the queue

Queuing in the LOB

- The volume available at a single price could be a combination of several LOs
- This collection of orders form a queue, with those at the front transacting first with an arriving MO
- A newly placed order is located at the back of the queue
- As orders in front are canceled or executed, the position moves forward

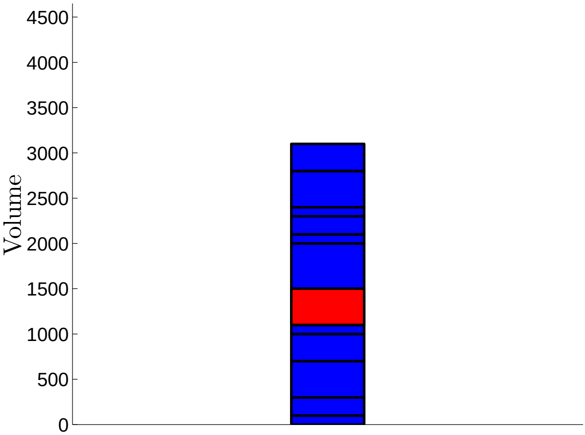
Queuing in the LOB

- Events alter size of queue and position of the agent order



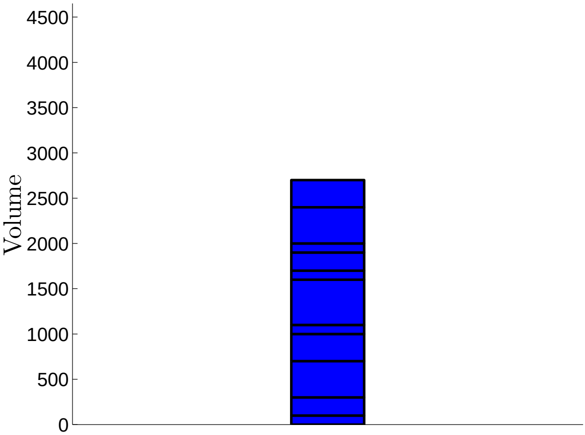
Queuing in the LOB

- Events alter size of queue and position of the agent order



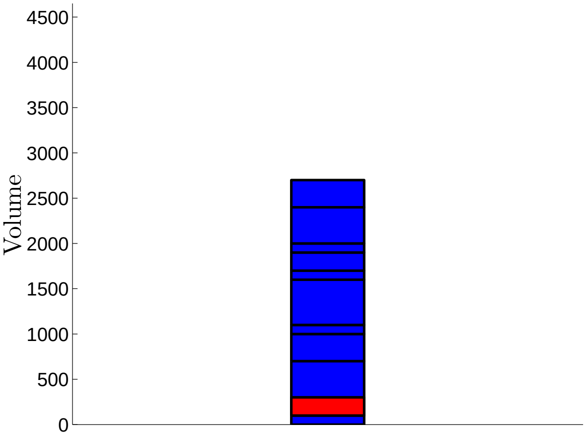
Queuing in the LOB

- Events alter size of queue and position of the agent order



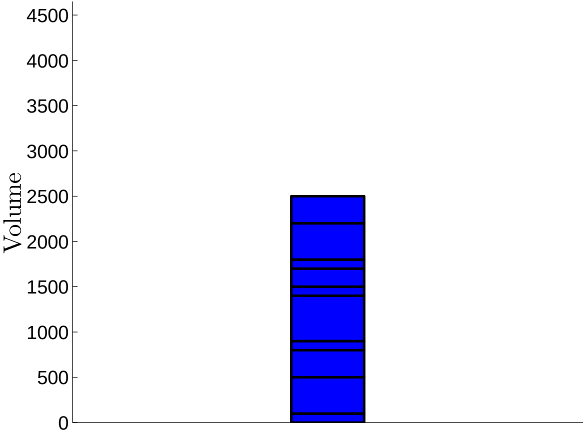
Queuing in the LOB

- Events alter size of queue and position of the agent order



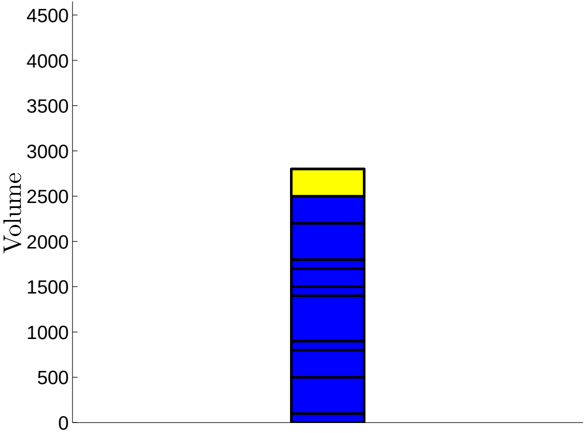
Queuing in the LOB

- Events alter size of queue and position of the agent order



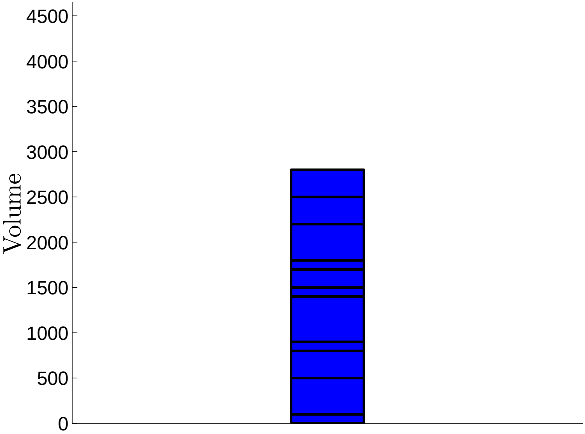
Queuing in the LOB

- Events alter size of queue and position of the agent order



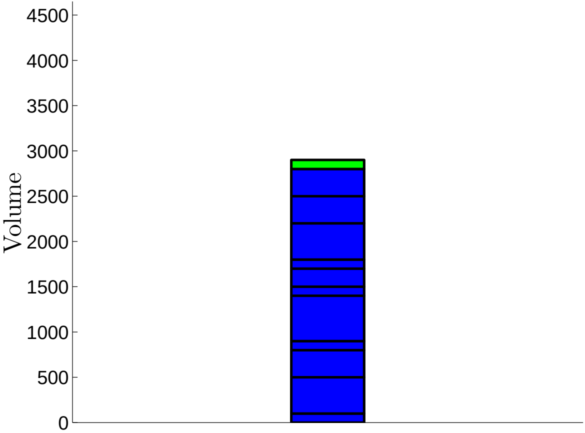
Queuing in the LOB

- Events alter size of queue and position of the agent order



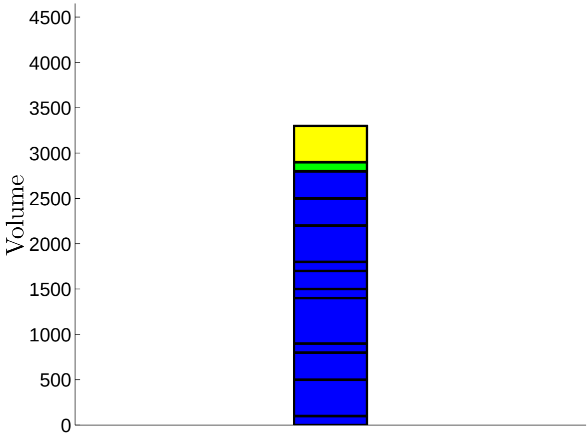
