Optimal Trade Execution using Limit Order Book Information

Rolf Waeber  Sasha Stoikov

Operations Research & Information Engineering
Cornell University, Ithaca, NY

Sunday October 14, 2012
INFORMS Annual Meeting
Phoenix, AZ
Optimal Liquidation Problem

How to liquidate \( X \) shares of an asset?

1. **Macroscopic** time scale:
   - Horizon \( \bar{T} > 0 \) over which the shares \( X \) need to be liquidated.
   - Depends on *long term* variables: average daily volume, strategic considerations, news events, ...

2. **Mesoscopic** time scale:
   - Trade schedule \( 0 \leq t_0 \leq t_1 \ldots \leq t_i \leq \ldots \leq t_n = \bar{T} \) for the “child” trades.
   - Depends on *medium term* variables: volatility of the stock, risk aversion of the trader, price impact considerations, ...
   - See Bertsimas and Lo (1998), Almgren and Chriss (2001), Schied et al. (2010), Gatheral and Schied (2012), ...

3. **Microscopic** time scale:
   - Within a time interval \((t_i, t_{i+1}]\), what is the *timing* and the *type of order* used to liquidate the “child” trade?
   - Depends on *short term* variables: *limit order book information*.
The trade schedule

Almgren and Chriss (2001) show that \( x_t = \frac{\sinh(\kappa(T-t))}{\sinh(\kappa T)} X \), for \( t \in \{0, \ldots, T\} \), and \( \kappa \) depends on price volatility, risk aversion, and price impact.
Microscopic Time Scale

- We assume that the trade schedule is given.
- The goal is then to liquidate one lot (the shares $x_t$) in the time window $(t_i, t_{i+1}]$, i.e., what is the optimal time $\tau$ in $[0, T]$ to sell the lot, where $T = t_{i+1} - t_i > 0$.
- We only consider market orders.
- $T$ is typically short, e.g., 1 minute.
- For such short time periods, observing the limit order book can be very advantageous in identifying good liquidation times.
- However, latency in the trade execution can diminish this advantage!
Outline

1. Optimal liquidation of one lot in the time interval $[0, T]$:
   - The efficient price process.
   - Optimal stopping problem.
   - The trade and no-trade regions.

2. Trading with latency.

3. Dynamic programming approximation.

4. Backtesting strategy on TAQ data.

5. Conclusions and future research.
Optimal Liquidation
The Efficient Price Process

- The **microprice** process

\[ S(t) = S^b(t) + \theta(t), \]

where \( S^b(t) \) is the bid price and \( \theta(t) \) is the **imbalance process**.

- The imbalance process can be constructed using limit order book information, e.g.,

\[ \theta(t) = \frac{B(t)}{A(t) + B(t)}. \]
The Optimal Liquidation Problem

• Submitting a sell order at time $t$ yields payoff

$$S^b(t) = \lfloor S(t) \rfloor \leq S(t).$$

• **Goal:** Identify an optimal time $\tau$ in $[0, T]$ to sell the share and in turn to receive $\lfloor S(t) \rfloor$, i.e.,

$$V(t, s) = \sup_{t \leq \tau \leq T} E[\lfloor S(\tau) \rfloor | S(t) = s],$$

for $s \in \mathbb{R}$ and $t \in [0, T]$, and $\tau \in \mathcal{T}$, where $\mathcal{T}$ is the set of stopping times with respect to $\sigma(S(t))_{t \geq 0}$.

