Algorithmic Trading

PIMS Summer School 2016

Order Imbalance[1]

Álvaro Cartea & Sebastian Jaimungal

University of Oxford & University of Toronto

July, 2016

[1]Based on "Enhancing Trading Strategies with Order Book Signals" Cartea, Donnelly, Jaimungal (2015)

Outline

- Volume order imbalance as an indicator of market behaviour.
- Imbalance model and market model.
- Optimal trading problem.
- Historical simulations.

Volume Order Imbalance

Volume Order Imbalance

- Volume order imbalance is the proportion of best interest on the bid side.
- Defined as:

$$\rho_t = \frac{V_t^b - V_t^a}{V_t^b + V_t^a}\,.$$

- V_t^b is the volume at the best bid at time t.
- V_t^a is the volume at the best ask at time t.
- $\rho_t \in [-1, 1]$.

Predictive Power of Volume Imbalance - MO type

- Consider the types of market orders that are placed depending on the level of imbalance.
- More market **buys** when imbalance is **high**, more market **sells** when imbalance is **low**.

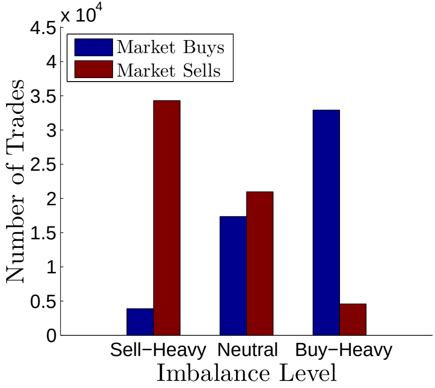


Figure : INTC: one month of NASDAQ trades. Imbalance ranges are $[-1,-0.33)$, $[-0.33, 0.33]$, and $(0.33, 1]$.

Predictive Power of Volume Imbalance - MO type

- Consider the types of market orders that are placed depending on the level of imbalance.
- More market **buys** when imbalance is **high**, more market **sells** when imbalance is **low**.

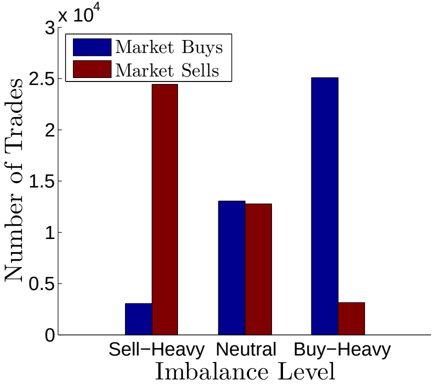


Figure : ORCL: one month of NASDAQ trades. Imbalance ranges are $[-1, -0.33)$, $[-0.33, 0.33]$, and $(0.33, 1]$.

Predictive Power of Volume Imbalance - Midprice Change

- Distribution of midprice change 10*ms* after a market order.

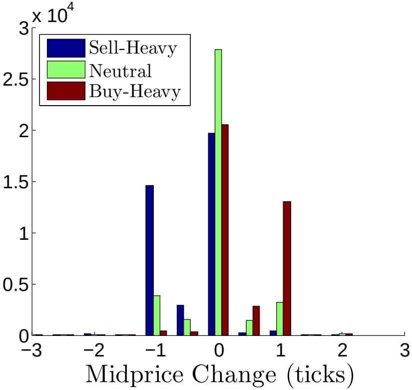


Figure : INTC: one month of NASDAQ trades. Imbalance ranges are $[-1, -0.33)$, $[-0.33, 0.33]$, and $(0.33, 1]$.

Predictive Power of Volume Imbalance - Midprice Change

- Distribution of midprice change 10*ms* after a market order.

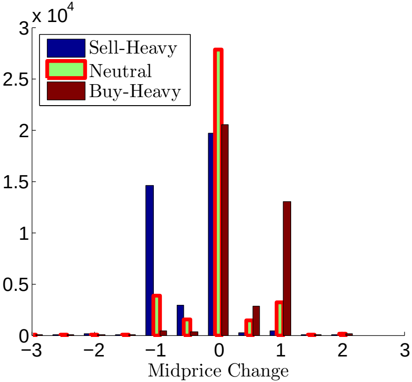


Figure : INTC: one month of NASDAQ trades. Imbalance ranges are $[-1, -0.33)$, $[-0.33, 0.33]$, and $(0.33, 1]$.