Queuing in the LOB

- Events alter size of queue and position of the agent order



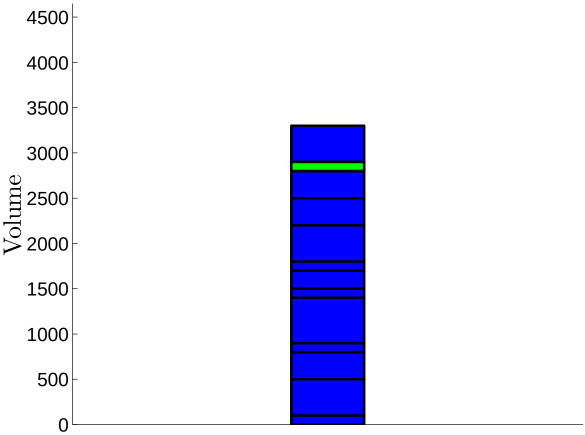
Queuing in the LOB

- Events alter size of queue and position of the agent order



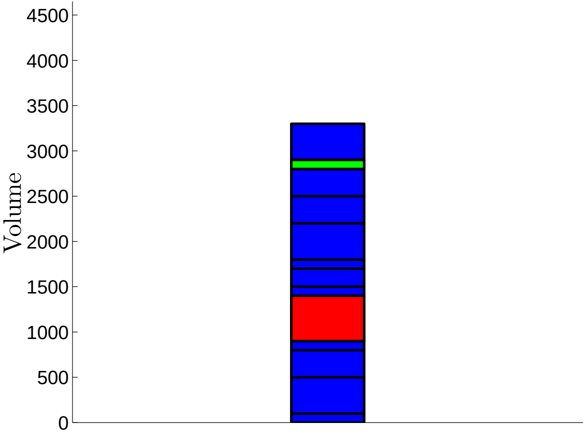
Queuing in the LOB

- Events alter size of queue and position of the agent order



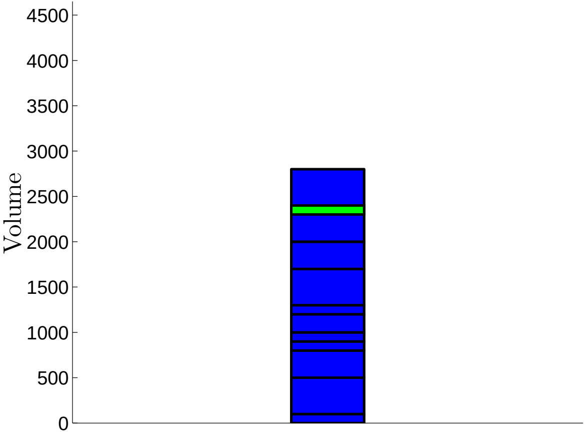
Queuing in the LOB

- Events alter size of queue and position of the agent order



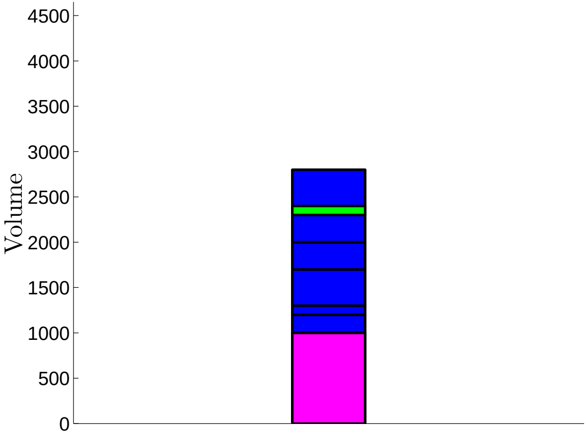
Queuing in the LOB

- Events alter size of queue and position of the agent order



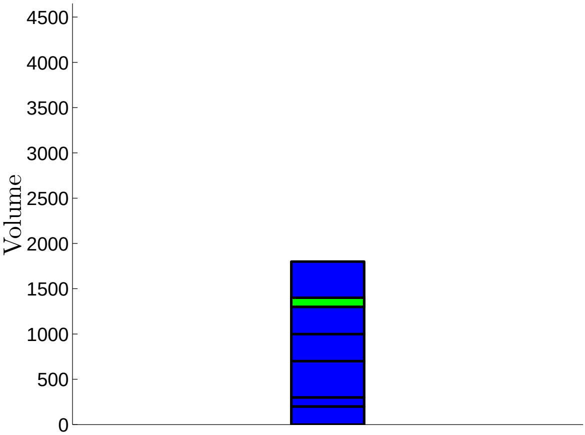
Queuing in the LOB

- Events alter size of queue and position of the agent order



Queuing in the LOB

- Events alter size of queue and position of the agent order



Advanced Order Types

- This summary is a significant oversimplification of a real LOB
- No mention of advanced order types:
  - stop-loss
  - stop-limit
  - immediate-or-cancel
  - fill-or-kill
  - all-or-none
- Each of these can be roughly classified as an aggressive or passive order type

Real LOB Snapshots

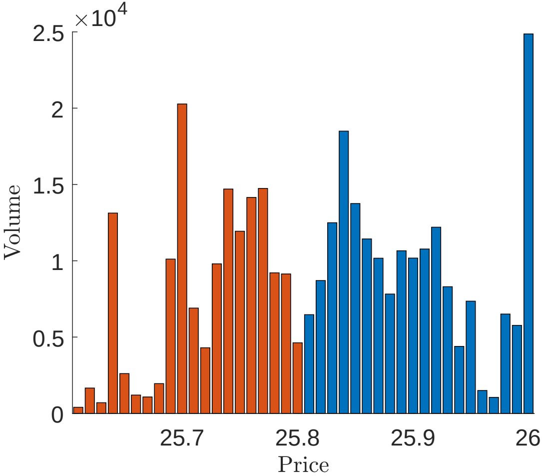


Figure: Snapshot of Nasdaq LOB for INTC on 01/04/2014 at 11:00am. Displayed price range is $0.40.

