Order Flows and Execution Costs

Mike Ludkovski

UC Santa Barbara

IPAM Workshop on High Frequency Trading, April 14, 2015

joint w/Kyle Bechler, Sebastian Jaimungal

Optimal Execution Modeling

Typical OE models are "reduced-form"

- Ingredients include
  - *reference price* (e.g. Brownian motion)
  - realized price (e.g. reference + offset) ;
  - price impact specification (e.g. transient linear impact w/exponential resilience);
  - liquidity state, etc.,
- Trading is continuous: diffusion setting
- Scheduling time-scale: 5–10 minutes

Data

- Actual data is based on the Limit Order Book
- Full information about all ticker events and snapshots of the book
- Mathematically represented in terms of point processes or queues
- Time-scale of $\ll 1$ second

How to reconcile/connect these frameworks?

OE vs LOB

- Existing OE models are **oblivious** to LOB states

  Despite these valuable proposals to model the dynamics of the order book at small time scales, no direct use of such effects in an optimal trading framework is currently available. —CA Lehalle, Handbook of Systemic Risk 2013

- Typical OE algorithms are “open-loop” – calibrate and forget

OE vs LOB

- Existing OE models are **oblivious** to LOB states

  Despite these valuable proposals to model the dynamics of the order book at small time scales, no direct use of such effects in an optimal trading framework is currently available. —CA Lehalle, Handbook of Systemic Risk 2013
- Typical OE algorithms are "open-loop" – calibrate and forget
- Wish to come up with a higher-level picture of LOB at the time-scale of the scheduler to build a "closed-loop" system:
- LOB $\rightarrow$ Liquidity/Market State $\rightarrow$ OE algorithm

What is Liquidity?

Challenges:

- How to measure liquidity of the LOB?
- How to measure price impact via the LOB?
- How to reconcile Market Orders and Limit Orders in LOB?

What is Liquidity?

Challenges:

- How to measure liquidity of the LOB?
- How to measure price impact via the LOB?
- How to reconcile Market Orders and Limit Orders in LOB?

- Static snapshots of the LOB have been extensively studied
- For example short-term effects from book imbalance and “micro price” which in turn drives mid-price movements
- Liquidity can be statically measured in terms of
  - Bid-Ask Spread
  - Depth
  - Execution Cost
- But these features are too detailed/fleeting at the OE scale
- Dynamic view is provided by the order flows (time-scale of minutes)

Outline

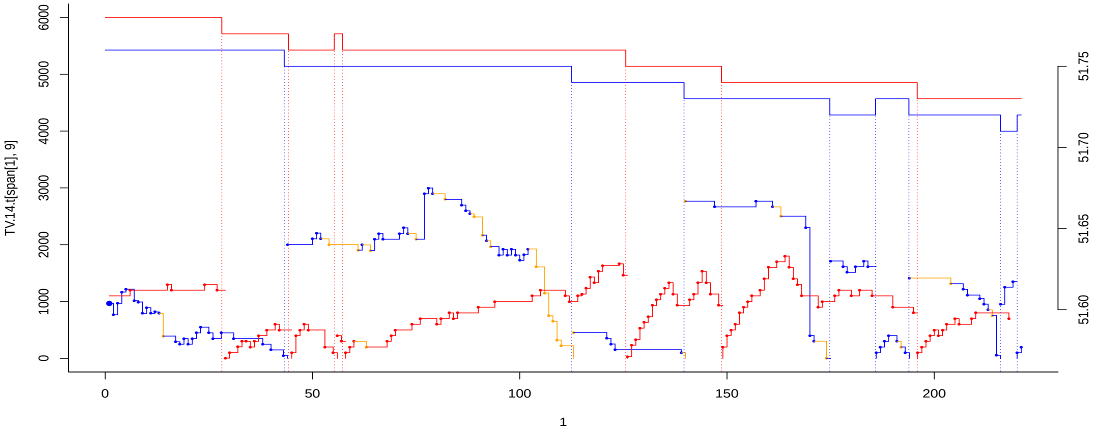
- Aim: include LOB info into OE schedulers
- **PART I:** Initial empirical analysis with a view to explain what is (statistically) important for order execution
- **PART II:** A simple model to incorporate order flow into OE
- For another take, see *Incorporating Order-Flow into Optimal Execution* by Cartea and Jaimungal (on SSRN, 2015)

Order Flow

- Need to aggregate LOB data to understand the more persistent features relevant for OE
- Moving from microstructure to the meso-scale (complements line of research that aggregates from point processes to diffusions)
- Aggregation naturally leads to consideration of Order Flows
- Related to metrics such as *Market Toxicity* or *Liquidity State*
- There is a lot of evidence that flows are persistent (hours or even days)

$\Rightarrow$ Trading shows up as a disturbance of order flow – links to price impact

