Empirical Estimation of Market Impact and Calibration of Optimal Execution Strategies

Enzo Busseti

J.P. Morgan

October 27, 2011
In this presentation I describe my work on market impact and optimal execution.

- I introduce the propagator model for market impact:
  - different time settings (trade time, real time);
  - estimation procedure with trade and quote data;
  - empirical results.

- Then describe the framework for optimal execution:
  - minimizing impact costs;
  - minimizing impact costs & risk;
  - show the resulting optimal trading schedules.
Propagator model for market impact
  The model
  Estimation procedure with trade and quote data
  Empirical results

Optimal execution
  The model
  Minimizing impact costs
  Minimizing impact costs & risk
Impact propagator model

The price dynamics is given by

\[ p_j = p_0 + \sum_{k=0}^{j-1} [\eta_k + f(v_k)G_0(j - k)], \]

where \( \eta_k \) are random shocks.

We have that:

- the *impact function* \( f(v_k) \) models the volume dependence of impact;
- the *propagator* \( G_0 \) models the temporal structure.

---

Generalizing to different time settings

\[ p_j = p_0 + \sum_{k=0}^{j-1} \left[ \eta_k + f(v_k) G_0(j - k) \right], \]

**Trade time:** \( p_n \) is the mid-price just before the \( n \)-th transaction, \( v_n \) the signed volume of the \( n \)-th transaction.

**d-Aggregated trade time:** \( p_n^d \) is the mid-price just before the \( (n \times d) \)-th transaction, \( v_n^d \) the sum of the signed volumes of transactions \([n \times d]\) to \([(n + 1) \times d]\).

**Real time:** \( p_n^{RT} \) is the mid-price at the start of the interval (i.e. every 5 minutes), \( v_n^{RT} \) the *normalized volume imbalance* in the interval.
Propagator model for market impact

- The model
- Estimation procedure with trade and quote data
- Empirical results

Optimal execution

- The model
- Minimizing impact costs
- Minimizing impact costs & risk
The estimation procedure

Two steps:

▶ estimate the impact function $f(v)$, as the average 1-lag return conditioned to transaction volume $v$

$$f(v) = \mathbb{E}[r|v];$$

▶ estimate the propagator $G_0$ with a (custom) linear regression of the returns on the $f(v)$ series:

$$(p_{n+1} - p_n) = r_n = \sum_{k=0}^{n} [G_0(k+1) - G_0(k)]f(v_{n-k}) + \eta_n.$$
Data

We estimate the model on four stocks from two different trade and quote dataset (on the LSE):

- AZN and VOD,
  - 32 months data from May, 2000 to December, 2002;
  - $\approx 10^3$ transactions per day.

- BP and BARC,
  - 4 months data from January 2011 to April 2011;
  - $\approx 10^4$ transactions per day.
Propagator model for market impact
  The model
  Estimation procedure with trade and quote data
  Empirical results

Optimal execution
  The model
  Minimizing impact costs
  Minimizing impact costs & risk
Impact function $f(v)$ estimate, trade time

Fit: $f(v) = \theta \, \text{sign}(v) \cdot |v|^\phi$ (a step function would be fine).
Propagator $G_0$ estimate, trade time

$$G_0^{fit}(l) = \frac{\Gamma_0}{(l^2 + l^2)^{\beta/2}}.$$
Impact function $f(v)$ estimate, 8-aggregated trade time

Fit: $f(v) = \theta \arctan(\phi v)$. 
Propagator $G_0$ estimate, 8-aggregated trade time

The model
Estimation procedure with trade and quote data
Empirical results

Fit $G_0^{\text{fit}}(l) = \frac{\Gamma_0}{(l_0^2 + l^2)^{\beta/2}}$. 
Impact function $f(v)$ estimate, real time

Fit: $f(v) = \theta \arctan(\phi v)$. 
Propagator $G_0$ estimate, real time

\begin{align*}
\text{Fit } G_0^{\text{fit}}(l) &= \frac{\Gamma_0}{(l_0^2 + l^2)^{3/2}}.
\end{align*}
Fraction of volatility explained by the model

Estimated \( R^2 = 1 - \frac{\sum n^2}{\sum r^2} \) coefficients

<table>
<thead>
<tr>
<th>Symbol</th>
<th>trade time</th>
<th>8 aggregated transactions</th>
<th>64 aggregated transactions</th>
<th>1 minute aggregation</th>
<th>5 minutes aggregation</th>
</tr>
</thead>
<tbody>
<tr>
<td>AZN</td>
<td>0.224</td>
<td>0.245</td>
<td>0.304</td>
<td>0.192</td>
<td>0.204</td>
</tr>
<tr>
<td>VOD</td>
<td>0.257</td>
<td>0.329</td>
<td>0.440</td>
<td>0.209</td>
<td>0.292</td>
</tr>
<tr>
<td>BP</td>
<td>0.083</td>
<td>0.070</td>
<td>0.047</td>
<td>0.054</td>
<td>0.039</td>
</tr>
<tr>
<td>BARC</td>
<td>0.090</td>
<td>0.069</td>
<td>0.044</td>
<td>0.070</td>
<td>0.040</td>
</tr>
</tbody>
</table>