• **Assumptions:**
  - $S(t)$ is a Lévy process and distribution of jump size has a density,
  - $S(t)$ is a martingale,
  - $S(t) \in \mathbb{R}$, $S^b(t) \in \mathbb{Z}$ and $\theta(t) \in [0, 1)$. 
Trade/No-trade Regions

- “Trade” and “No-trade” region

\[ D = \{(t, s) \in [0, T] \times \mathbb{R} : V(t, s) = \lfloor s \rfloor\}, \]
\[ C = \{(t, s) \in [0, T] \times \mathbb{R} : V(t, s) > \lfloor s \rfloor\}. \]

- Liquidation time

\[ \tau_D = \inf \{ t \geq 0 | S(t) \in D \}. \]

Proposition

\[ \tau_D \in \mathcal{T} \text{ and } \]
\[ V(t, s) = \mathbb{E}[\lfloor S(\tau_D) \rfloor | S(t) = s]. \]
Proposition

The function $V(t, s)$ satisfies the following properties:

(a) fix $t \in [0, T]$, then $V(t, s)$ is non-decreasing in $s$;

(b) fix $s \in \mathbb{R}$, then $V(t, s)$ is non-increasing in $t$;

(c) $V(t, s + z) = V(t, s) + z$ for all $s \in \mathbb{R}$, $t \in [0, T]$ and $z \in \mathbb{Z}$;

(d) $V(t, z) = z$ for all $t \in [0, T]$ and $z \in \mathbb{Z}$;

(e) $V(T, s) = \lfloor s \rfloor$ for all $s \in \mathbb{R}$.
State Space Reduction

1. Property (c) shows that

\[ V(s, t) = \sup_{t \leq \tau \leq T} \mathbb{E}[S(\tau) \mid S(t) = s] = \sup_{t \leq \tau^\theta \leq T} \mathbb{E}[S(\tau^\theta) \mid S(t) = s], \]

where \( \tau^\theta \) are stopping times adapted to \( (\theta(t))_{t \geq 0} \).

2. Further, for \( \tau \in \mathcal{T} \):

\[ \mathbb{E}[S(\tau) \mid S(t) = s] = \mathbb{E}[S(\tau) - \theta(\tau) \mid S(t) = s] = s - \mathbb{E}[\theta(\tau) \mid S(t) = s]. \]

Hence, the problem \( V(t, s) \) is equivalent to

\[ V^\theta(t, u) = \inf_{t \leq \tau^\theta \leq T} \mathbb{E}[\theta(\tau^\theta) \mid \theta(0) = u], \]

and \( V(s, t) = s - V^\theta(t, s - \lfloor s \rfloor) \).
Optimal Liquidation based on Imbalance

Define

\[
D^\theta = \{(t, u) \in [0, T] \times [0, 1) : V^\theta(t, u) = u\},
\]
\[
C^\theta = \{(t, u) \in [0, T] \times [0, 1) : V^\theta(t, u) < u\}.
\]

Proposition

There exists a non-decreasing function \( w^*: [0, T] \to [0, 1] \) with \( w^*(T) = 1 \), such that

\[
D^\theta = \{(u, t) \in [0, 1) \times [0, T] : u \leq w^*(t)\}.
\]

Proposition

If \( S(t) \) has almost sure continuous sample paths, then \( w^*(t) = 0 \) for all \( t \in [0, T) \) and \( w^*(T) = 1 \).
Trade/no Trade Regions

Jump Process, $\lambda(T) = 500$, $K = 0.4$, $\sigma = 0.005$

Latency = 0T

Timesteps: 10000
States: 500
Sensitivity of the Trade Region

As the volatility of the price process $S(t)$ increases one can liquidate less aggressively (assuming risk-neutral liquidation).
Latency
A trade triggered at time $t$ is executed at time $t + l$ for $l > 0$.

Consider

$$V^l(t, s) = \sup_{t \leq \tau \leq T - l} \mathbb{E}[S^b(\tau + l) | S(t) = s],$$

where $\tau \in \mathcal{T}$.
Latency is Costly

**Proposition**

Fix \( t \in [0, T], s \in \mathbb{R} \), then \( V^l(t, s) \) is non-increasing in \( l \) for \( l \in [0, T] \).
Reducing the State Space

Analogous to the no-latency case one can show that $V^l(t, s)$ is equivalent to

$$V^{l, \theta}(t, u) = \inf_{t \leq \tau^{\theta} \leq T - l} \mathbb{E}[G^{\theta, l}(\tau^{\theta}) | \theta(t) = u],$$

where $G^{\theta, l}(u) = \mathbb{E}[\theta(l) | \theta(0) = u]$. 