Predictive Power of Volume Imbalance - Midprice Change

- Distribution of midprice change 10*ms* after a market order.

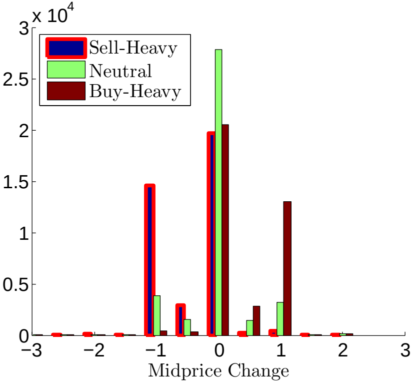


Figure : INTC: one month of NASDAQ trades. Imbalance ranges are $[-1, -0.33)$, $[-0.33, 0.33]$, and $(0.33, 1]$.

Predictive Power of Volume Imbalance - Midprice Change

- Distribution of midprice change 10*ms* after a market order.

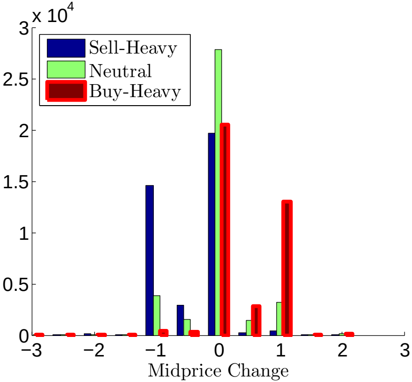


Figure : INTC: one month of NASDAQ trades. Imbalance ranges are $[-1, -0.33)$, $[-0.33, 0.33]$, and $(0.33, 1]$.

Predictive Power of Volume Imbalance - Midprice Change

- Distribution of midprice change 10*ms* after a market order.

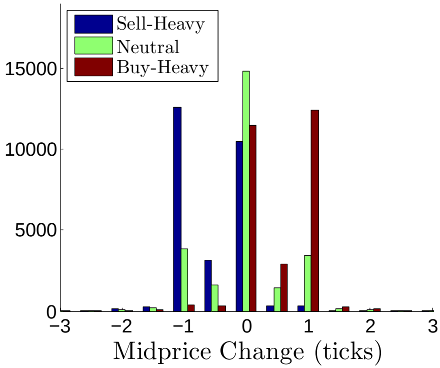


Figure : ORCL: one month of NASDAQ trades. Imbalance ranges are $[-1, -0.33)$, $[-0.33, 0.33]$, and $(0.33, 1]$.

Predictive Power of Volume Imbalance - Midprice Change

- Distribution of midprice change 10*ms* after a market order.

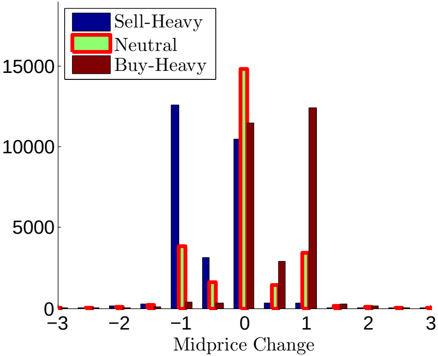


Figure : ORCL: one month of NASDAQ trades. Imbalance ranges are $[-1, -0.33)$, $[-0.33, 0.33]$, and $(0.33, 1]$.

Predictive Power of Volume Imbalance - Midprice Change

- Distribution of midprice change 10*ms* after a market order.