LOB Evolution



Tick-by-tick evolution of LOB queues at the touch. Market orders are in orange. Mid price is plotted at the top (TEVA, May 3, 2011)

Main Aspects of Order Flow

- Must separately treat Limit Orders and Market Orders
- Consequently, must also separately treat each side of the book (cross-effects)
- Aggregating by time-slicing or volume-slicing?
- Order Flow has intrinsic links to price behavior: mechanical (slippage) and informational (price discovery)
- Economically, OF is tied to informational costs:
  - Participants fear that other traders are better informed
  - Noise traders trade in all directions; informed traders are deliberate
  - $\Rightarrow$ When order flow is toxic market makers provide less liquidity
  - Example of a metric: VPIN by Easley, Lopez de Prado, O'Hara (2012abcd) – ties toxicity to flow imbalance = ratio of noise/informed traders

Empirical Order Flow

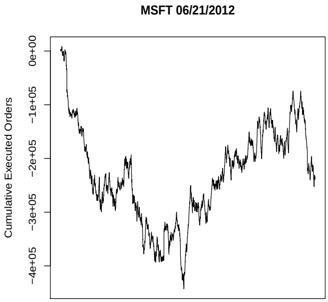


Figure: Cumulative net volume of MSFT executed trades over June 21, 2012

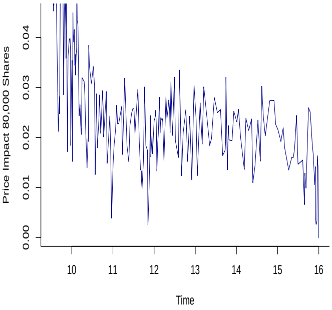


Figure: Price Impact for MSFT over same day

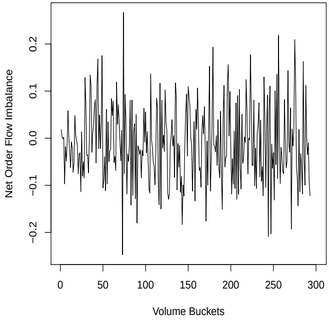


Figure: MSFT Net Order Flow over buckets of $1,000,000$

Notation

- LOB consists of pairs (p_t^j, v_t^j) indexed from the top queue.
- Mid-price: $P^M(t) := (p_1^A(t) + p_1^B(t))/2$
- Market orders (MO) and Limit orders (LO) form asynchronous time-series: time-stamp T_i, signed volume O_i, limit price S_i
- To synchronize & aggregate, construct **buckets**
- Slice the ticker according to intervals defined by (τ_k) (e.g. equal volume)
- Market buys: $VM_k^A = \sum_{i:\tau_k \le T_i^M \le \tau_{k+1}} O_i^M 1_{\{O_i^M > 0\}}$.
- $MI_k = VM_k^A - VM_k^B$
- Contemporaneous ask limit orders at the touch:

$$VL_k^A = \sum_{i:\tau_k \le T_i^L \le \tau_{k+1}} O_i^L 1_{\{S_i^L = p_1^B(T_i^L)\}}$$

How to Model Price Impact

- Permanent Price Impact
- Transient Price Impact
  1. Large orders walk the book (depends on the LOB shape)
  2. Liquidity providers replenish the queues and the book bounces back (resiliency)

How to Model Price Impact

- Permanent Price Impact
- Transient Price Impact
  1. Large orders walk the book (depends on the LOB shape)
  2. Liquidity providers replenish the queues and the book bounces back (resiliency)
- Obizhaeva/Wang: Realized execution price: $P_t = \tilde{P}_t + D_t$
- Unaffected $\tilde{P}$ is arithmetic BM
- Resilience D decays exponentially to zero
- Each term is linearly impacted by trading:

$$\tilde{P}_t = \tilde{P}_0 + \lambda MI_t + \sigma W_t, \qquad D_t = e^{-\rho t} D_0 + \nu MI_t.$$

- Assuming Gaussian D_0, gives a conditional Gaussian distribution for ΔP, linear in MI_t

LOB Perspective

- Immediate effect of a trade is deterministic given the LOB state. Cannot distinguish various sub-effects
- Difficult to measure long-term effects
- Have to describe the dynamic impact on LOB – eg arrival/cancellation rates

LOB Perspective

- Immediate effect of a trade is deterministic given the LOB state. Cannot distinguish various sub-effects
- Difficult to measure long-term effects
- Have to describe the dynamic impact on LOB – eg arrival/cancellation rates

Large Tick Assets

- Concentrate on liquid US equities
- Tick size is \$0.01; asset price is \$20-\$50
- Spread is almost always 1 tick
- The two queues at the touch are deep ($> 10AES$)
- Orders rarely “walk the book” and there is no *shape* to the LOB
- Typical immediate price impact is zero
- examples: MSFT, INTC, TEVA

