Empirical Estimation of Market Impact and Calibration of Optimal Execution Strategies

Enzo Busseti

J.P. Morgan

October 27, 2011

Outline

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.

Impact propagator model

The price dynamics is given by[1]:

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

where η_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.

[1] J.P. Bouchaud, J. Kockelkoren, and M. Potters. *Random walks, liquidity molasses and critical response in financial markets*, Quant. Financ., 6(2):115123, 2006.

Generalizing to different time settings

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

- **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.

The estimation procedure

Two steps:

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

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

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

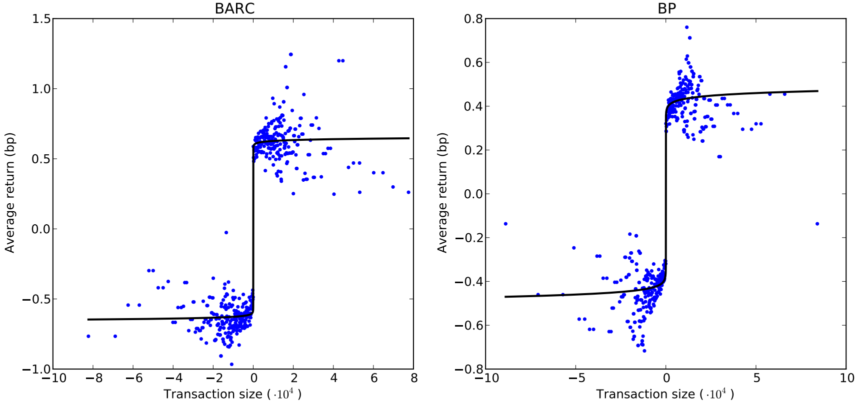
$$(p_{n+1} - p_n) = r_n = \sum_{k=0}^{n} [G_0(k+1) - G_0(k)] f(v_{n-k}) + \eta_n.$$

Data

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

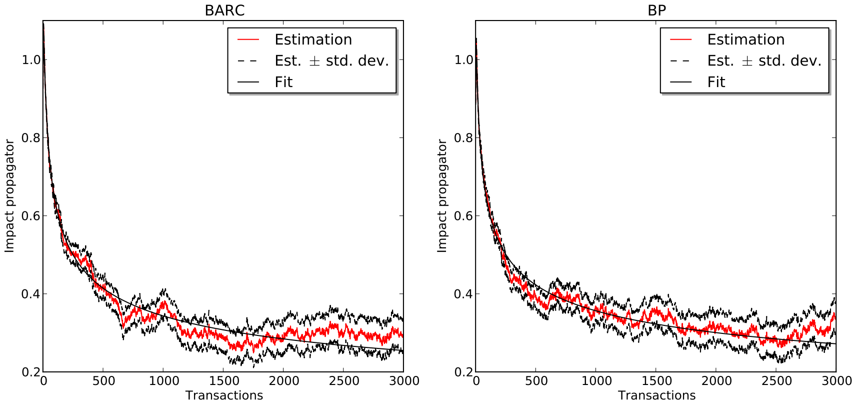
- AZN and VOD,
  - 32 months data from May, 2000 to December, 2002;
  - $\approx 10^3$ transactions per day.
- BP and BARC,
  - 4 months data from January 2011 to April 2011;
  - $\approx 10^4$ transactions per day.

Impact function $f(v)$ estimate, trade time



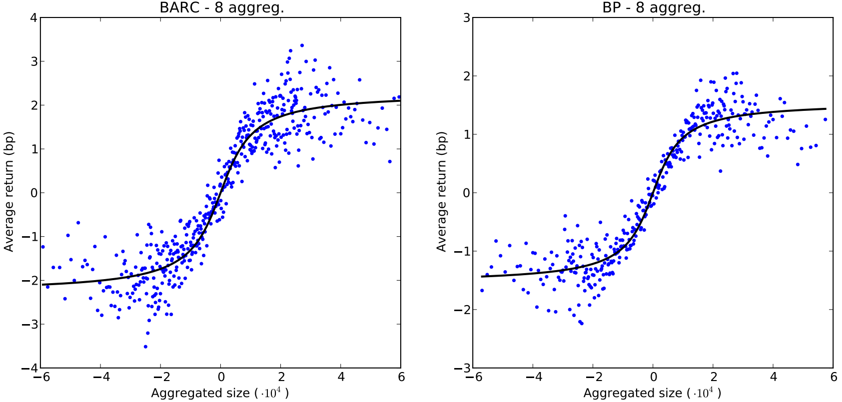
Fit: $f(v) = \theta \text{ sign}(v) \cdot |v|^{\phi}$ (a step function would be fine).