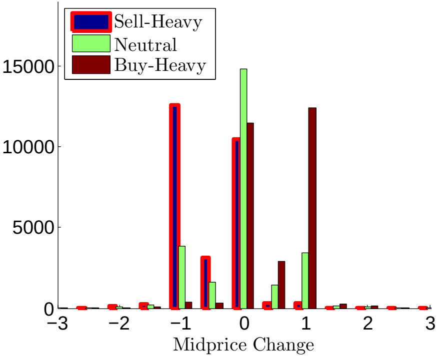


Figure : ORCL: one month of NASDAQ trades. Imbalance ranges are $[-1, -0.33)$, $[-0.33, 0.33]$, and $(0.33, 1]$.

Predictive Power of Volume Imbalance - Midprice Change

- Distribution of midprice change 10*ms* after a market order.

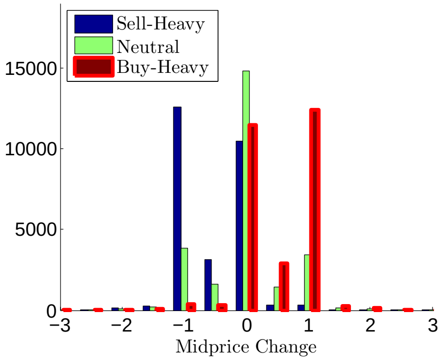


Figure : ORCL: one month of NASDAQ trades. Imbalance ranges are $[-1, -0.33)$, $[-0.33, 0.33]$, and $(0.33, 1]$.

Midprice Change After Buy Order

- Distribution of midprice change 10*ms* after a market buy order.

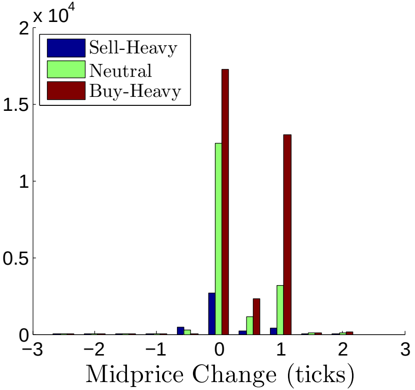


Figure : INTC: one month of NASDAQ trades. Imbalance ranges are $[-1, -0.33)$, $[-0.33, 0.33]$, and $(0.33, 1]$.

Midprice Change After Buy Order

- Distribution of midprice change 10*ms* after a market buy order.

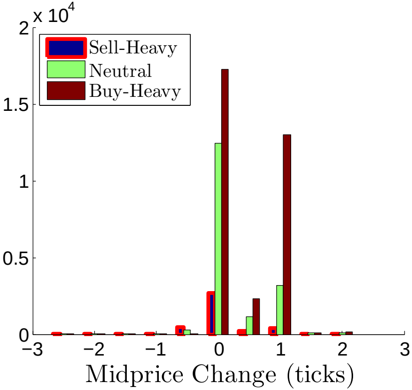


Figure : INTC: one month of NASDAQ trades. Imbalance ranges are $[-1, -0.33)$, $[-0.33, 0.33]$, and $(0.33, 1]$.

Midprice Change After Buy Order

- Distribution of midprice change 10*ms* after a market buy order.

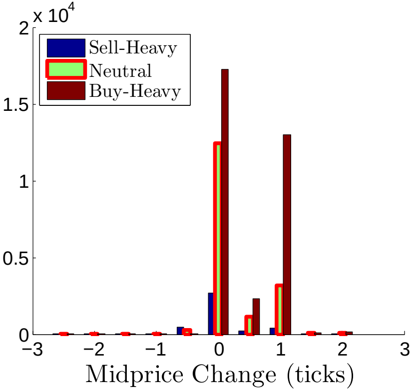


Figure : INTC: one month of NASDAQ trades. Imbalance ranges are $[-1, -0.33)$, $[-0.33, 0.33]$, and $(0.33, 1]$.

Midprice Change After Buy Order

- Distribution of midprice change 10*ms* after a market buy order.

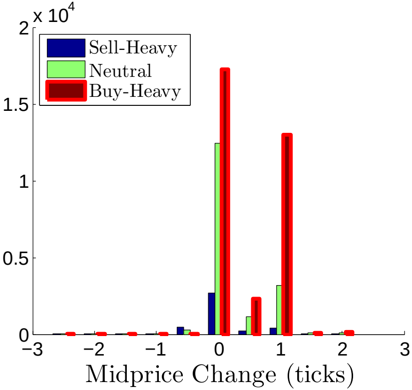


Figure : INTC: one month of NASDAQ trades. Imbalance ranges are $[-1, -0.33)$, $[-0.33, 0.33]$, and $(0.33, 1]$.

