Algorithmic Trading under the Effects of Volume Order Imbalance

7th General Advanced Mathematical Methods in Finance and

Swissquote Conference 2015

Lausanne, Switzerland

Ryan Donnelly

ryan.donnelly@epfl.ch

Swiss Finance Institute @ EPFL
joint work with

Álvaro Cartea, University College London
Sebastian Jaimungal, University of Toronto

September 9, 2015

Outline

- The limit order book.

- Volume order imbalance as an indicator of market behaviour.

- Imbalance model and market model.

- Optimal trading problem.

- Historical simulations.

The Limit Order Book

▶ The limit order book is a record of collective interest to buy or sell certain quantities of an asset at a certain price.
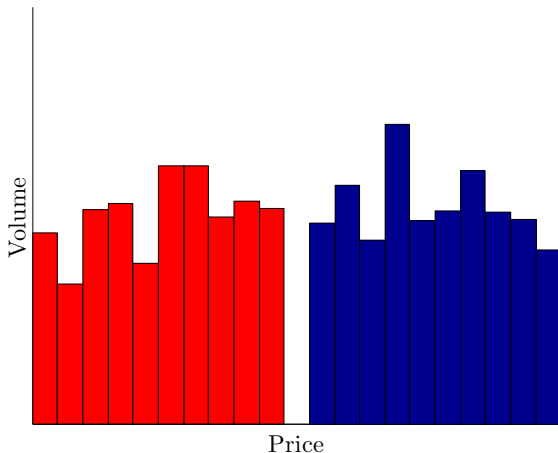
Buy Orders		Sell Orders	
Price	Volume	Price	Volume
60.00	80	60.10	75
59.90	100	60.20	75
59.80	90	60.30	50

▶ Graphical representation of the limit order book:
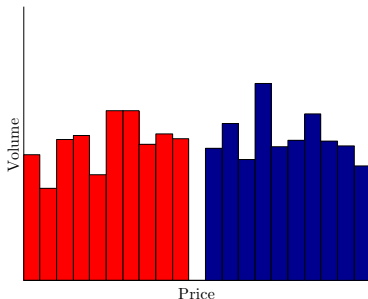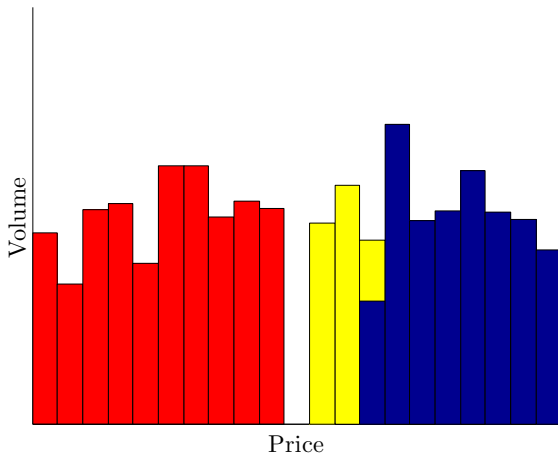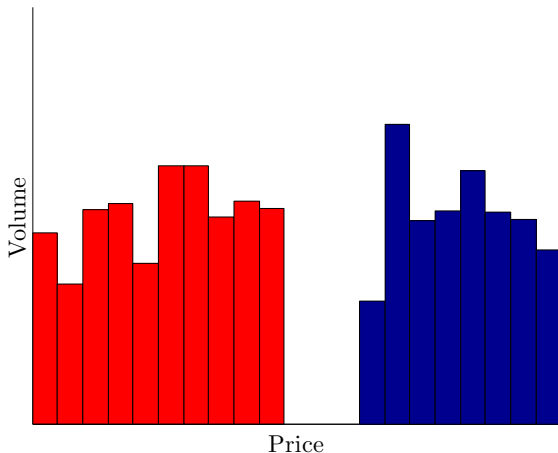
Market Orders

- An incoming market order lifts limit orders from the book.