Propagator G_0 estimate, trade time



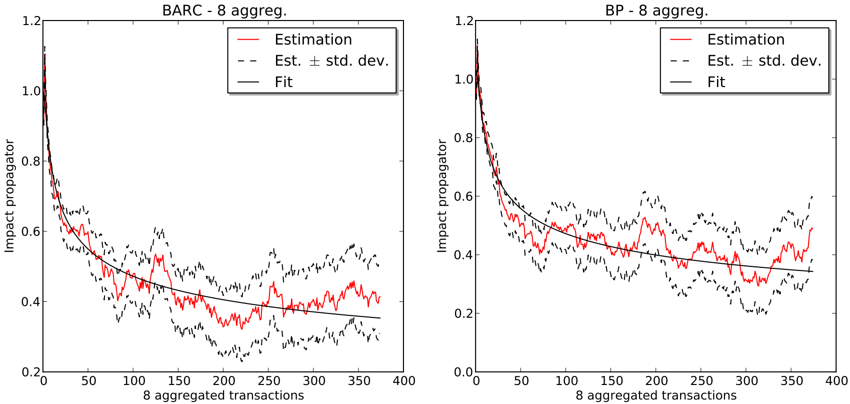
Fit $G_0^{fit}(l) = \frac{\Gamma_0}{(l_0^2 + l^2)^{\beta/2}}$.

Impact function $f(v)$ estimate, 8-aggregated trade time



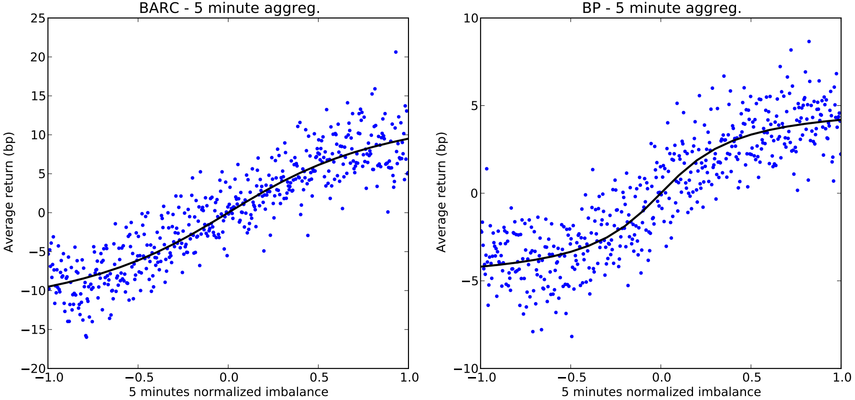
Fit: $f(v) = \theta \ \arctan(\phi v)$.

Propagator G_0 estimate, 8-aggregated trade time



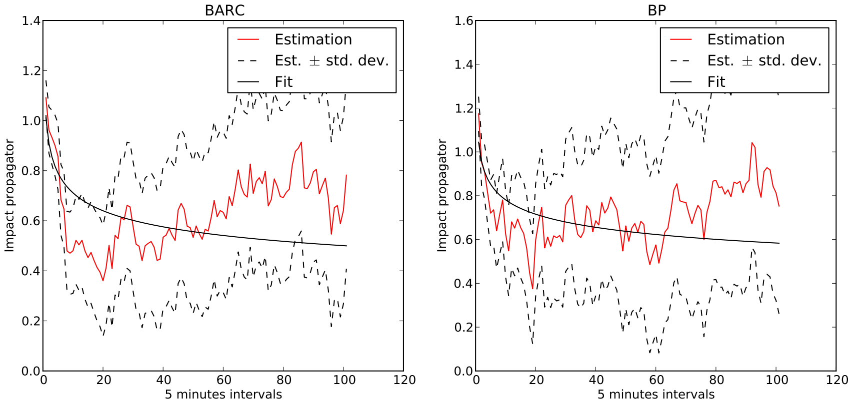
Fit $G_0^{fit}(l) = \frac{\Gamma_0}{(l_0^2 + l^2)^{\beta/2}}$.