Tick Activity

	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 : Number of MOs that touch only the first tick or go beyond the first tick. Data taken from a full month of trading in January, 2014 (first and last 30 minutes of each day removed).

Market Model

Market Model

- Rather than model imbalance directly, a finite state imbalance regime process is considered, $Z_t \in \{1, \dots, n_Z\}$.
- Z_t will act as an approximation to the true value of imbalance.
- The interval $[-1, 1]$ is subdivided in to n_Z subintervals. $Z_t = k$ corresponds to ρ_t lying within the k^{th} subinterval.
- The spread, Δ_t, also takes values in a finite state space, $\Delta_t \in \{1, \dots, n_\Delta\}$.

Market Model

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

- M_t^+ and M_t^-, the number of market buy and sell orders up to time t, are given by:

$$M_t^{\pm} = \int_0^t \int_{\bar{y} \in \mathbb{R}^3} \mu^{\pm}(d\bar{y}, du)$$

- The midprice, S_t, together with Z_t and Δ_t are modelled as:

$$S_t = S_0 + \int_0^t \int_{\bar{y} \in \mathbb{R}^3} y_1 (\mu^l + \mu^+ - \mu^-)(d\bar{y}, du) \,,$$

$$Z_t = Z_0 + \int_0^t \int_{\bar{y} \in \mathbb{R}^3} (y_2 - Z_{u^-})(\mu^l + \mu^+ + \mu^-)(d\bar{y}, du) \,,$$

$$\Delta_t = \Delta_0 + \int_0^t \int_{\bar{y} \in \mathbb{R}^3} (y_3 - \Delta_{u^-})(\mu^l + \mu^+ + \mu^-)(d\bar{y}, du) \,.$$

Main features of this model

- All three μ^i are conditionally independent given (Z_t, Δ_t) and have compensators of the form:

$$\nu^i(d\bar{y}, dt) = \lambda^i(Z_t, \Delta_t) F^i_{Z_t, \Delta_t}(d\bar{y}) dt$$

- This makes the joint process (Z_t, Δ_t) a continuous time Markov chain.
- $\lambda^{\pm}(Z, \Delta)$ and $F^{\pm}_{Z,\Delta}(d\bar{y})$ are chosen to reflect realistic dependence of market order arrivals and jumps after market orders on imbalance and spread.
- $F^{l}_{Z,\Delta}$ is chosen to have support only on $y_1 = \pm\frac{y_3 - \Delta}{2}$. Limit order activity must change the midprice and spread simultaneously.

Agent's Wealth and Inventory

- The agent may post bid and ask orders at the touch.
- Wealth has dynamics:

$$dX_t = \gamma_t^+ \left(S_{t^-} + \frac{\Delta_{t^-}}{2} \right) dM_t^+ - \gamma_t^- \left(S_{t^-} - \frac{\Delta_{t^-}}{2} \right) dM_t^- \,,$$

  where $\gamma_t^\pm \in \{0, 1\}$ are the agent's control processes.

- Inventory has dynamics:

$$dq_t = -\gamma_t^+ \, dM_t^+ + \gamma_t^- \, dM_t^- \,.$$

- Controls $\gamma_t^\pm$ are chosen such that inventory is constrained, $\underline{Q} \leq q_t \leq \overline{Q}$:

Optimal Trading

The Optimal Trading Problem

- The agent attempts to maximise 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 equation:

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

Value Function Ansatz

- Making the ansatz $H(t,x,q,S,Z,\Delta) = x + qS + h(t,q,Z,\Delta)$ allows for a corresponding equation for h to be written:

$$
\begin{aligned}
&\partial_t h - \phi q^2 + \lambda^l(Z,\Delta)(q\epsilon^l(Z,\Delta) + \Sigma^l(t,q,Z,\Delta)) \\
&+ \sup_{\gamma^+\in\{0,1\}} \lambda^+(Z,\Delta)\left(\gamma^+\frac{\Delta}{2} + (q-\gamma^+)\epsilon^+(Z,\Delta) + \Sigma^+_{\gamma^+}(t,q,Z,\Delta)\right) \\
&+ \sup_{\gamma^-\in\{0,1\}} \lambda^-(Z,\Delta)\left(\gamma^-\frac{\Delta}{2} - (q+\gamma^-)\epsilon^-(Z,\Delta) + \Sigma^-_{\gamma^-}(t,q,Z,\Delta)\right) = 0 \\
&h(T,q,Z,\Delta) = -q\ell(q,\Delta)
\end{aligned}
$$

- This is a system of ODE's of dimension $n_Z n_\Delta(\overline{Q} - \underline{Q} + 1)$.

Feedback Controls

- Feedback controls can be written as:

$$\gamma^{\pm}(t,q,Z,\Delta) = \begin{cases} 1, & \frac{\Delta}{2} - \epsilon^{\pm}(Z,\Delta) + \Sigma_1^{\pm}(t,q,Z,\Delta) > \Sigma_0^{\pm}(t,q,Z,\Delta) \\ 0, & \frac{\Delta}{2} - \epsilon^{\pm}(Z,\Delta) + \Sigma_1^{\pm}(t,q,Z,\Delta) \leq \Sigma_0^{\pm}(t,q,Z,\Delta) \end{cases}$$

where

$$\epsilon^{\pm}(Z,\Delta) = \sum_{y_1,y_2,y_3} y_1 F^{\pm}_{Z,\Delta}(y_1,y_2,y_3)$$

$$\Sigma^{\pm}_{\gamma^{\pm}}(t,q,Z,\Delta) = \sum_{y_1,y_2,y_3} (h(t,q \mp \gamma^{\pm}, y_2, y_3) - h(t,q,Z,\Delta))F^{\pm}_{Z,\Delta}(y_1,y_2,y_3)$$

Optimal Trading Strategy – Parameters

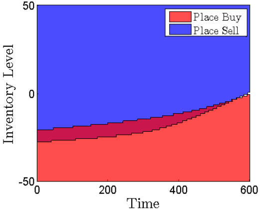
- Allow three possible states of imbalance: $Z_t \in \{1, 2, 3\}$
- Two possible spreads: $\Delta_t \in \{1, 2\}$
- MO arrival rates and price impact account for imbalance:

$$\overline{\lambda}^{+} = \begin{pmatrix} 0.050 & 0.091 & 0.242 \\ 0.057 & 0.051 & 0.095 \end{pmatrix} \qquad \overline{\varepsilon}^{+} = \begin{pmatrix} 0.247 & 0.556 & 0.710 \\ 0.539 & 0.959 & 1.036 \end{pmatrix}$$

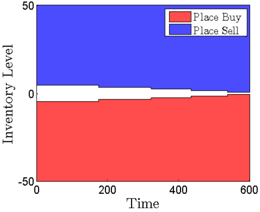
$$\overline{\lambda}^{-} = \begin{pmatrix} 0.242 & 0.091 & 0.050 \\ 0.095 & 0.051 & 0.057 \end{pmatrix} \qquad \overline{\varepsilon}^{-} = \begin{pmatrix} 0.710 & 0.556 & 0.247 \\ 1.036 & 0.959 & 0.539 \end{pmatrix}$$

- Terminal penalty function chosen to be $\ell(q) = sgn(q)\frac{\Delta}{2}$.