Price Impact of Market Flow

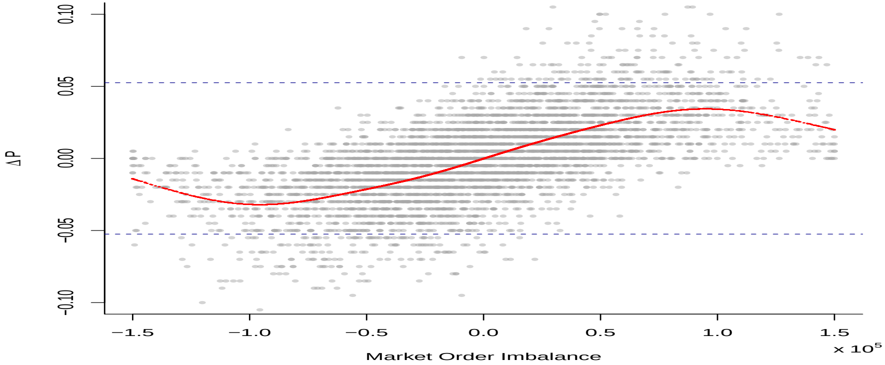


Figure: Relationship between mid-price change ΔP and net market order flow MI across **volume-aggregated** slices of 150,000 shares (1/80 of ADV) for MSFT during Jan-Mar 2011. The average price change was computed using Loess regression.

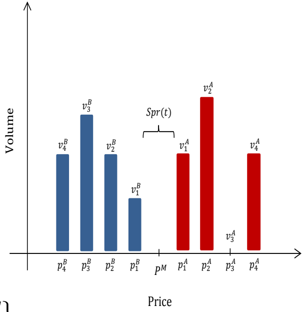
Explaining Price Movement based on Flows

- One sided flow consumes liquidity asymmetrically so the price moves
- Lots of buying $\rightarrow$ price goes up
- (Extreme MI typically anticipates low ΔP – observed S-shape)
- BUT: Even under balanced flow the price can make significant moves
- Good state: $MI = 50,000$ and $\Delta P = 0$.
- Bad state: $MI = 20,000$ and $\Delta P = 0.04$ large price move based on small net volume.
- Potential causes of such weak LOB resilience: are:
  1. Illiquid LOB
  2. Cancellations of limit orders (vanishing liquidity)

Static Liquidity of the LOB

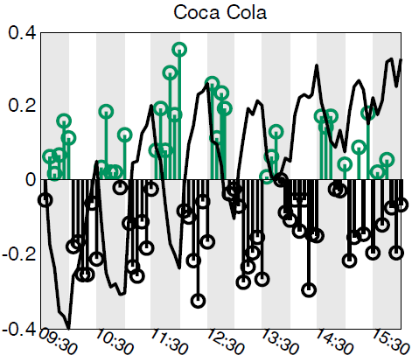
- Spread: $p_1^A - p_1^B$. Always 1-2 ticks – not indicative
- Static volume imbalance: $|v_1^A(t) - v_1^B(t)|$ - jumps each time the mid-price moves
- Depth: $\sum_i v_i^j$ (e.g. top 2 levels of the book)
- Price Impact: total effect relative to touch if instantaneously sell M shares:

$$PI_i := M^{-1} \sum_{i=1}^{i^M} v_i (S_i - p_1) + \bar{v_{i^M}} (S_{i^M} - p_1), \; i^M = \min\{i : \sum_{j<i} v_j > M\}$$



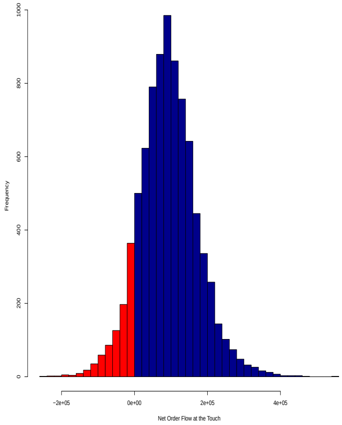
Vanishing Liquidity

- Static depth can also be misleading
- One-sided market flows can induce wild swings in additions/cancellations and generate a lot of mid-price volatility
- This depends on the correlation between VL amd MO and further latent regimes
- Picture from Lehalle et al (2012) – July 19, 2012



Limit Order Flows

- Limit order flow should generally be positive (more additions than cancellations) – counteracts market orders
- Instances where VL is negative are highly indicative of major price moves
- This is not surprising from the LOB mechanics
- But shows that time series of limit flows is crucial to identify reduced liquidity/toxicity