Market Orders

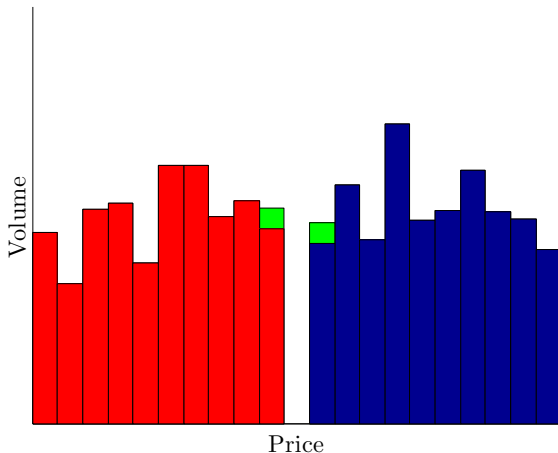▶ An incoming market order lifts limit orders from the book.

Market Orders

- An incoming market order lifts limit orders from the book.

Agent's Goal

▶ Optimally place limit orders in the limit order book (LOB)

Agent's Goal

Optimally placing limit orders in the limit order book requires the agent to specify dynamics of the market, namely:

- ▶ Dynamics of the midprice.

- ▶ Dynamics of the spread.

- ▶ Dynamics of incoming market buy and sell orders.

- ▶ Interaction between the agent's limit orders and incoming market orders.

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.
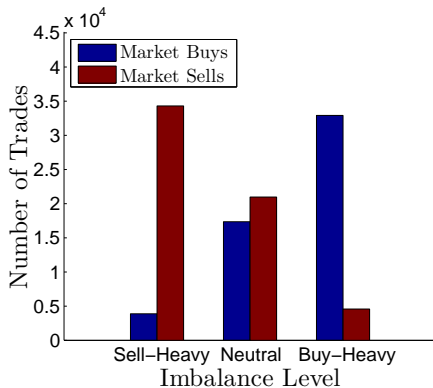
Volume Order Imbalance

Volume Order Imbalance

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

- Defined as:

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

- 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.

- $I_t \in [0, 1]$.
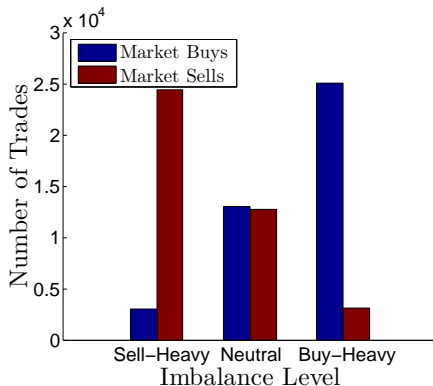
Predictive Power of Volume Imbalance - MO type

- ▶ 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 $[0, 0.33)$, $[0.35, 0.67]$, and $(0.67, 1]$.

Predictive Power of Volume Imbalance - MO type
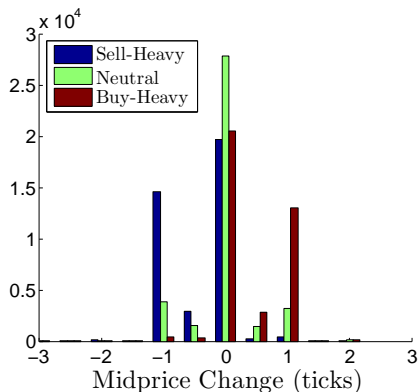
▶ Consider the types of market orders that are placed depending on the level of imbalance.

▶ More market **buys** when imbalance is **high**, more market **sells** when imbalance is **low**.

Figure : ORCL: one month of NASDAQ trades. Imbalance ranges are $[0, 0.33)$, $[0.35, 0.67]$, and $(0.67, 1]$.

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 [0, 0.33), [0.35, 0.67], and (0.67, 1].

Predictive Power of Volume Imbalance - Midprice Change

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

Figure : INTC: one month of NASDAQ trades. Imbalance ranges are $[0, 0.33)$, $[0.35, 0.67]$, and $(0.67, 1]$.

Predictive Power of Volume Imbalance - Midprice Change

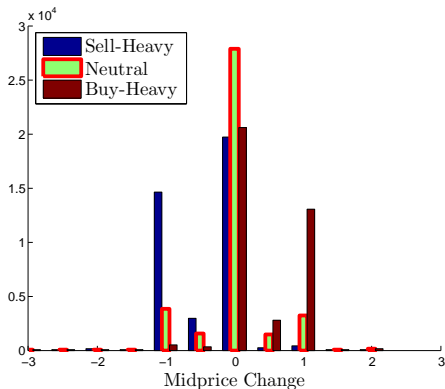- Distribution of midprice change $10ms$ after a market order.