Real LOB Snapshots

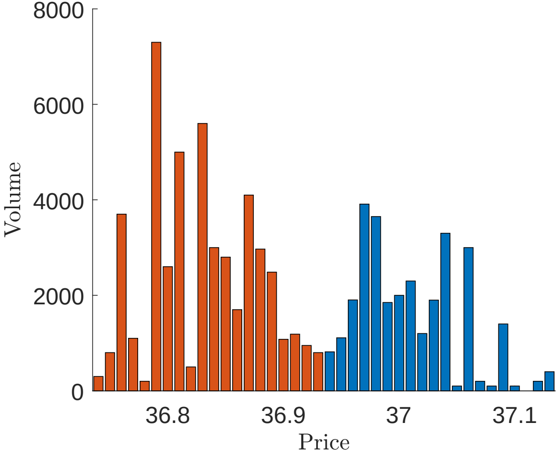


Figure: Snapshot of Nasdaq LOB for NTAP on 01/04/2014 at 11:00am. Displayed price range is $0.40.

KING'S College LONDON

Real LOB Snapshots

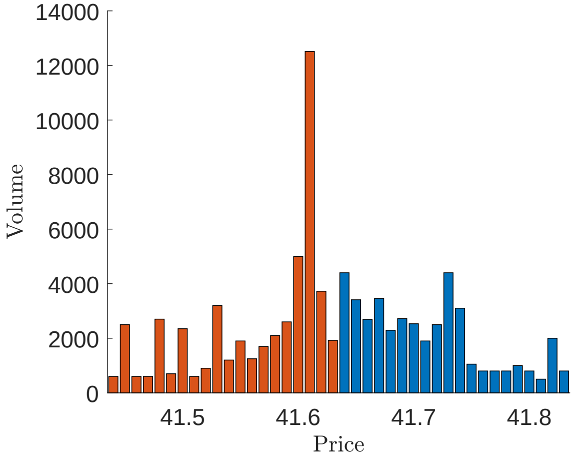


Figure: Snapshot of Nasdaq LOB for ORCL on 01/04/2014 at 11:00am. Displayed price range is $0.40.

Real LOB Snapshots

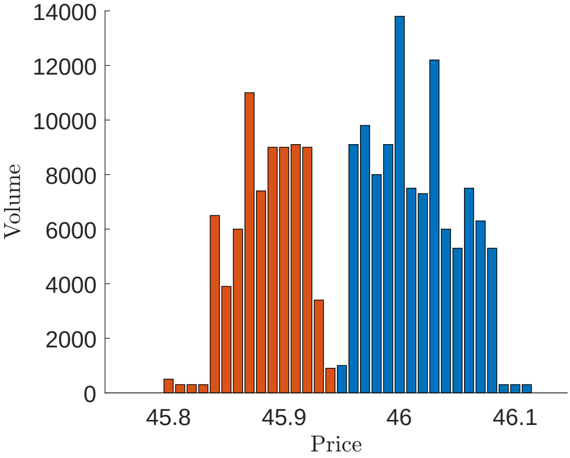


Figure: Snapshot of Nasdaq LOB for SMH on 01/04/2014 at 11:00am. Displayed price range is $0.40.

Real LOB Snapshots

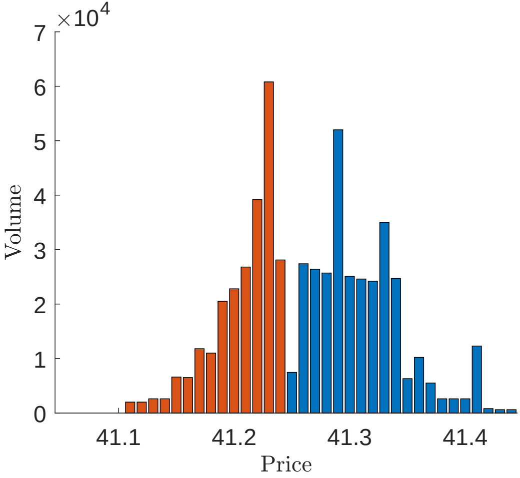


Figure: Snapshot of Nasdaq LOB for EEM on 01/04/2014 at 11:00am. Displayed price range is $0.40.

KING'S College LONDON

Real LOB Snapshots

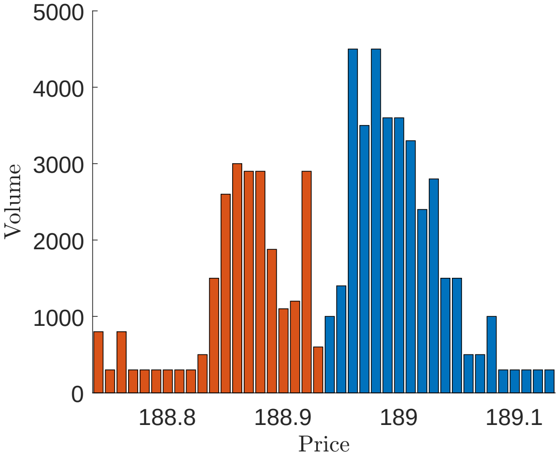


Figure: Snapshot of Nasdaq LOB for IVV on 01/04/2014 at 11:00am. Displayed price range is $0.40.

Real LOB Snapshots

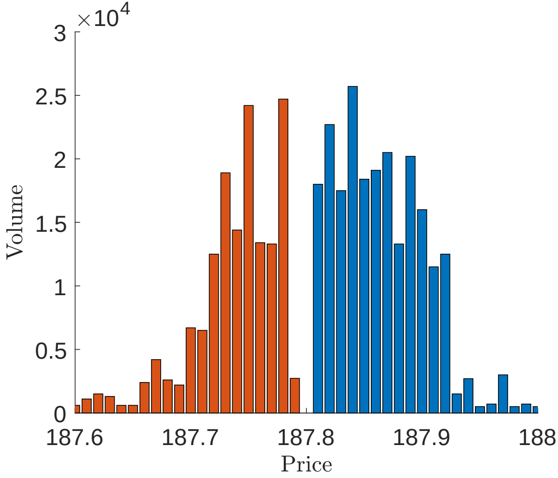


Figure: Snapshot of Nasdaq LOB for SPY on 01/04/2014 at 11:00am. Displayed price range is $0.40.

Real LOB Snapshots

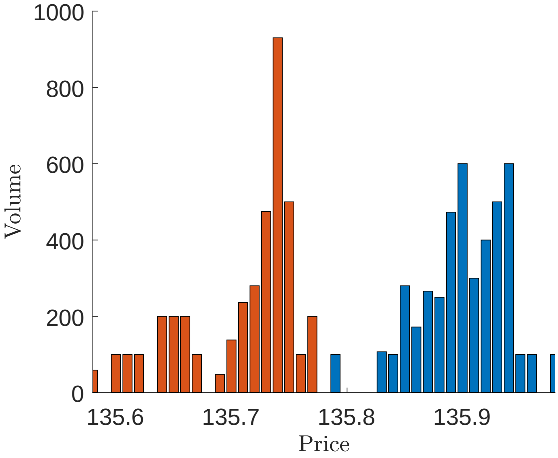


Figure: Snapshot of Nasdaq LOB for MMM on 01/04/2014 at 11:00am. Displayed price range is $0.40.