Limit Order Flows

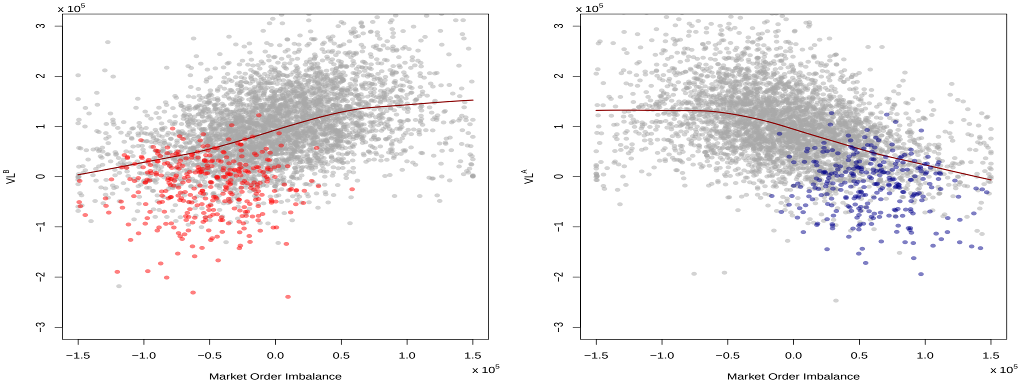
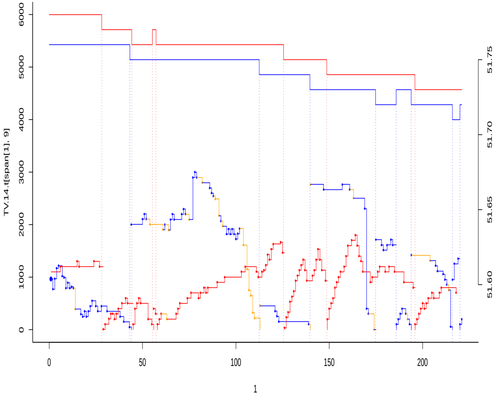


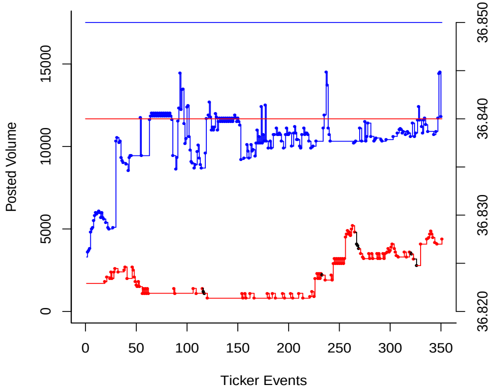
Figure: Relationship between net market order flow *MI* and concurrent top-level limit flow VL^j for MSFT during Jan-Mar 2011. Left: Bid, right: Ask. Highlighted points indicate buckets where the price change *exceeded 5 ticks*.

Can We Predict Scarce Liquidity?

- Generalized additive model to explain ΔP
- (GAM) Logistic regression to predict large price moves aka scarce liquidity
- Most important predictor is MO flow
- In both cases LO have much more explanatory power than price impact or book imbalance.
- This is not actionable (LO is contemporaneous)
- There is also some memory of scarce liquidity (based on ACF for volume-sliced time series)

Dependence Between Limit & Market Orders



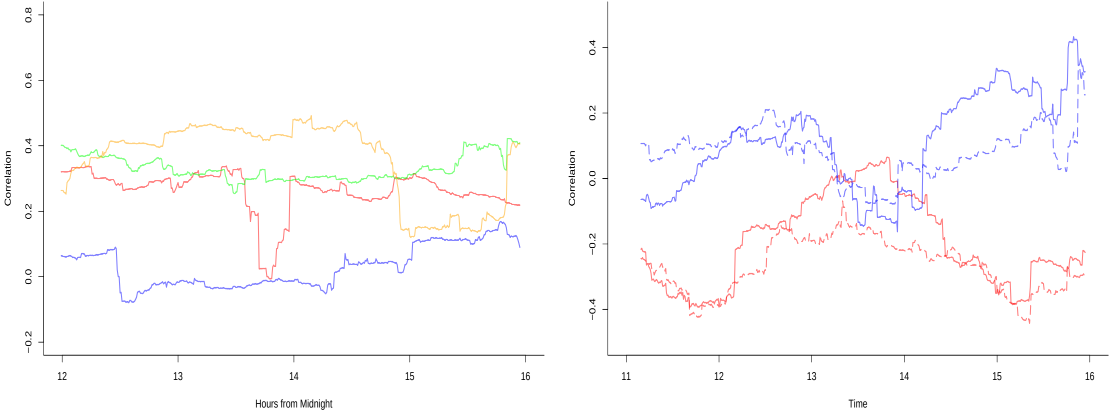


$-$*ve* Corr (TEVA) $+$*ve* Corr (ORCL)

Dependence Between Limit & Market Orders

- Could represent via regimes: resilient (*VL* and *MI* positive correlation)
- Scarce: negative correlation (price fade/"front running")
- Limit order flow is a key factor to explain price move conditional on MO
- **Open Problem:** How to define model-free dependence of 4 asynchronous marked point processes?