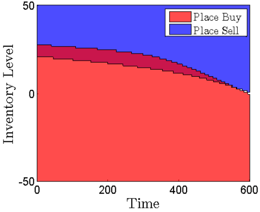
Optimal Trading Strategy



Low Imbalance

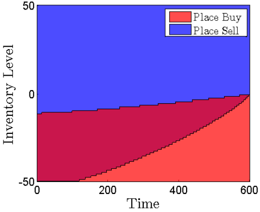


Middle Imbalance

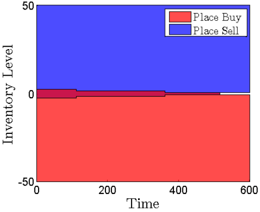


High Imbalance

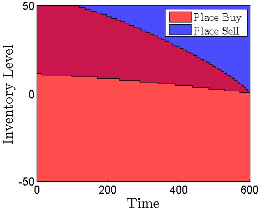
$\Delta = 1$



Low Imbalance



Middle Imbalance



High Imbalance

$\Delta = 2$

Historical Simulations

The Value of Knowing Imbalance

- The number of imbalance regimes is an important modelling choice.
- A large number of regimes can begin to cause observation and parameter estimation problems.
- A small number of regimes will not benefit as much from the predictive information.
- How does the performance of an agent depend on the number imbalance regimes in the model?

Historical Simulations

- We analyze the performance of the strategy tested on historical data.
- The strategy is executed based on 1, 3, and 5 different states of imbalance.
- We compare to a zero-intelligence strategy which consists of always posting limit orders at the best bid and ask, regardless of the state of the limit order book.
- Data consists of all trading days from July to December 2014 divided into 30 minute intervals. The first and last interval of each day are excluded.

Parameter Forecasting

- The historical simulations are performed out-of-sample.
- We employ a simple method of forecasting model parameters based on intraday seasonality.

$$\boldsymbol{\lambda}^i_{m,n} = \alpha_n^{\lambda^i} + \beta_n^{\lambda^i} \boldsymbol{\lambda}^i_{m,n-1}$$
$$\boldsymbol{\epsilon}^i_{m,n} = \alpha_n^{\epsilon^i} + \beta_n^{\epsilon^i} \boldsymbol{\epsilon}^i_{m,n-1}$$

- Factor loadings are obtained by regression using data from January to June 2014.
- The improvement in performance over the naive strategy is substantial, and a more elegant forecasting method would likely give further improvements.

Seasonality - Market Order Intensity

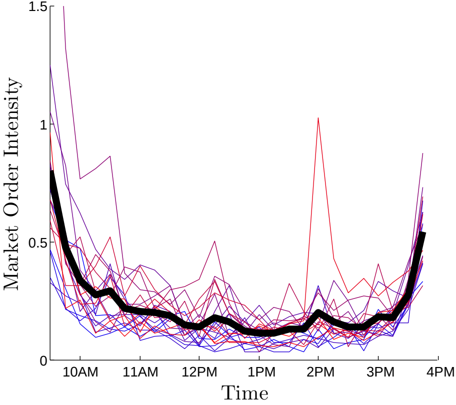


Figure : Market order intensity as a function of time for one month of INTC trades. Interval lengths are 15 minutes.

Seasonality - Market Order Intensity

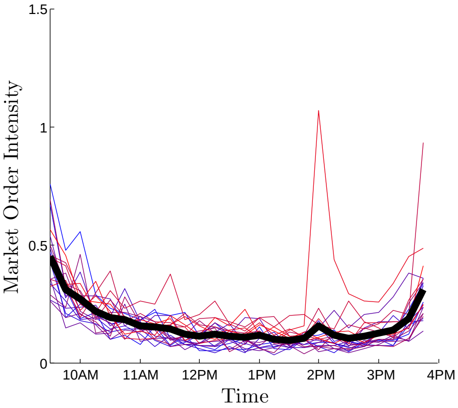


Figure : Market order intensity as a function of time for one month of ORCL trades. Interval lengths are 15 minutes.