KING'S College LONDON

Real LOB Snapshots

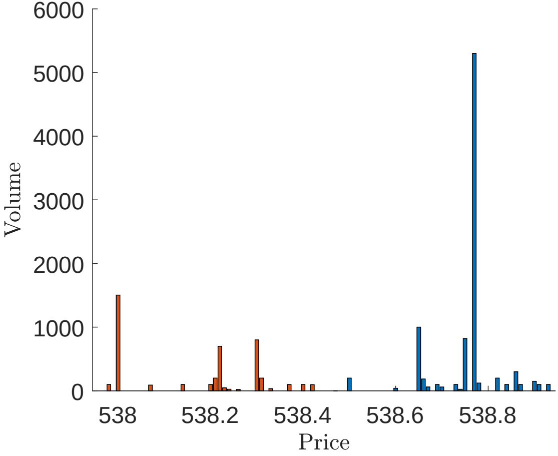


Figure: Snapshot of Nasdaq LOB for AAPL on 01/04/2014 at 11:00am. Displayed price range is $1.00.

Real LOB Snapshots

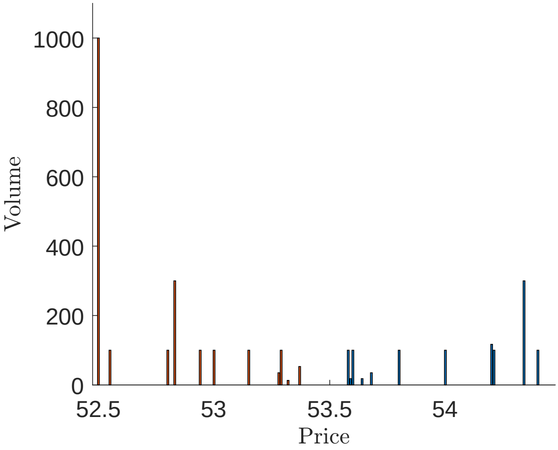


Figure: Snapshot of Nasdaq LOB for FARO on 01/04/2014 at 11:00am. Displayed price range is $2.00.

Real LOB Snapshots

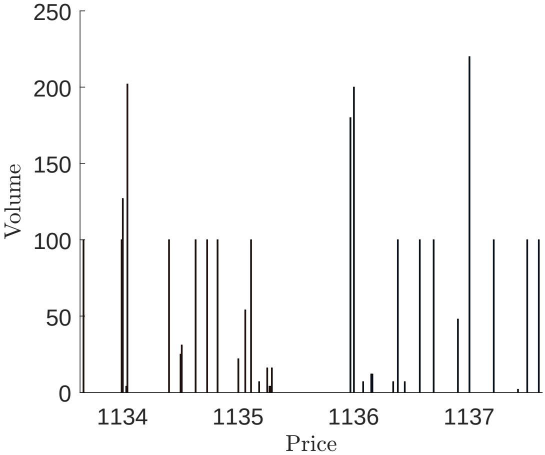


Figure: Snapshot of Nasdaq LOB for GOOG on 01/04/2014 at 11:00am. Displayed price range is $4.00.

Order Book Activity

- Every "event" in the LOB is logged
  - LO placement, LO cancellation, or MO

Order Book Activity

- Every "event" in the LOB is logged
  - LO placement, LO cancellation, or MO

	Number of Events	Market Orders	Percentage MO (%)
INTC			
NTAP			
ORCL			
SMH			
EEM			
IVV			
SPY			
MMM			
AAPL			
FARO			
GOOG			

Table: Nasdaq events on 1 April 2014 (first and last 30 minutes of each day removed). Percentage MO is the proportion of all events which are MOs.

Order Book Activity

- Every "event" in the LOB is logged
  - LO placement, LO cancellation, or MO

	Number of Events	Market Orders	Percentage MO (%)
SPY	1,611,668	10,439	0.65
IVV	700,388	1,540	0.22
ORCL	669,617	5,007	0.75
EEM	356,151	1,812	0.51
AAPL	269,849	6,224	2.30
INTC	269,208	2,106	0.78
NTAP	198,662	1,909	0.96
MMM	168,565	1,979	1.17
GOOG	121,519	2,534	2.09
SMH	107,583	514	0.48
FARO	23,802	144	0.60

Table: Nasdaq events on 1 April 2014 (first and last 30 minutes of each day removed). Percentage MO is the proportion of all events which are MOs.

KING'S College LONDON

Percentage of MOs that Walk the Book

- To "walk the book" is to execute a trade across multiple price levels

Percentage of MOs that Walk the Book

- To "walk the book" is to execute a trade across multiple price levels

	First Tick Only		Beyond First Tick		$\mathbb{P}(V_{MO} \leq V_{LO})$
	Buys	Sells	Buys	Sells	
AAPL	100,362	105,655	4,581	4,527	0.958
FARO	1,745	2,374	64	109	0.960
GOOG	32,096	34,969	3,085	3,075	0.916
INTC	35,595	38,451	54	50	0.999
MMM	22,996	25,745	130	118	0.995
NTAP	28,519	27,118	104	123	0.996
ORCL	30,001	27,502	41	45	0.999
SMH	3,087	3,084	7	4	0.998

Table: Nasdaq trades in January, 2014 (first and last 30 minutes of each day removed). $\mathbb{P}(V_{MO} \leq V_{LO})$ is probability that an MO has smaller volume than all limit orders posted at the best price.

LOB Activity Clustering

- Consider only the brief time intervals which occur immediately after an MO (up to 50ms)

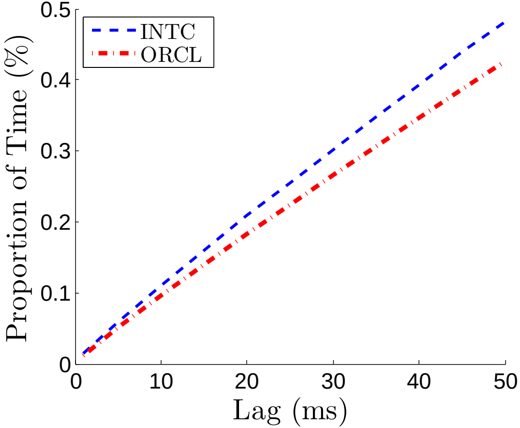


Figure: Percentage of trading day contained immediately after MO. Nasdaq trades in January, 2014.