Figure : INTC: one month of NASDAQ trades. Imbalance ranges are
$[0, 0.33)$, $[0.35, 0.67]$, and $(0.67, 1]$.

Predictive Power of Volume Imbalance - Midprice Change

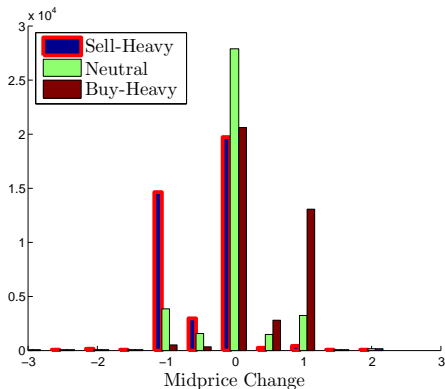▶ Distribution of midprice change $10ms$ after a market order.

Figure : INTC: one month of NASDAQ trades. Imbalance ranges are $[0, 0.33)$, $[0.35, 0.67]$, and $(0.67, 1]$.

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
$[0, 0.33)$, $[0.35, 0.67]$, and $(0.67, 1]$.

Predictive Power of Volume Imbalance - Midprice Change

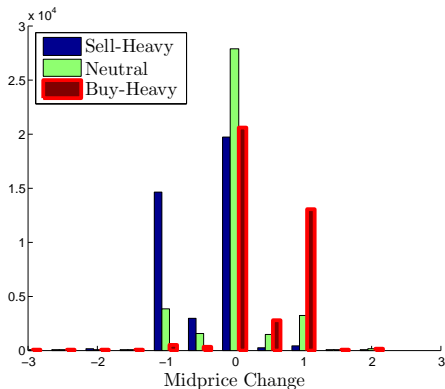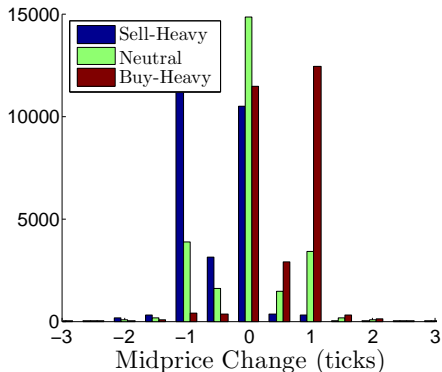▶ Distribution of midprice change 10ms after a market order.

Figure : ORCL: one month of NASDAQ trades. Imbalance ranges are $[0, 0.33)$, $[0.35, 0.67]$, and $(0.67, 1]$.
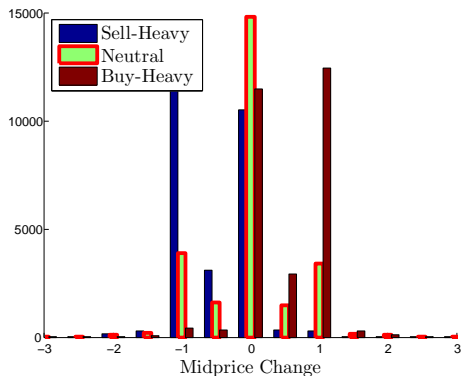
Predictive Power of Volume Imbalance - Midprice Change

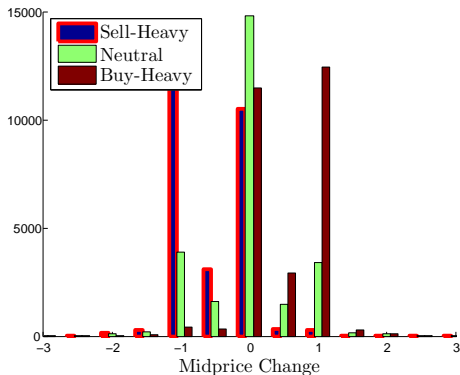▶ Distribution of midprice change 10*ms* after a market order.