Correlation Between Limit & Market Orders



Smoothed contemporaneous correlation between *VL* and *MI* over the past 2.5 hours using 30-second buckets (4 days in 2011). Left: TEVA. Right: MSFT. Solid: bid-side. Dashed: ask-side

Another View of Persistency in Liquidity

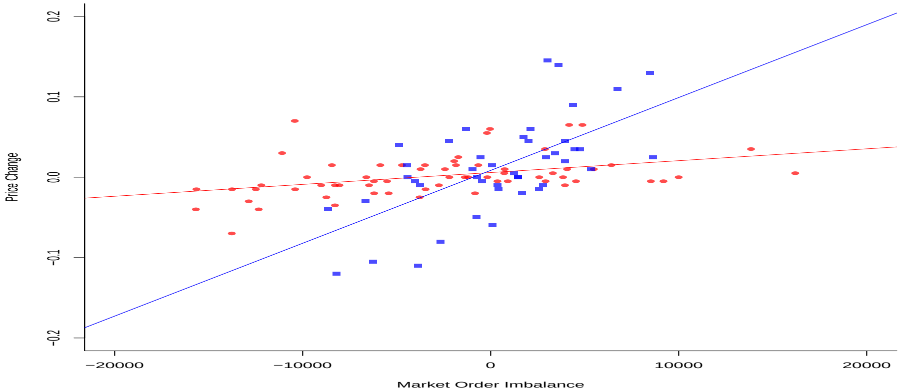


Figure: ΔP against MI and accompanying linear best fit lines for Feb 10, 2011 and March 9, 2011 in TEVA. Higher correlation implies stronger MI impact (also see Cartea & Jaimungal, 2015).

Current Work

- Quantify observed stylized features using statistical models (nonparametric regression)
- Suggest some ways of measuring the temporal correlation between limit and market order flows
- Investigate existence of statistical regimes (hidden Markov factors??)

Part II: Incorporating OF into OE

Two main concerns when executing a large order:

1. Price Impact

- Spatial effect: consume liquidity in terms of standing limit orders

Part II: Incorporating OF into OE

Two main concerns when executing a large order:

1. Price Impact
   - Spatial effect: consume liquidity in terms of standing limit orders
2. Information Leakage
   - Temporal effect: executed trade appears on the ticker tape
   - Other traders react and adjust their behavior (eg “front-running”) – will receive worse price in the future

AIM: integrate these ideas into a dynamic optimal execution framework

Treat order flow as a (controlled) state variable

Literature Landscape

- Work with continuous trading rates
- Assume a parametric form for inventory risk (Gatheral-Schied)
- Risk neutral agents (Gueant & Lehalle, ...)
- No fill risk (in contrast to Cartea & Jaimungal, Gueant et al, Bayraktar-L, ...)
- Treat only market orders (in the future: hybrid models like in Guilbaud & Pham, Carmona & Webster, ...)
- Postulate one-sided trading (in the future: allow two-sided to consider potential for market manipulation)
- Flat (constant depth) LOB $\rightarrow$ linear price impact
- Flat LOB $\rightarrow$ quadratic instantaneous execution cost
- (Order book resilience (Obizhaeva & Wang, Alfonsi et al))
- (Empirical evidence: Farmer, Bouchaud, ...)

Contributions

- Introduce order flow imbalance as a state variable
- Suggest a simple mechanism for informational costs (inspired by ELO12) – temporal impact beyond usual spatial impact
- Will also endogenize the execution horizon T – can trade faster under favorable conditions
- Derive closed-form approximate strategies for the resulting optimization problem

Execution Model

- Inventory x_t: liquidate x_0
- $t \mapsto x_t$ is absolutely continuous; trading rate is α_t
- $dx_t = -\alpha_t dt$
- Unaffected price S_t: martingale

Execution Model

- Inventory x_t: liquidate x_0
- $t \mapsto x_t$ is absolutely continuous; trading rate is α_t
- $dx_t = -\alpha_t dt$
- Unaffected price S_t: martingale
- Order flow imbalance Y_t:
  - Mean-reverting to zero
  - Stationary in long-run
  - Affected by execution algorithms
- $Y_t > 0$: buyers-market; $Y_t < 0$: sellers market