LOB Activity Clustering - INTC

- What percentage of all LO activity occurs within these time intervals?

LOB Activity Clustering - INTC

- What percentage of all LO activity occurs within these time intervals?

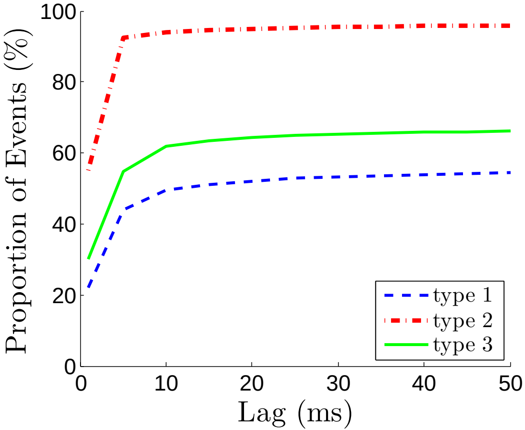


Figure: Percentage of LO activity occurring immediately after MO. Nasdaq trades for INTC in January, 2014.

LOB Activity Clustering - ORCL

- What percentage of all LO activity occurs within these time intervals?

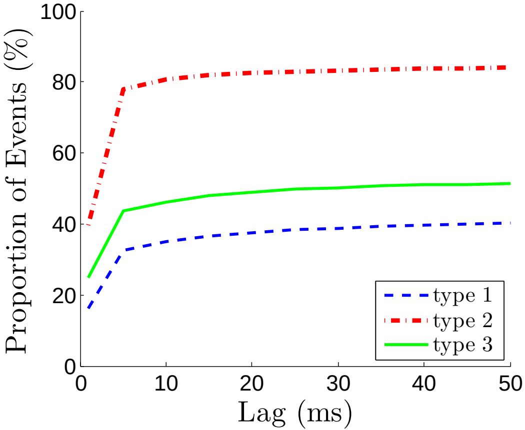


Figure: Percentage of LO activity occurring immediately after MO. Nasdaq trades for ORCL in January, 2014.

Volume Order Imbalance

- Denote by V_t^b and V_t^a the volume of orders at the best bid and ask prices at time t
- Define volume order imbalance by

$$I_t = \frac{V_t^b - V_t^a}{V_t^b + V_t^a}$$

- Then $I_t \in [-1, 1]$
- It measures the proportion of best interest on the bid side
- We look at events that occur when I_t is within the ranges $[-1, -\frac{1}{3})$, $[-\frac{1}{3}, \frac{1}{3}]$, and $(\frac{1}{3}, 1]$.

KING'S College LONDON

Trade Type vs. Imbalance - INTC

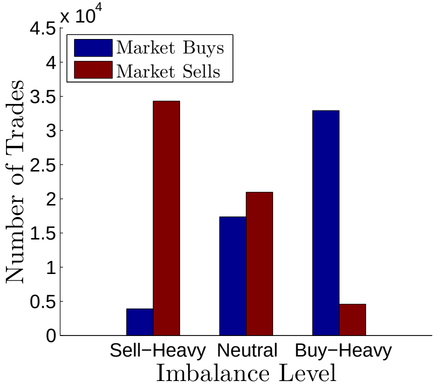


Figure: INTC: one month of NASDAQ trades. Imbalance ranges are $[-1, -\frac{1}{3})$, $[-\frac{1}{3}, \frac{1}{3}]$, and $(\frac{1}{3}, 1]$.

Trade Type vs. Imbalance - ORCL

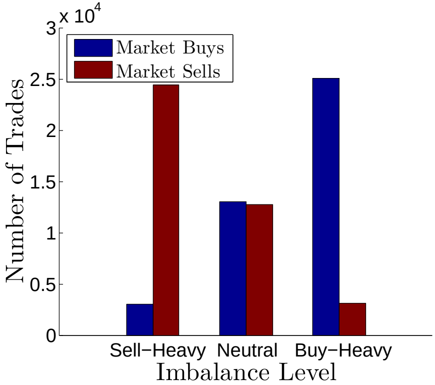


Figure: ORCL: one month of NASDAQ trades. Imbalance ranges are $[-1, -\frac{1}{3})$, $[-\frac{1}{3}, \frac{1}{3}]$, and $(\frac{1}{3}, 1]$.

Midprice Change vs. Imbalance - INTC

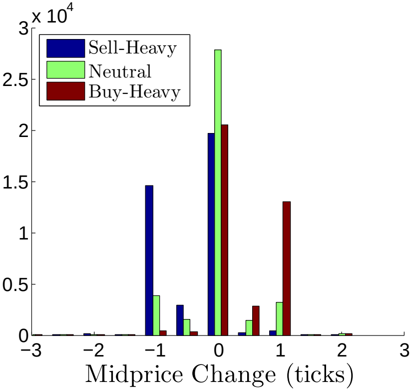


Figure: INTC: one month of NASDAQ trades. Imbalance ranges are $[-1, -\frac{1}{3})$, $[-\frac{1}{3}, \frac{1}{3}]$, and $(\frac{1}{3}, 1]$.

Midprice Change vs. Imbalance - INTC

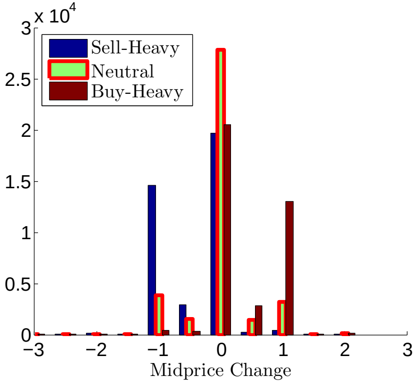


Figure: INTC: one month of NASDAQ trades. Imbalance ranges are $[-1, -\frac{1}{3})$, $[-\frac{1}{3}, \frac{1}{3}]$, and $(\frac{1}{3}, 1]$.

Midprice Change vs. Imbalance - INTC

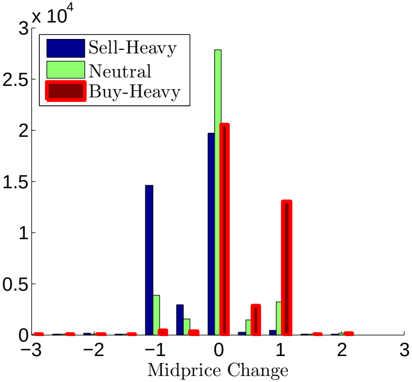


Figure: INTC: one month of NASDAQ trades. Imbalance ranges are $[-1, -\frac{1}{3})$, $[-\frac{1}{3}, \frac{1}{3}]$, and $(\frac{1}{3}, 1]$.