Figure : ORCL: one month of NASDAQ trades. Imbalance ranges are
$[0, 0.33)$, $[0.35, 0.67]$, and $(0.67, 1]$.

Predictive Power of Volume Imbalance - Midprice Change

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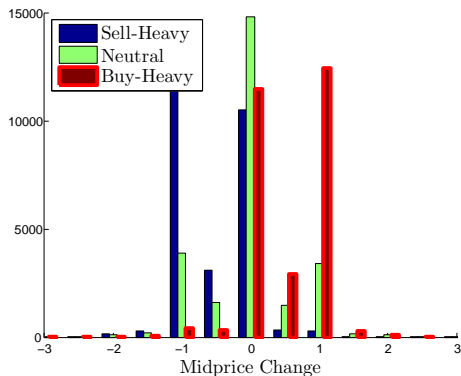

Figure : ORCL: one month of NASDAQ trades. Imbalance ranges are $[0, 0.33)$, $[0.35, 0.67]$, and $(0.67, 1]$.

Tick Activity

- ▶ Number of market orders that take place at ticks from midprice.

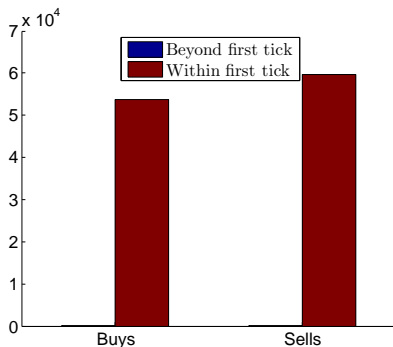Figure : INTC: one month of NASDAQ trades.

Tick Activity

▶ Number of market orders that take place at ticks from midprice.
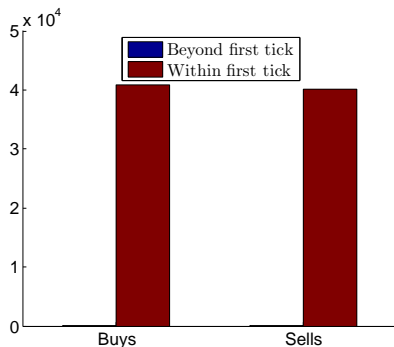
Figure : ORCL: one month of NASDAQ trades.

Market Model

Market Model

- Rather than model imbalance directly, a finite state imbalance regime process is considered, $Z_t \in \{1, \ldots, n_Z\}$.

- Z_t will act as an approximation to the true value of imbalance.

- The interval $[0, 1]$ is subdivided in to n_Z subintervals. $Z_t = k$ corresponds to I_t lying within the kth subinterval.

- The spread, Δ_t, is also takes values in a finite state space, $\Delta_t \in \{1, \ldots, n_\Delta\}$.

Market Model

- Let μ^I, 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:

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

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

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

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:

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

  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 maximize expected terminal wealth, penalized by cumulative inventory position:

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

▶ This value function has associated equation:

$$\partial_t H - \phi q^2 + \lambda^l(Z, \Delta)\mathbb{E}[\mathcal{D}^l H | Z, \Delta]$$

$$+ \sup_{\gamma^+ \in \{0,1\}} \lambda^+(Z, \Delta)\mathbb{E}[\mathcal{D}^+ H | Z, \Delta]$$

$$+ \sup_{\gamma^- \in \{0,1\}} \lambda^-(Z, \Delta)\mathbb{E}[\mathcal{D}^- H | Z, \Delta] = 0$$

$$H(T, x, q, S, Z) = x + q(S - \ell(q, \Delta))$$

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:

$$\partial_t h - \phi q^2 + \lambda^I(Z, \Delta)(q\epsilon^I(Z, \Delta) + \Sigma^I(t, q, Z, \Delta))$$

$$+ \sup_{\gamma^+ \in \{0,1\}} \lambda^+(Z, \Delta)\left(\gamma^+ \frac{\Delta}{2} + (q - \gamma^+)\epsilon^+(Z, \Delta) + \Sigma^+_{\gamma^+}(t, q, Z, \Delta)\right)$$

$$+ \sup_{\gamma^- \in \{0,1\}} \lambda^-(Z, \Delta)\left(\gamma^- \frac{\Delta}{2} - (q + \gamma^-)\epsilon^-(Z, \Delta) + \Sigma^-_{\gamma^-}(t, q, Z, \Delta)\right) = 0$$