Order Flow Dynamics

- Unaffected order flow: $dY_t^0 = -\beta Y_t^0 dt + \sigma\, dW_t$
- With execution:
$$dY_t = (-\beta Y_t - \phi(\alpha_t))\, dt + \sigma\, dW_t$$
- i.e. $Y_t = Y_t^0 + \int_0^t e^{-\beta(t-s)} \alpha_s ds$
- Typical cases:
  - $\phi(\alpha) = \phi_t$ (deterministic information cost)
  - $\phi(\alpha) = \eta\alpha$ (linear in trading rate)
- Assume that flow is independent of price (empirical relationship is not clear) - more on this later

Optimization Problem

- Objective: $v(x,y) := \inf_{\alpha\in\mathcal{A}} \mathbb{E}_{x,y}\left[\int_0^{T_0} g(\alpha_s) + \lambda(x_s) + \kappa Y_s^2 ds\right]$
- Realized horizon $T_0 := \inf\{t : x_t^\alpha = 0\}$ – endogenous to the strategy α

Optimization Problem

- Objective: $v(x,y) := \inf_{\alpha \in \mathcal{A}} \mathbb{E}_{x,y}\left[\int_0^{T_0} g(\alpha_s) + \lambda(x_s) + \kappa Y_s^2 ds\right]$
- Realized horizon $T_0 := \inf\{t : x_t^{\alpha} = 0\}$ – endogenous to the strategy α
- $g(\alpha)$: price impact
- $g(\alpha) = \alpha^2$ (constant-depth LOB)

Optimization Problem

- Objective: $v(x,y) := \inf_{\alpha \in \mathcal{A}} \mathbb{E}_{x,y}\left[\int_0^{T_0} g(\alpha_s) + \lambda(x_s) + \kappa Y_s^2 ds\right]$
- Realized horizon $T_0 := \inf\{t : x_t^\alpha = 0\}$ – endogenous to the strategy α
- $g(\alpha)$: price impact
- $g(\alpha) = \alpha^2$ (constant-depth LOB)
- $\lambda(x)$: inventory risk
- $\lambda(x) = cx^2$ (Almgren-Chriss criterion) / $\lambda(x) = cx$ (similar to Gatheral-Schied)

Optimization Problem

- Objective: $v(x,y) := \inf_{\alpha\in\mathcal{A}} \mathbb{E}_{x,y}\left[\int_0^{T_0} g(\alpha_s) + \lambda(x_s) + \kappa Y_s^2 ds\right]$
- Realized horizon $T_0 := \inf\{t : x_t^\alpha = 0\}$ – endogenous to the strategy α
- $g(\alpha)$: price impact
- $g(\alpha) = \alpha^2$ (constant-depth LOB)
- $\lambda(x)$: inventory risk
- $\lambda(x) = cx^2$ (Almgren-Chriss criterion) / $\lambda(x) = cx$ (similar to Gatheral-Schied)
- κY^2 : information cost
- Unbalanced order flow: Higher liquidity costs

HJB Equation

- $0 = -\beta Y v_Y + \frac{1}{2}\sigma^2 v_{YY} + \inf_{\alpha \geq 0}\{g(\alpha) - \alpha v_x - \phi(\alpha) v_Y\} + \kappa Y^2 + \lambda(x)$
- Finite-fuel boundary condition: $v(0, y) = 0$ for all y
- Nonlinear parabolic PDE
- Hard to understand the structure
- Positivity constraint on α is challenging

HJB Equation

- $0 = -\beta Y v_Y + \frac{1}{2}\sigma^2 v_{YY} + \inf_{\alpha \geq 0}\{g(\alpha) - \alpha v_x - \phi(\alpha) v_Y\} + \kappa Y^2 + \lambda(x)$
- Finite-fuel boundary condition: $v(0, y) = 0$ for all y
- Nonlinear parabolic PDE
- Hard to understand the structure
- Positivity constraint on α is challenging
- To gain insights: build approximating problems by
  - (i) solving the fixed-horizon problem
  - (ii) optimizing over T

Fixed Horizon Problem

$$u(T,x,y) = \inf_{(\alpha_t)\in\mathcal{A}(T,x)} \mathbb{E}_{x,y}\left[\int_0^T \alpha_s^2 + \lambda(x_s^\alpha) + \kappa Y_s^2\, ds\right]$$

HJB equation becomes

$$u_T = \frac{1}{2}\sigma^2 u_{yy} + \kappa y^2 + \lambda(x) - \beta y u_y + \inf_\alpha \left\{\alpha^2 - \alpha u_x - \eta\alpha u_y\right\} \tag{1}$$

Singular boundary condition: $\lim_{T\downarrow 0} u(T,x,y) = \infty$ if $x \neq 0$

Proposition

The solution of (1) *has the form*

$$u(T,x,y) = x^2 A(T) + y^2 B(T) + xyC(T) + D(T), \tag{2}$$

where A, B, C, D solve a matrix Riccati ordinary differential equation.

