

# Optimal order execution

Jim Gatheral

(including joint work with Alexander Schied and Alla Slynko)



The City University of New York

JOIM Fall Conference, Boston, October 4, 2011

# References



[Almgren] Robert Almgren, Equity market impact, *Risk* July 2005, 57–62.



[Almgren and Chriss] Robert Almgren and Neil Chriss, Optimal execution of portfolio transactions, *Journal of Risk* 3 5–40 (2001).



[Alfonsi, Fruth and Schied] Aurélien Alfonsi, Antje Fruth and Alexander Schied, Optimal execution strategies in limit order books with general shape functions, *Quantitative Finance* 10(2) 143–157 (2010).



[Alfonsi, Schied and Slynko] Aurélien Alfonsi, Alexander Schied and Alla Slynko, Order book resilience, price manipulation, and the positive portfolio problem, [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1498514](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1498514) (2009).



[Forsyth et al.] P.A. Forsyth, J.S. Kennedy, S. T. Tse, and H. Windcliff, Optimal Trade Execution: A Mean - Quadratic-Variation Approach, University of Waterloo (2011).



[Gatheral] Jim Gatheral, No-dynamic-arbitrage and market impact, *Quantitative Finance* 10(7) 749–759 (2010).

# References



[Gatheral and Schied] Jim Gatheral and Alexander Schied, Optimal Trade Execution under Geometric Brownian Motion in the Almgren and Chriss Framework, *International Journal of Theoretical and Applied Finance* **14**(3) 353–368 (2011).



[Gatheral, Schied and Slynko] Jim Gatheral, Alexander Schied and Alla Slynko, Transient linear price impact and Fredholm integral equations, *Mathematical Finance* forthcoming  
[http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1531466](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1531466) (2011).



[Obizhaeva and Wang] Anna Obizhaeva and Jiang Wang, Optimal trading strategy and supply/demand dynamics *MIT working paper* (2005).



[Predoiu, Shaikhet and Shreve] Silviu Predoiu, Gennady Shaikhet and Steven Shreve, Optimal execution in a general one-sided limit-order book, *SIAM Journal on Finance Mathematics* **2** 183–212 (2011).



[Tóth et al.] Bence Tóth, Yves Lempérière, Cyril Deremble, Joachim de Lataillade, Julien Kockelkoren, and Jean-Philippe Bouchaud, Anomalous price impact and the critical nature of liquidity in financial markets, <http://arxiv.org/abs/1105.1694v1> (2011).



[Weiss] Alexander Weiss, Executing large orders in a microscopic market model, <http://arxiv.org/abs/0904.4131v2> (2010).

# Overview of this talk

- Statement of the optimal execution problem
- The Almgren-Chriss framework and 2001 model
  - Statically and dynamically optimal strategies
  - Model-dependence of optimality of the static solution
- The Obizhaeva and Wang model
- The Alfonsi and Schied model
- Price manipulation and existence of optimal strategies
- Transient linear price impact
- The square-root process
  - Optimal strategies and their properties

## Overview of execution algorithm design

Typically, an execution algorithm has three layers:

- The macrotrader
  - This highest level layer decides how to slice the order: when the algorithm should trade, in what size and for roughly how long.
- The microtrader
  - Given a slice of the order to trade (a child order), this level decides whether to place market or limit orders and at what price level(s).
- The smart order router
  - Given a limit or market order, which venue should this order be sent to?

In this lecture, we are concerned with the highest level of the algorithm: How to slice the order.

# Statement of the problem

- Given a model for the evolution of the stock price, we would like to find an optimal strategy for trading stock, the strategy that minimizes some cost function over all permissible strategies.
  - We will specialize to the case of stock liquidation where the initial position  $x_0 = X$  and the final position  $x_T = 0$ .
- A *static* strategy is one determined in advance of trading.
- A *dynamic* strategy is one that depends on the state of the market during execution of the order, *i.e.* on the stock price.
  - Delta-hedging is an example of a dynamic strategy. VWAP is an example of a static strategy.
- It will turn out, surprisingly, that in many models, a statically optimal strategy is also dynamically optimal.

# An observation from Predoiu, Shaikhet and Shreve

Suppose the cost associated with a strategy depends on the stock price only through the term

$$\int_0^T S_t dx_t.$$

with  $S_t$  a martingale. Integration by parts gives

$$\mathbb{E} \left[ \int_0^T S_t dx_t \right] = \mathbb{E} \left[ S_T x_T - S_0 x_0 - \int_0^T x_t dS_t \right] = -S_0 X$$

which is independent of the trading strategy and we may proceed as if  $S_t = 0$ .

Quote from [Predoiu, Shaikhet and Shreve]

“...there is no longer a source of randomness in the problem. Consequently, without loss of generality we may restrict the search for an optimal strategy to nonrandom functions of time”.