![Payoff Function $G^{\theta, l}(\theta(t))$](image)
The “trade region” is still connected, but the “no-trade” region does not need to be connected anymore:

**Proposition**

There exists a non-decreasing function \( w_i^* : [0, T] \rightarrow [0, 1] \) and a non-increasing function \( v_i^* : [0, T] \rightarrow [0, 1] \), with \( v_i^* \leq w_i^* \), \( w_i^*(t) = 1 \) for \( t \in [T - l, T] \) and \( v_i^* = 0 \) for \( t \in [T - l, T] \), such that

\[
D_{\theta, l} = \{(t, u) \in [0, T] \times [0, 1] : v_i^*(t) \leq u \leq w_i^*(t)\}.
\]
Trade/No-Trade Regions with Latency

The red line shows the optimal policy without adjustment.
Dynamic Programming Approximation
Discretization Approximation

- Knowing $V^l(t, s)$, resp., $V^\theta, l(t, u)$, is enough to identify good liquidation times.
- For general Lévy-processes no closed-form solutions exist, hence we rely on a time- and space-discretization.
- Let $N, E \in \mathbb{N}$. Define,

$$k : [0, T] \to K = \{0, \ldots, N\}$$

$$t \mapsto k(t) = \sup \{n \in \{0, \ldots, N\} \mid nT/N \leq t\},$$

$$h : [0, 1) \to H = \{1, \ldots, E\}$$

$$x \mapsto h(x) = \lfloor Ex \rfloor + 1.$$

- These mappings transform the original state space $[0, T] \times [0, 1)$ into a discrete state space with $(N + 1)E$ states.
Discretization Approximation cont.

- Consider the discrete-time, discrete-space version of $\theta(t)$, i.e.,

$$\tilde{\theta}(j(t)) = h(\theta(t)).$$

- Let $P_{\tilde{\theta}}$ denote the transition matrix of the homogenous Markov chain $\tilde{\theta}(n)$, i.e., the matrix with entries

$$p_{ij} = \mathbb{P}\left(\tilde{\theta}(n + 1) = j|\tilde{\theta}(n) = i\right),$$

for $i, j \in \{1, \ldots, E\}$ and $n \in \{0, \ldots, N\}$.

- Approximate $V^\theta,l(t, u)$ and $D^\theta,l$ with

$$\tilde{V}_{E,N}^\theta,l(n, i) = \inf_{\tau \in \tilde{T}^\theta,l} \mathbb{E}[\tilde{G}^\theta,l(\tau)|\tilde{\theta}(n) = i],$$

$$\tilde{D}_{E,N}^{l,\theta} = \left\{ (n, i) \in K \times H : \tilde{V}_{E,N}^\theta,l(n, i) = \tilde{G}^{\theta,l}(i) \right\}.$$
Dynamic Program

- Bellman's recursion:

\[
\tilde{V}_{E,N}^{\theta,l}(n, i) = \max \left\{ \tilde{G}^{\theta,l}(i), \mathbb{E}[\tilde{V}_{E,N}^{\theta,l}(n + 1, \tilde{\theta}(n + 1))|\tilde{\theta}(n) = i] \right\},
\]

for \( i \in \{0, \ldots, N\} \) and \( n \in \{0, \ldots, N\} \).

- Conditional probability:

\[
\mathbb{E}[\tilde{V}_{E,N}^{\theta,l}(\tilde{\theta}(n + 1), n + 1)|\tilde{\theta}(n) = i] = \sum_{k=1}^{E} p_{ik} \tilde{V}_{E,N}^{\theta,l}(n + 1, k).
\]
As $N \to \infty$ and $E \to \infty$ the boundary between trade and no-trade region converges to a smooth curve.
Empirical Backtesting
Set Up

- Backtesting on TAQ data for 5-years US treasury bonds for 21 days (July 2010).
- Assume that one lot is traded per minute.
- Need a price model for 
  \[ S(t) = S^b(t) + \theta(t), \]
  the bid price \( S^b(t) \) can be observed in TAQ data.
- Many possibilities to construct \( \theta(t) \) based on limit order book data.
- \( \theta(t) \) is the trade secret of many algorithmic trading companies.
Imbalance Process