Note: Riccati equations parameterized in terms of time-to-maturity τ

Execution Speed

- The corresponding optimal rate of liquidation is

$$\alpha_t^D = \frac{x_t(2A(\tau) + \eta C(\tau)) + Y_t(C(\tau) + 2\eta B(\tau))}{2}.$$

- Execution rate is linear in x_t and in Y_t (generalizes Almgren-Chriss)
- The Proposition only treats the unconstrained case $\alpha \in \mathbb{R}$: if T is large relative to x_0 or Y_t is negative enough then $\alpha^D < 0$
- As $t \to T$, the dynamic trading rate stabilizes, resembling a VWAP strategy.

Myopic Strategies $(\phi(\alpha) = \phi_t)$

- Suppose agent myopically optimizes only against price impact:
  $u^M(T,x,y) := \inf_{(x_t)} \left(\int_0^T \dot{x}_s^2 + \lambda(x_s) ds \right) + \int_0^T \kappa \mathbb{E}_y[Y_s^2] ds =: \mathcal{I} + \mathcal{O}$
- If $\lambda(x) = cx^2$ solution is $\left\{ \begin{aligned} x_t^{MH} &= \frac{x \sinh(\sqrt{c}(T-t))}{\sinh(\sqrt{c}T)} \\ \alpha_t^{MH} &= \frac{\sqrt{c}x \cosh(\sqrt{c}(T-t))}{\sinh(\sqrt{c}T)} \end{aligned} \right\}$
- Now $\phi_t = \eta \alpha_t^{MH}$ is the above deterministic function of $t \to Y_t$ is Gaussian with known moments
- $\mathcal{I}^{MH}(T,x) = \sqrt{c}x^2 \coth(\sqrt{c}T)$
  $\mathcal{O}^{MH}(T,x,y) = \kappa \int_0^T \left(ye^{-\beta t} - \int_0^t e^{-\beta(t-s)} \eta \alpha_s^{MH} ds \right)^2 + \frac{\sigma^2}{2\beta}(1 - e^{-2\beta t})\, dt$

Myopic Strategies $(\phi(\alpha) = \phi_t)$

- Suppose agent myopically optimizes only against price impact:
$$u^M(T,x,y) := \inf_{(x_t)} \left(\int_0^T \dot{x}_s^2 + \lambda(x_s) ds \right) + \int_0^T \kappa \mathbb{E}_y[Y_s^2] ds =: \mathcal{I} + \mathcal{O}$$
- If $\lambda(x) = cx^2$ solution is $\left\{ \begin{aligned} x_t^{MH} &= \frac{x \sinh(\sqrt{c}(T-t))}{\sinh(\sqrt{c}T)} \\ \alpha_t^{MH} &= \frac{\sqrt{c} x \cosh(\sqrt{c}(T-t))}{\sinh(\sqrt{c}T)} \end{aligned} \right\}$
- Now $\phi_t = \eta \alpha_t^{MH}$ is the above deterministic function of $t \to Y_t$ is Gaussian with known moments
- $\mathcal{I}^{MH}(T,x) = \sqrt{c} x^2 \coth(\sqrt{c}T)$
$$\mathcal{O}^{MH}(T,x,y) = \kappa \int_0^T \left(y e^{-\beta t} - \int_0^t e^{-\beta(t-s)} \eta \alpha_s^{MH}\, ds \right)^2 + \frac{\sigma^2}{2\beta}(1 - e^{-2\beta t})\, dt$$
- Similarly have closed-form expressions for cases $\lambda(x) = cx$ (Quadratic) and $\lambda(x) = 0$ (VWAP)
- Next step: Take existing closed-form expressions u^M and optimize over T

Optimizing the Horizon

- $T^* = \arg\min_T \; u(T, x, y)$
- Lemma: $T^* \in (0, \infty)$ (closed-form for $T \mapsto u(T, x, y)$)
- Open-loop (static): find T^* at the outset and implement $\alpha_t(T^*(x, y), x_t, Y_t)$

Optimizing the Horizon

- $T^* = \arg\min_T \; u(T, x, y)$
- Lemma: $T^* \in (0, \infty)$ (closed-form for $T \mapsto u(T, x, y)$)
- Open-loop (static): find T^* at the outset and implement $\alpha_t(T^*(x, y), x_t, Y_t)$
- Closed-loop (dynamic): continuously recompute T^*:

$$\tilde{\alpha}_t^M(x, y) := \alpha^M(T^*(x_t, Y_t), x_t, Y_t)$$

- Realized horizon $T_0(x, y)$ becomes random
- Dynamically recomputing T^* - adapt to changing Y_t without the indefinite horizon finite-fuel problem

Next up: we show these are in fact good approximations!