# Corollary

- This observation enables us to easily determine whether or not a statically optimal strategy will be dynamically optimal.
  - In particular, if the price process is of the form

$$S_t = S_0 + \text{impact of prior trading} + \text{noise},$$

and if there is no risk term, a statically optimal strategy will be dynamically optimal.

- If there is a risk term independent of the current stock price, a statically optimal strategy will again be dynamically optimal.
- It follows that the statically optimal strategy is dynamically optimal in the following models:
  - Almgren and Chriss
  - Almgren (2005)
  - Obizhaeva and Wang



# Almgren and Chriss

[Almgren and Chriss] model market impact and slippage as follows.  
The stock price  $S_t$  evolves as

$$dS_t = \sigma dZ_t$$

and the price  $\tilde{S}_t$  at which we transact is given by

$$\tilde{S}_t = S_t + \eta v_t$$

where  $v_t := -\dot{x}_t$  is the rate of trading.

- In this model, temporary market impact decays instantaneously and does not affect the market price  $S_t$ .

# The statically optimal strategy

The statically optimal strategy  $v_s$  is the one that minimizes the cost function

$$\mathcal{C} = \mathbb{E} \left[ \int_0^T \tilde{S}_t v_t dt \right] = \mathbb{E} \left[ \int_0^T (S_t + \eta v_t) v_t ds \right] = \eta \int_0^T v_t^2 dt$$

again with  $v_t = -\dot{x}_t$ .

The Euler-Lagrange equation is then

$$\partial_t v_t = -\partial_{t,t} x_t = 0$$

with boundary conditions  $x_0 = X$  and  $x_T = 0$  and the solution is obviously

$$v_t = \frac{X}{T}; \quad x_t = X \left( 1 - \frac{t}{T} \right)$$

## Adding a risk term

[Almgren and Chriss] add a risk term that penalizes the variance of the trading cost.

$$\text{Var}[C] = \text{Var} \left[ \int_0^T x_t dS_t \right] = \sigma^2 \int_0^T x_t^2 dt$$

The expected risk-adjusted cost of trading is then given by

$$C = \eta \int_0^T \dot{x}_t^2 dt + \lambda \sigma^2 \int_0^T x_t^2 dt$$

for some price of risk  $\lambda$ .

- Note the analogies to physics and portfolio theory.
  - The first term looks like kinetic energy and the second term like potential energy.
  - The expression looks like the objective in mean-variance portfolio optimization.

The Euler-Lagrange equation becomes

$$\ddot{x} - \kappa^2 x = 0$$

with

$$\kappa^2 = \frac{\lambda \sigma^2}{\eta}$$

The solution is a linear combination of terms of the form  $e^{\pm \kappa t}$  that satisfies the boundary conditions  $x_0 = X$ ,  $x_T = 0$ . The solution is then

$$x(t) = X \frac{\sinh \kappa(T - t)}{\sinh \kappa T}$$

Once again, the statically optimal solution is dynamically optimal.

# What happens if we change the risk term?

Suppose we penalize average VaR instead of variance. This choice of risk term has the particular benefit of being linear in the position size. The expected risk-adjusted cost of trading is then given by

$$C = \eta \int_0^T \dot{x}_t^2 dt + \lambda \sigma \int_0^T x_t dt$$

for some price of risk  $\lambda$ .

The Euler-Lagrange equation becomes

$$\ddot{x} - A = 0$$

with

$$A = \frac{\lambda \sigma}{2 \eta}$$

The solution is a quadratic of the form  $A t^2/2 + B t + C$  that satisfies the boundary conditions  $x_0 = X$ ,  $x_T = 0$ . The solution is then

$$x(t) = \left( X - \frac{A T}{2} t \right) \left( 1 - \frac{t}{T} \right) \quad (1)$$

In contrast to the previous case where the cost function is monotonic decreasing in the trading rate and the optimal choice of liquidation time is  $\infty$ , in this case, we can compute an optimal liquidation time. When  $T$  is optimal, we have

$$\frac{\partial C}{\partial T} \propto \dot{x}_T + A x_T = 0$$

from which we deduce that  $\dot{x}_T = 0$ .

Substituting into (1) and solving for the optimal time  $T^*$  gives

$$T^* = \sqrt{\frac{2X}{A}}$$

With this optimal choice  $T = T^*$ , the optimal strategy becomes

$$\begin{aligned} x(t) &= X \left(1 - \frac{t}{T}\right)^2 \\ u(t) &= -\dot{x}(t) = 2X \left(1 - \frac{t}{T}\right) \end{aligned}$$

One can verify that the static strategy is dynamically optimal, independent of the stock price.

# ABM vs GBM

- One of the reasons that the statically optimal strategy is dynamically optimal is that the stock price process is assumed to be arithmetic Brownian motion (ABM).
- If for example geometric Brownian motion (GBM) is assumed, the optimal strategy depends on the stock price.
- How dependent is the optimal strategy on dynamical assumptions for the underlying stock price process?