Zero-Intelligence Performance

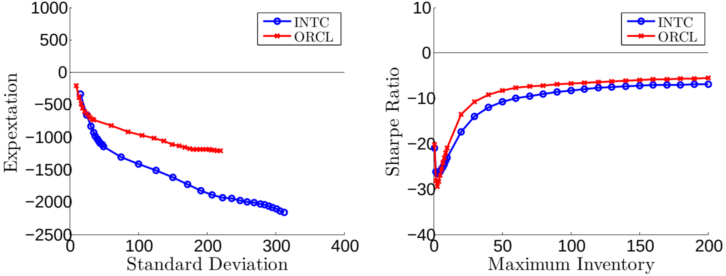
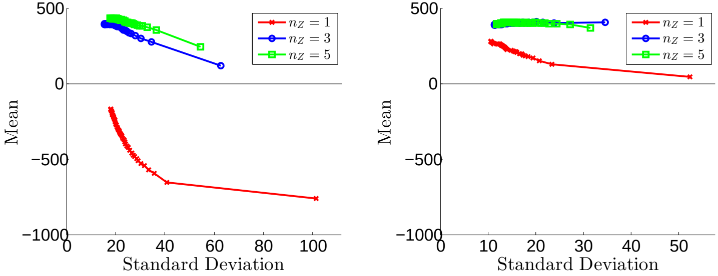


Figure : Naive strategy: annualised mean vs. standard deviation and annualised Sharpe ratio for various values of maximum inventory constraint from 1 to 200.

Performance of Historical Tests (INTC and ORCL)



(a) INTC (b) ORCL

Figure : Imbalance based strategy: Annualised Expectation vs. Standard Deviation of trading strategies based on different number of imbalance states and $\overline{Q} = -\underline{Q} = 50$. Each point represents a different value of ϕ, ranging from 0 to 10^{-5} (larger values of ϕ correspond to smaller values of standard deviation).

Sharpe Ratio of Historical Tests (INTC and ORCL)

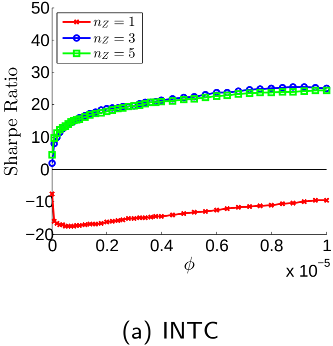


Figure : Imbalance based strategy: annualised Sharpe ratio for difference numbers of observable imbalance states and various inventory penalisations ϕ.

Zero-Intelligence

	Maximum Inventory Q							
	1	2	4	10	20	40	100	200
AA	-29.62	-33.77	-31.57	-21.89	-15.00	-10.18	- 6.97	- 6.41
AMAT	-34.20	-37.90	-33.66	-22.27	-15.15	-11.60	- 8.81	- 7.47
ARCC	-14.46	-16.07	-14.11	-11.28	- 9.13	- 6.95	- 5.30	- 4.85
BXS	-33.87	-27.24	-22.25	-17.15	-14.49	-12.12	-11.96	-11.96
CSCO	- 8.50	-18.03	-20.56	-15.71	-11.24	- 9.26	- 7.04	- 6.18
EBAY	-17.11	-20.55	-21.21	-18.08	-13.36	- 8.58	- 5.55	- 4.17
FMER	-33.12	-34.11	-28.45	-18.63	-13.44	-11.30	-10.25	-10.22
IMGN	-49.92	-40.68	-29.62	-18.39	-13.70	-10.54	- 8.33	- 7.90
INTC	-20.97	-26.15	-26.53	-23.11	-17.37	-11.98	- 8.24	- 6.90
NTAP	-45.96	-44.68	-37.70	-23.71	-15.08	- 9.75	- 8.06	- 8.04
ORCL	-20.24	-28.23	-28.36	-20.89	-13.59	- 9.17	- 6.71	- 5.53

Table : Annualised Sharpe ratio of zero-intelligence trading strategies based on various values of Q.