- Assume $S(t) = S^b(t) + \theta(t)$ is a pure jump process with symmetric jumps.
- We then use

$$\theta(t) = g \left( \frac{B(t)}{A(t) + B(t)} \right),$$

where $A(t)$ (or $B(t)$) is the best ask (or bid) size and $g(\cdot)$ is a cubic polynomial with constraints $g(0) = 0$, $g(0.5) = 0.5$ and $g(1) = 1$ (leaving 1 degree of freedom).
- This transformation makes the stationary distribution almost uniform, which is necessary when $S(t)$ is a pure jump process with symmetric jumps.
Calibrate Price Model

- Jumps occur according a homogenous Poisson process with intensity parameter $\lambda = 355$ jumps per minute.
- A two-sided generalized Pareto distribution with shape parameter $K = 0.4$ and scale parameter $\sigma = 0.01$ provides a good fit of the jump distribution.

![Jump Density Comparison](image-url)
Optimal Stopping vs. TWAP Strategy

- The time-weighted average price (TWAP) strategy liquidates one share per minute independently of the state of the limit order book.
- Consider residuals $R = S^b(\tau) - S^b(0)$, where $\tau$ is the stopping time from the optimal stopping problem $V(t,s)$ calibrated to a pure jump process $S(t)$.
- Compare 5,649 intervals of length 1 minute.
- Without latency the optimal liquidation strategy saves on average 33 $ per share, i.e., more than a 1/3 of the spread (Spread is 78$ for 5 yrs US-treasury bonds):

<table>
<thead>
<tr>
<th>Optimal policy vs. TWAP</th>
<th>$\mathbb{E}[R]$</th>
<th>$\sigma(R)$</th>
<th>$P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>33.01 $</td>
<td>51.56 $</td>
<td>$2.64 \cdot 10^{-200}$</td>
</tr>
</tbody>
</table>
Realized Imbalances

\( \theta_N \)

\( \theta_\tau \)

Frequency

Optimal Liquidation Problem

Optimal liquidation

Latency

Dynamic programming

Backtesting on TAQ data

Conclusions

References
Empirical Evidence for Price Model

Empirical observed imbalance $\theta(t)$ conditioned a trade occurs on the next quote.
Effect of Trading Horizon $T$
Cost of Latency

- **Cost of latency:**
  \[ COL = \mathbb{E}[S_b(\tau^0 + l) - S_b(\tau^0)] , \]
  where \( \tau^0 \) is the stopping time induced by \( V(t, s) \).

- **Adjusted cost of latency:**
  \[ COL_{adj} = \mathbb{E}[S_b(\tau^l + l) - S_b(\tau^0)] , \]
  where \( \tau^l \) is the stopping time induced by the adjusted problem \( V^l(t, s) \).

- **Note,** we do not calculate the COL with respect to the TWAP strategy, but with respect to the **optimal strategy with no latency.**
• 10ms latency $\approx 10\$ per share.

• For latencies $\geq 2000\text{ms}$ (i.e., 2 secs) the advantage of observing the limit order book diminishes (performance becomes similar to TWAP).

• Adjusting the liquidation policy to latency brings only minor improvement in the performance.
Conclusions
Conclusions

- We consider an optimal stopping problem that depends on:
  - Information found in the order book;
  - The time left to catch up with the TWAP algorithm;
  - Latency.

- The solution comes in the form of a trade/no-trade regions in the imbalance process.

- We estimate model parameters using limit order book information.

- We find that our optimal liquidation algorithm significantly outperforms a TWAP algorithm.

- We quantify the cost of latency.

- **Future research:**
  1. Modeling limit order executions.
  2. Game-theoretic considerations, i.e., prevent front-running of such a liquidation algorithm.
THANK YOU!


J. Gatheral and A. Schied (2012): Dynamical models of market impact and algorithms for order execution. *Available at SSRN 2034178*.