Impact function $f(v)$ estimate, real time



Fit: $f(v) = \theta \ \arctan(\phi v)$.

Propagator G_0 estimate, real time



$$\text{Fit } G_0^{fit}(l) = \frac{\Gamma_0}{(l_0^2 + l^2)^{\beta/2}}.$$

Fraction of volatilty explained by the model

Estimated $R^2 = 1 - \frac{\sum \eta_n^2}{\sum r_n^2}$ coefficients

	R^2 coefficients of the linear regression				
Symbol	trade time	8 aggregated transactions	64 aggregated transactions	1 minute aggregation	5 minutes aggregation
AZN	0.224	0.245	0.304	0.192	0.204
VOD	0.257	0.329	0.440	0.209	0.292
BP	0.083	0.070	0.047	0.054	0.039
BARC	0.090	0.069	0.044	0.070	0.040

AZN and VOD belong to the 2000-2002 dataset, BP and BARC to the 2011 dataset.

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:

$$\mathbf{E}\left[C(\vec{v})\right].$$

Optimal execution with market impact propagator

With the propagator impact model we have:

$$\mathbf{E}[c(\vec{v})] = \sum_{n=0}^{N-1} v_n [\sum_{k=0}^{n} f(v_k) G_0(n-k)]$$

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

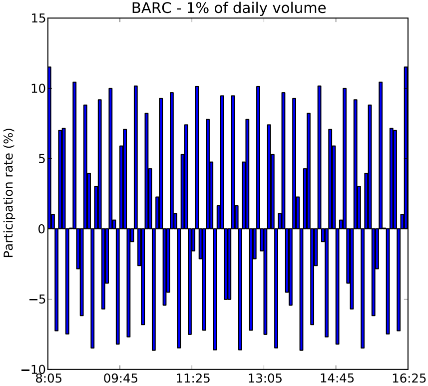
$$\mathcal{I}_{i,j} \equiv \begin{cases} \theta_i G_0(i-j) & i \geq j \\ 0 & i < j \end{cases}$$

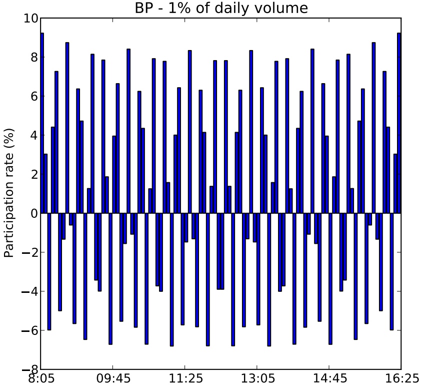
So that the execution costs are given by:

$$\mathbf{E}[c(\vec{v})] = \vec{v}^T \mathcal{I} \vec{v}.$$

Optimal schedule with the minimization procedure

We minimize analytically $\mathbf{E}[c(\vec{v})] = \vec{v}^T \mathcal{I} \vec{v}$ (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:

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

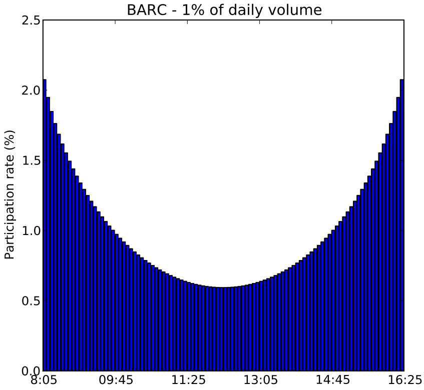
The problem is not linear anymore, and there is no analytical solution.