Midprice Change vs. Imbalance - INTC

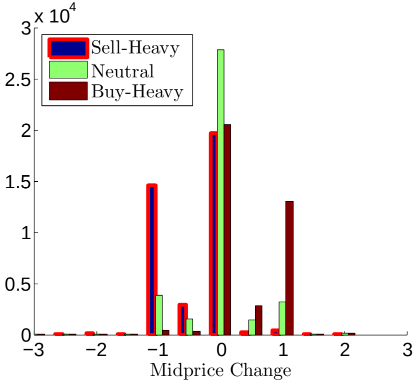


Figure: INTC: one month of NASDAQ trades. Imbalance ranges are $[-1, -\frac{1}{3})$, $[-\frac{1}{3}, \frac{1}{3}]$, and $(\frac{1}{3}, 1]$.

Midprice Change vs. Imbalance - ORCL

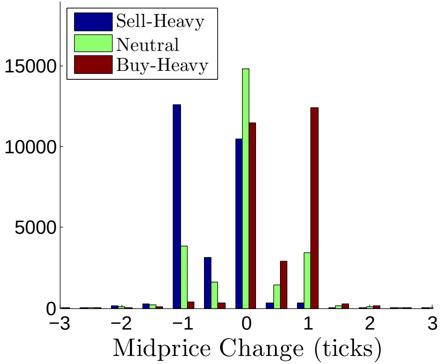


Figure: ORCL: one month of NASDAQ trades. Imbalance ranges are $[-1, -\frac{1}{3})$, $[-\frac{1}{3}, \frac{1}{3}]$, and $(\frac{1}{3}, 1]$.

Midprice Change vs. Imbalance - ORCL

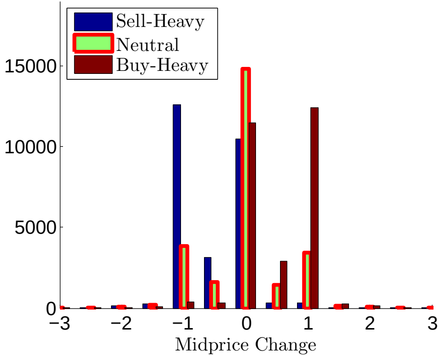


Figure: ORCL: one month of NASDAQ trades. Imbalance ranges are $[-1, -\frac{1}{3})$, $[-\frac{1}{3}, \frac{1}{3}]$, and $(\frac{1}{3}, 1]$.

Midprice Change vs. Imbalance - ORCL

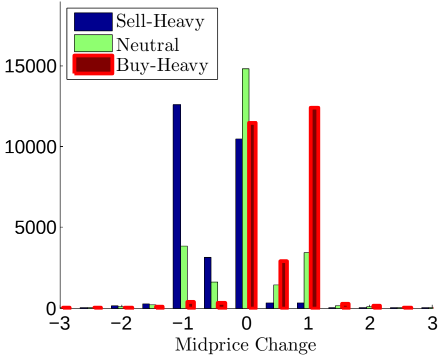


Figure: ORCL: one month of NASDAQ trades. Imbalance ranges are $[-1, -\frac{1}{3})$, $[-\frac{1}{3}, \frac{1}{3}]$, and $(\frac{1}{3}, 1]$.

Midprice Change vs. Imbalance - ORCL

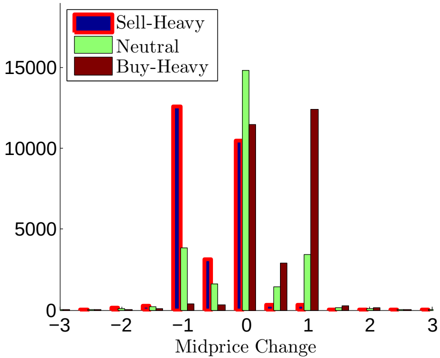


Figure: ORCL: one month of NASDAQ trades. Imbalance ranges are $[-1, -\frac{1}{3})$, $[-\frac{1}{3}, \frac{1}{3}]$, and $(\frac{1}{3}, 1]$.

Midprice Change after Buy vs. Imbalance - INTC

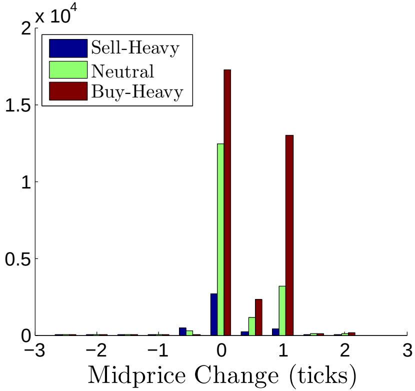


Figure: INTC: one month of NASDAQ trades. Imbalance ranges are $[-1, -\frac{1}{3})$, $[-\frac{1}{3}, \frac{1}{3}]$, and $(\frac{1}{3}, 1]$.

Midprice Change after Buy vs. Imbalance - INTC

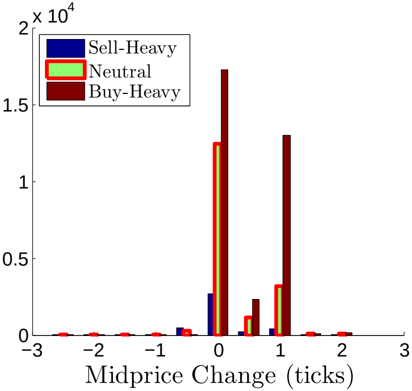


Figure: INTC: one month of NASDAQ trades. Imbalance ranges are $[-1, -\frac{1}{3})$, $[-\frac{1}{3}, \frac{1}{3}]$, and $(\frac{1}{3}, 1]$.

Midprice Change after Buy vs. Imbalance - INTC

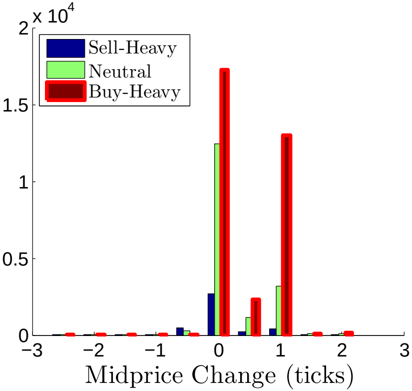


Figure: INTC: one month of NASDAQ trades. Imbalance ranges are $[-1, -\frac{1}{3})$, $[-\frac{1}{3}, \frac{1}{3}]$, and $(\frac{1}{3}, 1]$.