AZN and VOD belong to the 2000-2002 dataset, BP and BARC to the 2011 dataset.
Propagator model for market impact
   The model
   Estimation procedure with trade and quote data
   Empirical results

Optimal execution
   The model
   Minimizing impact costs
   Minimizing impact costs & risk
Optimal execution model

- We trade $X$ shares in a time horizon $T$ (1 market day in my simulations);
- divide $T$ into $N$ intervals of length $\tau = \frac{T}{N}$ (5 minutes in my simulations);
- define a *trading schedule* $\vec{v}$;
- express the *execution costs*:

$$C(\vec{v}) = \sum_{k=1}^{N} v_k P_k - XP_0;$$

- minimize the expected value:

$$\mathbb{E}[C(\vec{v})].$$
Propagator model for market impact

The model
Estimation procedure with trade and quote data
Empirical results

Optimal execution

The model
Minimizing impact costs
Minimizing impact costs & risk
Optimal execution with market impact propagator

With the propagator impact model we have:

\[
E[c(\vec{v})] = \sum_{n=0}^{N-1} \nu_n \left[ \sum_{k=0}^{n} f(v_k) G_0(n-k) \right]
\]

If we approximate a linear \( f(v) \) we can introduce the impact matrix \( \mathcal{I} \):

\[
\mathcal{I}_{i,j} \equiv \begin{cases} 
\theta_i G_0(i-j) & i \geq j \\
0 & i < j 
\end{cases}
\]

So that the execution costs are given by:

\[
E[c(\vec{v})] = \vec{v}^T \mathcal{I} \vec{v}.
\]
Propagator model for market impact
Optimal execution

Optimal schedule with the minimization procedure

We minimize analytically $E[c(\vec{v})] = \vec{v}^T \mathcal{I} \vec{v}$ (linear problem). The impact parameters come from the previous estimates.
Adding bid-ask spread costs

An important addition to the model is to account for bid-ask spread costs:

$$E[c^{total}(\vec{v})] = \vec{v}^T I \vec{v} + \delta 1^T |\vec{v}|.$$ 

The problem is not linear anymore, and there is no analytical solution.

$$\downarrow$$

I used a numerical optimization routine.
Optimal schedule with the minimization procedure

We minimize numerically $(\vec{v}^T \mathcal{I} \vec{v} + \delta \mathbf{1}^T |\vec{v}|)$. The impact parameters come from the previous estimates.
Performances

<table>
<thead>
<tr>
<th>Symbol</th>
<th>“flat” solution</th>
<th>“U-shaped” solution</th>
<th>(difference)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AZN</td>
<td>4.36</td>
<td>4.29</td>
<td>1.63%</td>
</tr>
<tr>
<td>VOD</td>
<td>9.82</td>
<td>9.76</td>
<td>0.61%</td>
</tr>
<tr>
<td>BP</td>
<td>2.36</td>
<td>2.33</td>
<td>1.29%</td>
</tr>
<tr>
<td>BARC</td>
<td>3.69</td>
<td>3.64</td>
<td>1.37%</td>
</tr>
</tbody>
</table>
Propagator model for market impact
The model
Estimation procedure with trade and quote data
Empirical results

Optimal execution
The model
Minimizing impact costs
Minimizing impact costs & risk
Optimal execution with risk aversion - Almgren & Chriss\textsuperscript{2}

We take variance $\mathbf{V}[C(\bar{v})]$ as a proxy for risk, $\lambda$ is the risk aversion parameter.

One finds the solutions to:

$$\min_{\mathbf{v}_t} (\mathbf{E}[C(\bar{v})] + \lambda \mathbf{V}[C(\bar{v})]).$$

With the classic impact model (temporary + permanent) one obtains $\Rightarrow$

Optimal schedule with risk-aversion I

With our impact model and estimated parameters we solve numerically: \( \min_{\nu_t} \left( \mathbb{E}[C(\vec{\nu})] + \delta \vec{1}^T |\vec{\nu}| + \lambda \mathbf{V}[C(\vec{\nu})] \right) . \)
Optimal schedule with risk-aversion II

With higher coefficients of risk-aversion $\lambda$ the optimal schedule is more \textit{front loaded}:
Efficient frontier of optimal execution

We compare the performances of our solutions and Almgren Chriss’.

BARC - 1% of daily volume

BP - 1% of daily volume