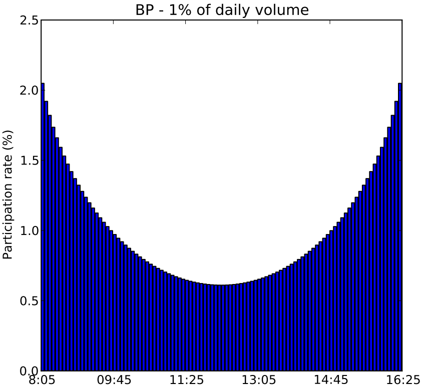
$$\Downarrow$$

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

	1% daily volume, expected impact costs (bp)		
Symbol	"flat" solution	"U-shaped" solution	(difference)
AZN	4.36	4.29	1.63%
VOD	9.82	9.76	0.61%
BP	2.36	2.33	1.29%
BARC	3.69	3.64	1.37%

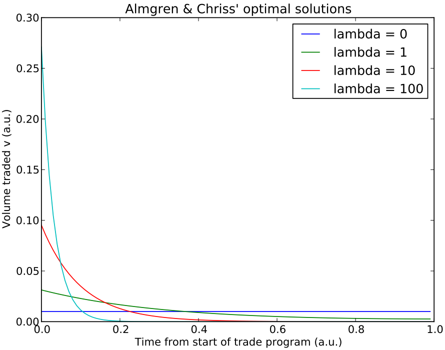
Optimal execution with risk aversion - Almgren & Chriss[2]

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

One finds the solutions to:

$$\min_{v_t} \left(\mathbf{E}[C(\vec{v})] + \lambda \mathbf{V}[C(\vec{v})]\right).$$

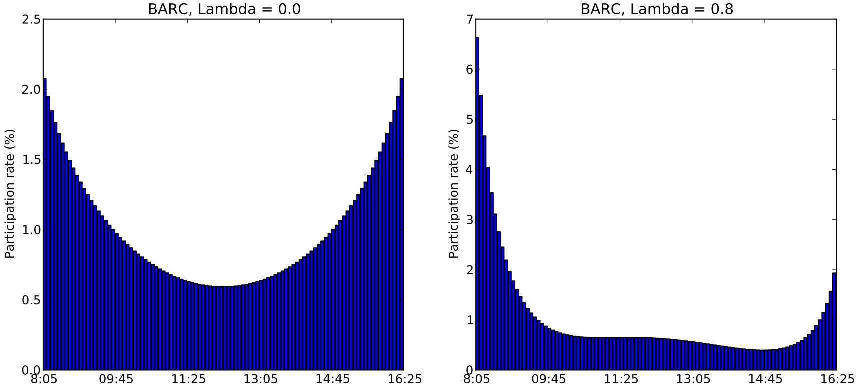
With the classic impact model (temporary + permanent) one obtains $\Rightarrow$



[2] R. Almgren and N. Chriss. Optimal execution of portfolio transactions. *J. Risk*, 3:540, 2001.

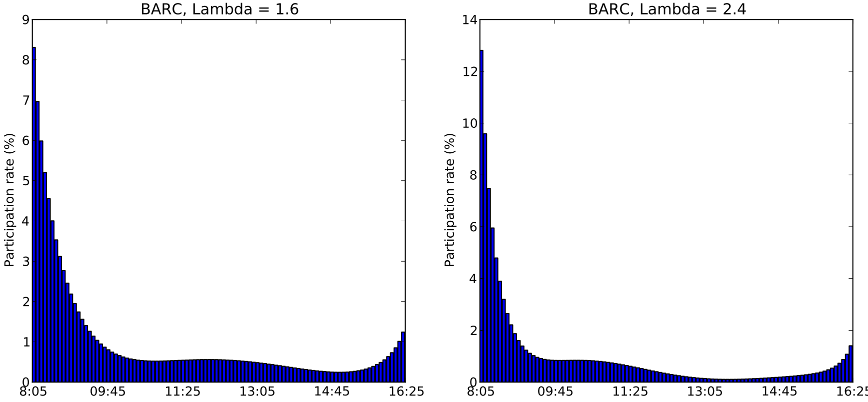
Optimal schedule with risk-aversion I

With our impact model and estimated parameters we solve numerically: $\min_{v_t}\left(\mathbf{E}[C(\vec{v})] + \delta\vec{1}^T|\vec{v}| + \lambda\mathbf{V}[C(\vec{v})]\right).$



Optimal schedule with risk-aversion II

With higher coefficients of risk-aversion λ the optimal schedule is more *front loaded*:



Efficient frontier of optimal execution

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