$$h(T, q, Z, \Delta) = -q\ell(q, \Delta)$$

- 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

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

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

Optimal Trading Strategy – Parameters

- Allow three possible states of imbalance: $Z_t \in \{1, 2, 3\}$

- Two possible spreads: $\Delta_t \in \{1, 2\}$
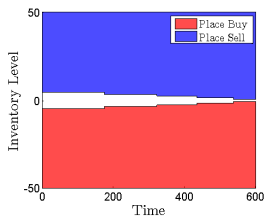
- MO arrival rates and price impact account for imbalance:

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

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

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

Optimal Trading Strategy
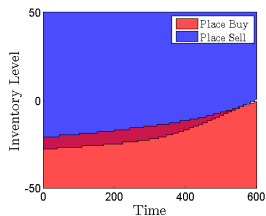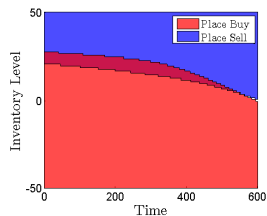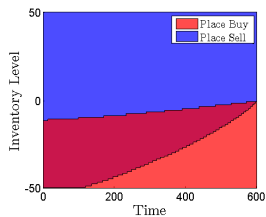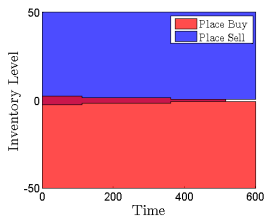

Low Imbalance

Middle Imbalance

$\Delta = 1$
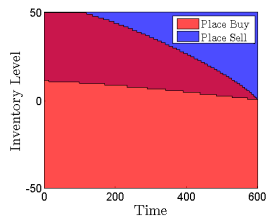
High Imbalance

Low Imbalance

Middle Imbalance

$\Delta = 2$

High Imbalance

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 naive 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.

Naive Results
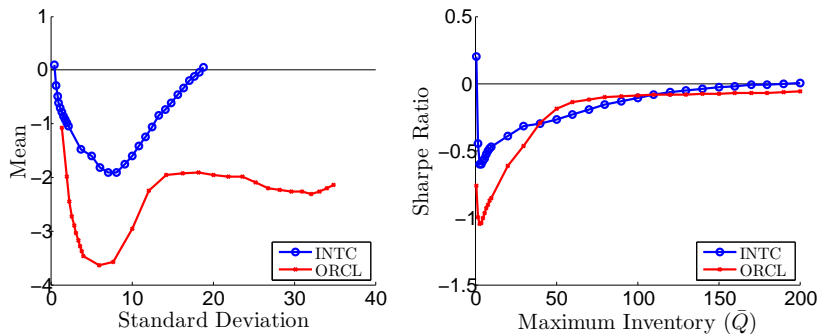

Figure : Naive strategy: mean vs. standard deviation and Sharpe ratio for various values of maximum inventory constraint from 1 to 200.
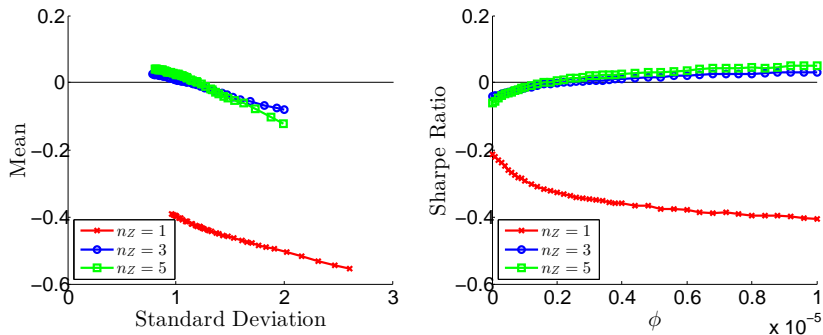
Imbalance Based Results: INTC