Midprice Change after Buy vs. Imbalance - INTC

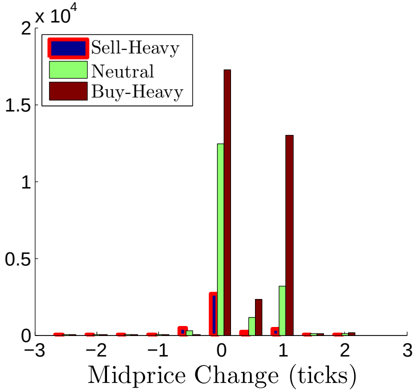


Figure: INTC: one month of NASDAQ trades. Imbalance ranges are $[-1, -\frac{1}{3})$, $[-\frac{1}{3}, \frac{1}{3}]$, and $(\frac{1}{3}, 1]$.

Midprice Change after Buy vs. Imbalance - ORCL

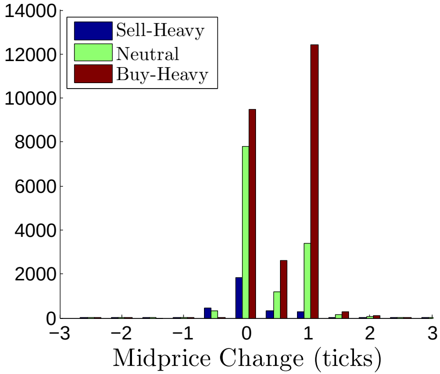


Figure: ORCL: one month of NASDAQ trades. Imbalance ranges are $[-1, -\frac{1}{3})$, $[-\frac{1}{3}, \frac{1}{3}]$, and $(\frac{1}{3}, 1]$.

Midprice Change after Buy vs. Imbalance - ORCL

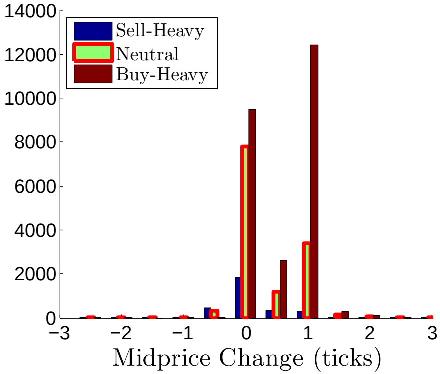


Figure: ORCL: one month of NASDAQ trades. Imbalance ranges are $[-1, -\frac{1}{3})$, $[-\frac{1}{3}, \frac{1}{3}]$, and $(\frac{1}{3}, 1]$.

Midprice Change after Buy vs. Imbalance - ORCL

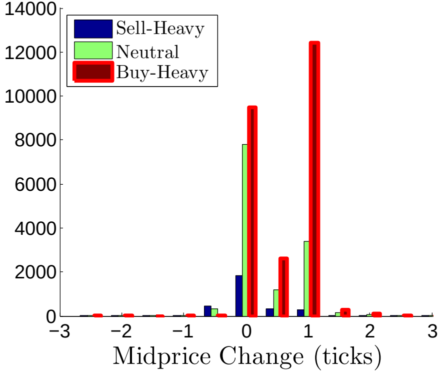


Figure: ORCL: one month of NASDAQ trades. Imbalance ranges are $[-1, -\frac{1}{3})$, $[-\frac{1}{3}, \frac{1}{3}]$, and $(\frac{1}{3}, 1]$.

Midprice Change after Buy vs. Imbalance - ORCL

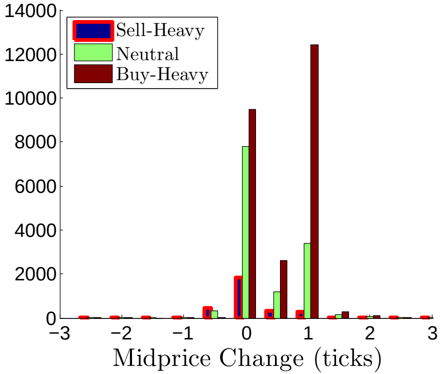


Figure: ORCL: one month of NASDAQ trades. Imbalance ranges are $[-1, -\frac{1}{3})$, $[-\frac{1}{3}, \frac{1}{3}]$, and $(\frac{1}{3}, 1]$.

Intraday Seasonality - MO Intensity - INTC

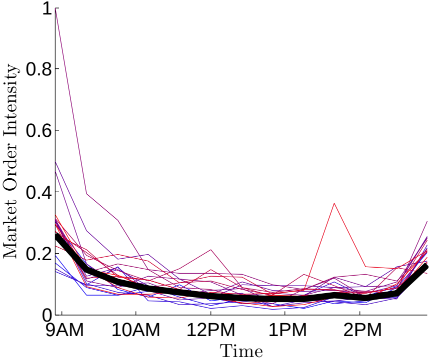


Figure: INTC: one month of NASDAQ trades with 30 minute intervals. Imbalance ranges are $[-1, -\frac{1}{3})$, $[-\frac{1}{3}, \frac{1}{3}]$, and $(\frac{1}{3}, 1]$.

Intraday Seasonality - MO Intensity - ORCL

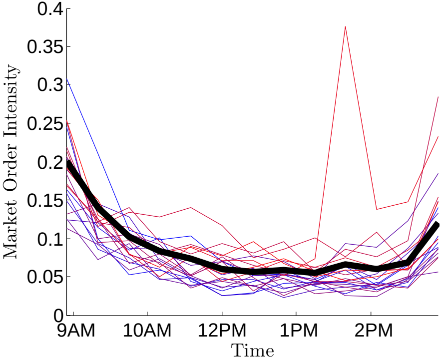


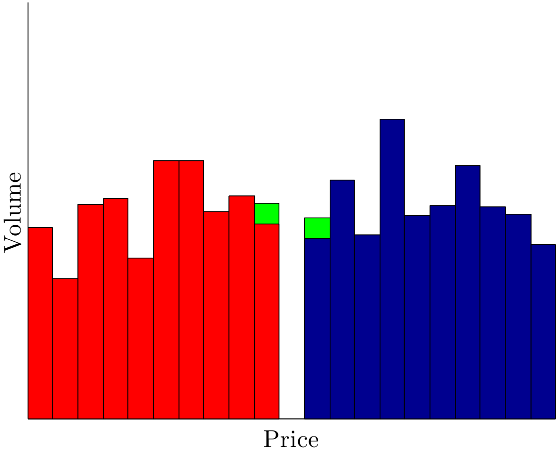
Figure: ORCL: one month of NASDAQ trades with 30 minute intervals. Imbalance ranges are $[-1, -\frac{1}{3})$, $[-\frac{1}{3}, \frac{1}{3}]$, and $(\frac{1}{3}, 1]$.

Now what?