Volume Imbalance Strategy: always first in queue

	Inventory Penalty Parameter ϕ							
	0	10^{-7}	$2 \cdot 10^{-7}$	$4 \cdot 10^{-7}$	10^{-6}	$2 \cdot 10^{-6}$	$4 \cdot 10^{-6}$	10^{-5}
AA	6.12	8.35	9.79	12.06	17.34	20.33	23.99	27.93
AMAT	6.81	9.49	10.91	12.79	16.53	19.67	23.72	28.49
ARCC	-0.66	-0.01	0.73	1.40	3.07	4.35	5.85	9.34
BXS	4.25	4.50	4.78	5.08	6.45	6.43	7.57	9.42
CSCO	9.48	11.60	12.91	15.38	18.94	22.73	27.14	32.85
EBAY	3.41	4.59	6.00	7.74	9.96	12.42	16.12	20.91
FMER	5.24	6.57	7.02	8.05	10.48	12.69	15.77	20.71
IMGN	2.83	2.75	2.73	2.43	2.05	2.06	2.49	3.15
INTC	4.49	9.68	11.25	12.85	15.60	17.93	20.88	24.45
NTAP	1.63	2.29	2.63	2.73	2.74	3.79	4.53	7.32
ORCL	11.75	14.28	16.09	18.32	21.90	25.20	28.99	33.29

Table : Annualised Sharpe ratio of trading strategies based on various values of ϕ, the number of imbalance states here is $n_Z = 5$ and $n_\Delta = 2$ and the maximum inventory constraint is $Q = 50$.

Fill Probability is not 100%

- Our simulations assumed that a market order lifts the agent's placed limit order with 100% probability.
- We modify this behaviour and perform the historical simulation again to see how significant this assumption is on performance.
- The probability of fill assigned to each market order is chosen to depend on imbalance at the time of the order according to:

$$\begin{aligned} p^{+} &= \begin{pmatrix} 0.20 & 0.35 & 0.50 & 0.65 & 0.80 \end{pmatrix}, \\ p^{-} &= \begin{pmatrix} 0.80 & 0.65 & 0.50 & 0.35 & 0.20 \end{pmatrix}. \end{aligned}$$

Volume Imbalance Strategy with Fill Probability

	Inventory Penalty Parameter ϕ							
	0	10^{-7}	$2 \cdot 10^{-7}$	$4 \cdot 10^{-7}$	10^{-6}	$2 \cdot 10^{-6}$	$4 \cdot 10^{-6}$	10^{-5}
AA	2.42	3.96	4.92	4.46	6.83	10.85	12.81	17.28
AMAT	3.90	4.17	4.66	5.83	8.99	8.94	12.88	17.83
ARCC	-2.22	-2.62	- 2.23	- 2.63	- 1.26	- 1.27	- 0.68	2.03
BXS	1.83	1.89	1.48	0.68	1.00	1.92	2.45	1.62
CSCO	6.46	7.68	7.56	10.00	11.95	14.12	17.18	20.79
EBAY	1.99	2.71	5.11	4.92	3.84	6.34	7.77	11.02
FMER	1.42	2.96	2.78	4.05	5.61	5.48	7.46	9.65
IMGN	0.79	0.47	0.19	0.11	1.54	0.28	1.52	2.08
INTC	0.43	5.23	5.87	7.63	9.96	9.66	12.57	15.91
NTAP	-0.22	1.62	- 0.57	1.89	- 0.47	- 0.34	0.70	0.81
ORCL	7.87	9.68	10.65	11.71	13.00	15.87	15.78	21.14

Table : Annualised Sharpe ratio of trading strategies with modified fill probabilities depending on the level of imbalance. The trading strategy here is based on a number of imbalance states equal to $n_Z = 5$ and $n_\Delta = 2$ and the maximum inventory constraint is $Q = 50$.

Model Issues / Future Endeavours

- Multiple events within the same millisecond.
- Markovian assumptions associated with the model may be oversimplifying (i.e. evolution of spread and imbalance, arrival of market orders).
- Overly simplistic assumption about queue priority, interaction between market orders and limit orders (always able to post at front of queue).
- Latency issues can make it difficult to accurately observe the imbalance and spread processes.
- The class of control processes is less suitable if we consider stocks which are not considered large-tick stocks.

Conclusions

- The willingness of an agent to post limit orders is strongly dependent on the value of imbalance.
- Agent's should post buy orders more aggressively and sell orders more conservatively when imbalance is high. This reflects taking advantage of short term speculation and protecting against adverse selection.
- Corresponding opposite behaviour applies when imbalance is low.
- The additional value of being able to more accurately observe imbalance appears to have diminishing returns, but initially the additional value is very steep and the information embedded in the imbalance process should not be ignored.