Figure : Imbalance based strategy: mean vs. standard deviation and Sharpe ratio for difference numbers of observable imbalance states and various inventory penalizations.
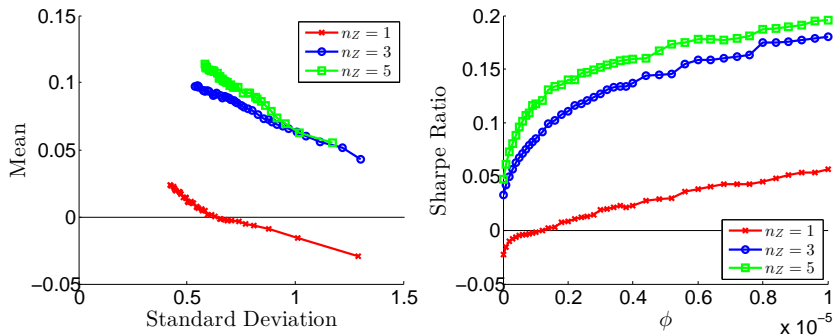
Imbalance Based Results: ORCL

Figure : Imbalance based strategy: mean vs. standard deviation and Sharpe ratio for difference numbers of observable imbalance states and various inventory penalizations.

Parameter Forecasting

- The historical simulations are performed out-of-sample.

- We employ a simple method of forecasting model parameters based on intraday seasonality.

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

- 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.

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.

Thanks for your attention!

Ryan Donnelly

ryan.donnelly@epfl.ch

Midprice Change After Buy Order

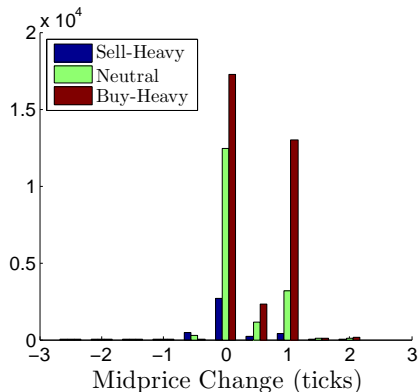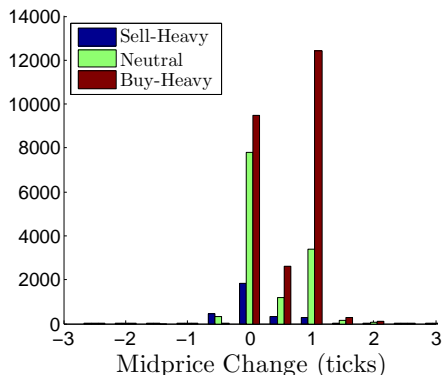- Distribution of midprice change 10*ms* after a market buy order.

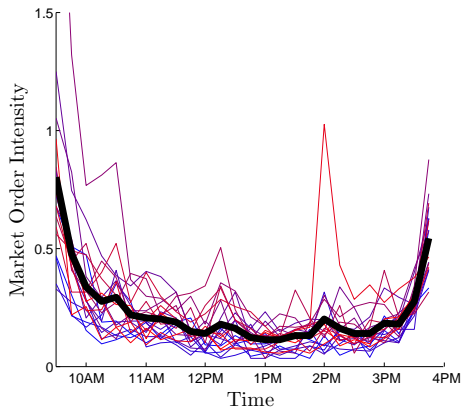Figure : INTC: one month of NASDAQ trades. Imbalance ranges are [0, 0.33), [0.35, 0.67], and (0.67, 1].

Midprice Change After Buy Order

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

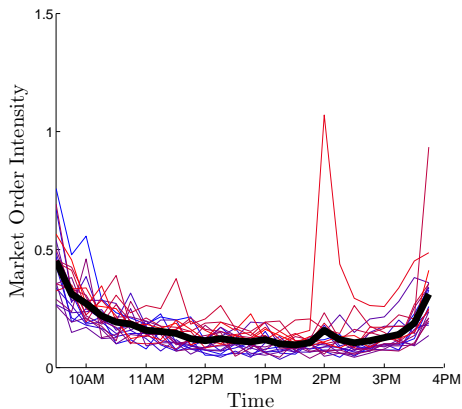Figure : ORCL: one month of NASDAQ trades. Imbalance ranges are [0, 0.33), [0.35, 0.67], and (0.67, 1].

Non-Markovian Behaviour - Market Order Intensity

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

Non-Markovian Behaviour - Market Order Intensity

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