- What do we do will all of this information?
  - Build a model where price dynamics accurately reflect tendencies observed in the data
  - Propose an optimization problem based on LO placement
  - Compare performance of the optimal strategy to naive LO placement

Our Objective

- Optimally place LOs at the best bid and ask



Models from Previous Literature

- Avellaneda and Stoikov (2008): midprice is BM, trades arrive according to Poisson process, exponential fill rate.
- Cartea and Jaimungal (2012): midprice jumps due to market orders, introduce risk control via inventory penalization.
- Fodra and Labadie (2012): midprice follows a diffusion process with general local drift and volatility terms, Poisson arrivals, exponential fill rate.
- Guilbaud and Pham (2013): discrete spread modelled as Markov chain, independent Levy process midprice, inventory penalization.
- Guéant, Lehalle, and Fernandez-Tapia (2013): midprice is BM, trades arrive according to Poisson process, exponential fill rate.
- Cartea, Jaimungal, and Ricci (2014): multi-factor mutually-exciting process jointly models arrivals, fill probabilities, and midprice drift.

Dynamic Model

- Let μ^l, μ^+, and μ^- be three doubly stochastic Poisson random measures.

- The midprice (S_t), imbalance regime (Z_t), spread (Δ_t), MO count $(M_t^\pm)$, inventory (q_t), and cash (X_t) are modelled by

$$
\begin{aligned}
M_t^\pm &= \int_0^t \int_{\bar{y}\in\mathbb{R}^3} \mu^\pm(d\bar{y}, du) \\
S_t &= S_0 + \int_0^t \int_{\bar{y}\in\mathbb{R}^3} y_1(\mu^l + \mu^+ - \mu^-)(d\bar{y}, du) \\
Z_t &= Z_0 + \int_0^t \int_{\bar{y}\in\mathbb{R}^3} (y_2 - Z_{u^-})(\mu^l + \mu^+ + \mu^-)(d\bar{y}, du) \\
\Delta_t &= \Delta_0 + \int_0^t \int_{\bar{y}\in\mathbb{R}^3} (y_3 - \Delta_{u^-})(\mu^l + \mu^+ + \mu^-)(d\bar{y}, du) \\
dX_t &= \gamma_t^+ (S_{t^-} + \frac{\Delta_{t^-}}{2}) dM_t^+ - \gamma_t^- (S_{t^-} - \frac{\Delta_{t^-}}{2}) dM_t^- \\
dq_t &= -\gamma_t^+ dM_t^+ + \gamma_t^- dM_t^-
\end{aligned}
$$

KING'S College LONDON

The Optimal Trading Problem

- The agent attempts to maximize expected terminal wealth, penalized by cumulative inventory position:

$$H(t, x, q, S, Z, \Delta) = \sup_{(\gamma_t^{\pm}) \in \mathcal{A}} \mathbb{E}\left[X_T + q_T \Big(S_T - \ell(q_T, \Delta_T) \Big) - \phi \int_t^T q_u^2 du \Big| \mathcal{F}_t \right]$$

- This value function has associated HJB equation:

$$\begin{aligned} &\partial_t H - \phi q^2 + \lambda^l(Z, \Delta)\mathbb{E}[\mathcal{D}^l H | Z, \Delta] \\ &\quad + \sup_{\gamma^+ \in \{0,1\}} \lambda^+(Z, \Delta)\mathbb{E}[\mathcal{D}^+ H | Z, \Delta] \\ &\quad + \sup_{\gamma^- \in \{0,1\}} \lambda^-(Z, \Delta)\mathbb{E}[\mathcal{D}^- H | Z, \Delta] = 0 \\ &\qquad\qquad H(T, x, q, S, Z) = x + q(S - \ell(q, \Delta)) \end{aligned}$$

Zero-Intelligence Performance

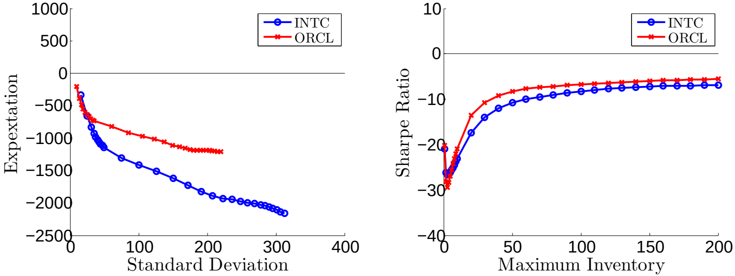
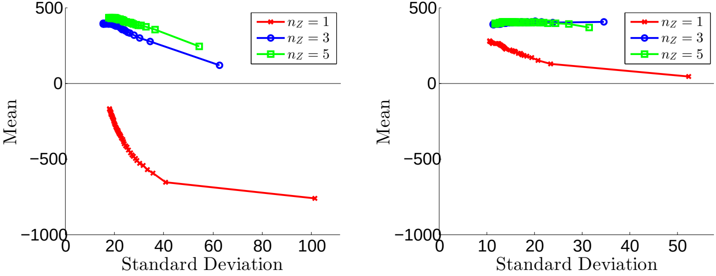


Figure: Naive strategy: annualized mean vs. standard deviation and annualized Sharpe ratio for various values of maximum inventory constraint from 1 to 200.

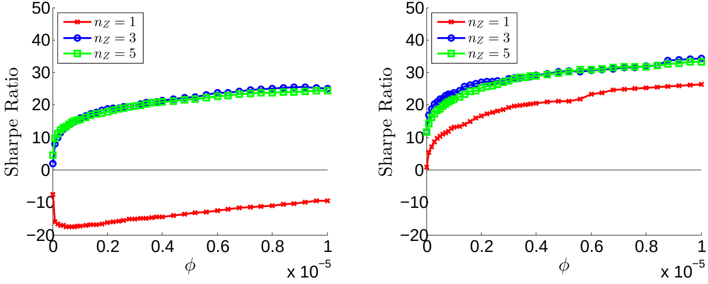
Performance of Historical Tests (INTC and ORCL)



(a) INTC (b) ORCL

Figure: Imbalance based strategy: annualized Sharpe ratio for difference numbers of observable imbalance states and various inventory penalizations ϕ.

Sharpe Ratio of Historical Tests (INTC and ORCL)



(a) INTC

(b) ORCL

Figure: Imbalance based strategy: annualized Sharpe ratio for difference numbers of observable imbalance states and various inventory penalizations ϕ.

What's next?

- There has been extensive work on using modern statistical and machine learning techniques to investigate LOB data
- Can the predictive signal of volume order imbalanced be improved?
  - Include volumes deeper in the LOB
  - Consider historical information in LOB
  - Include cross-asset information as a signal
- Can signals be constructed for interesting or useful features other than price changes?

KING'S College LONDON

Thanks for your attention!

Ryan Donnelly

ryan.f.donnelly@kcl.ac.uk

For more details see:

Cartea Á, R. Donnelly, S. Jaimungal (2018). Enhancing trading strategies with order bok signals. *Applied Mathematical Finance 25*(1), 1-35.