Comparison of costs

	Optimal Execution Cost				
	v	$\tilde{u}^D$	$\tilde{u}^{ML}$	u^D	u^{ML}
$\mathbb{E}[J(\alpha)]$	**4.257**	4.264	4.317	4.483	4.547
$SD(J(\alpha))$	1.50	1.45	1.39	1.77	1.84
$\mathbb{E}[T_0]$	3.87	3.70	3.48	3.43	3.43

Legend:

- v: directly from the HJB **pde** (fully numerical)
- $\tilde{u}(x, y)$: **closed-loop** optimization of T^*; T_0 is random
- $u(T^*, x, y)$: **static** optimization $T_0 = T^*$
- u^{ML}: **VWAP** on $[0, T^*]$
- ^{D} superscript refers to dynamic strategies based on Proposition 1 (**Riccati** eqns)

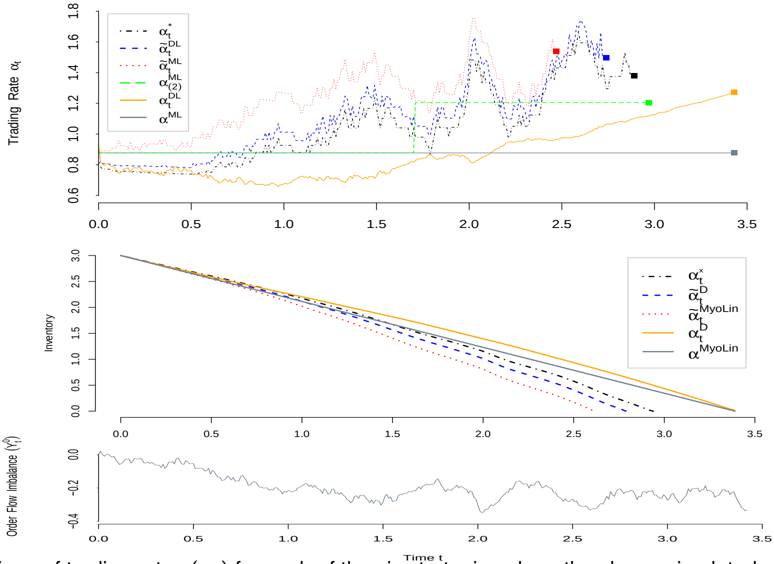


Figure: Comparison of trading rates (α_t) for each of the six strategies along the shown simulated path of (Y_t^0) (The realized (Y_t) depends on the strategy chosen).

Execution Paths

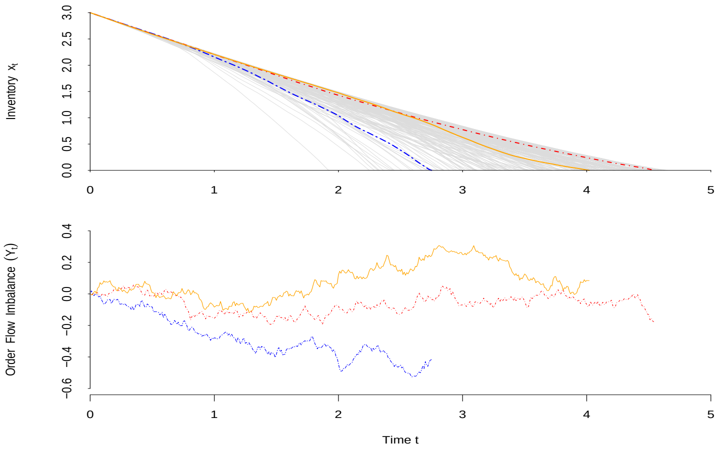


Figure: Top: 200 simulated trajectories from strategy $\tilde{\alpha}_t^D$. Highlighted are three trajectories resulting from different Y_t-paths. Bottom: Corresponding realizations of $t \mapsto Y_t$.

Extensions

- How to estimate/filter/infer Y_t from LOB data?
- information leakage $\phi(\alpha)$ depends on Y_t?
- Correlated S_t and Y_t: dependence between price movement and order flow

Extensions

- How to estimate/filter/infer Y_t from LOB data?
- information leakage $\phi(\alpha)$ depends on Y_t?
- Correlated S_t and Y_t: dependence between price movement and order flow

Thank You!

References

A. Cartea, S. Jaimungal
Incorporating Order-Flow into Optimal Execution
http://ssrn.com/abstract=2557457 (2015)

J. Gatheral and A. Schied
Dynamical models of market impact and algorithms for order execution
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2034178 (2013).

C.-A. Lehalle
Market Microstructure knowledge needed to control an intra-day trading process
Handbook on Systemic Risk, J-P Fouque and J Langsam Eds (2013)

K. Bechler and M. Ludkovski
Optimal Execution with Dynamic Order Flow Imbalance
Preprint, (2014) http://arxiv.org/abs/1409.2618