# Forsyth et al.

- [Forsyth et al.] solve the HJB equation numerically under geometric Brownian motion with variance as the risk term so that the (random) cost is given by

$$\mathcal{C} = \eta \int_0^T \dot{x}_t^2 dt + \lambda \sigma^2 \int_0^T S_t^2 x_t^2 dt$$

- The efficient frontier is found to be virtually identical to the frontier computed in the arithmetic Brownian motion case.
- The problem of finding the optimal strategy is ill-posed; many strategies lead to almost the same value of the cost function.
- It is optimal to trade faster when the stock price is high so as to reduce variance. The optimal strategy is aggressive-in-the-money when selling stock and passive-in-the-money when buying stock.

# Gatheral and Schied

[Gatheral and Schied] take time-averaged VaR as the risk term so that

$$C(T, X, S_0) = \inf_{v \in \mathcal{G}} \mathbb{E} \left[ \int_0^T v_t^2 dt + \lambda \int_0^T S_t x_t dt \right], \quad (2)$$

where  $\mathcal{G}$  is the set of admissible strategies.

$C(T, X, S)$  should then satisfy the following Hamilton-Jacobi-Bellman PDE:

$$C_T = \frac{1}{2} \sigma^2 S^2 C_{SS} + \lambda S X + \inf_{v \in \mathbb{R}} (v^2 - v C_X). \quad (3)$$

with initial condition

$$\lim_{T \downarrow 0} C(T, X, S) = \begin{cases} 0 & \text{if } X = 0, \\ +\infty & \text{if } X \neq 0. \end{cases} \quad (4)$$

# The optimal strategy under GBM

Solving the HJB equation explicitly gives

## Theorem

*The unique optimal trade execution strategy attaining the infimum in (2) is*

$$x_t^* = \left( \frac{T-t}{T} \right) \left[ X - \frac{\lambda T}{4} \int_0^t S_u du \right] \quad (5)$$

*Moreover, the value of the minimization problem in (2) is given by*

$$\begin{aligned} C &= \mathbb{E} \left[ \int_0^T \{ (\dot{x}_t^*)^2 + \lambda x_t^* S_t \} dt \right] \\ &= \frac{X^2}{T} + \frac{1}{2} \lambda T X S_0 - \frac{\lambda^2}{8 \sigma^6} S_0^2 \left( e^{\sigma^2 T} - 1 - \sigma^2 T - \frac{1}{2} \sigma^4 T^2 \right). \end{aligned}$$

# The optimal strategy under ABM

If we assume ABM,  $S_t = S_0 (1 + \sigma W_t)$ , instead of GBM, the risk term becomes

$$\hat{\lambda} S_0 \int_0^T x_t dt. \quad (6)$$

As we already showed, the optimal strategy under ABM is just the static version of the dynamic strategy (5) obtained by replacing  $S_t$  with its expectation  $\mathbb{E}[S_t] = S_0$ , a strategy qualitatively similar to the Almgren-Chriss optimal strategy.

# Comparing optimal strategies under ABM and GBM

As before, define the characteristic timescale

$$T^* = \sqrt{\frac{4X}{\lambda S_0}}$$

and choose the liquidation time  $T$  to be  $T^*$ .

With  $T = T^*$ , the optimal trading rate under ABM becomes

$$v^A(t) = \frac{x_t}{T-t} + \frac{X}{T^2} (T-t) = \frac{2X}{T} \left(1 - \frac{t}{T}\right) \quad (7)$$

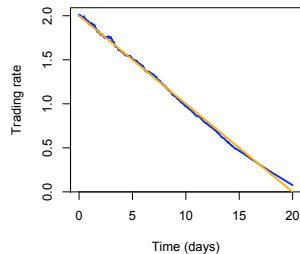
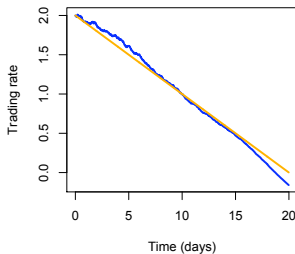
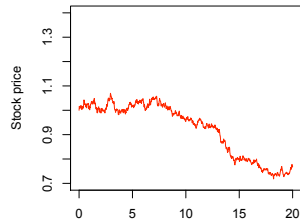
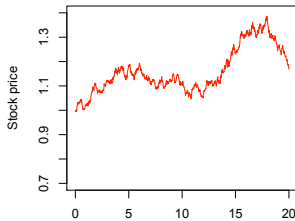
and the optimal trading rate under GBM becomes

$$v^G(t) = \frac{x_t}{T-t} + \frac{X}{T^2} \frac{S_t}{S_0} (T-t). \quad (8)$$

# Comparing optimal strategies under ABM and GBM

In the following slide:

- The upper plots show rising and falling stock price scenarios respectively; the trading period is 20 days and daily volatility is 4%.
- The lower plots show the corresponding optimal trading rates from (7) and (8); the optimal trading rate under ABM is in orange and the optimal trading rate under GBM is in blue.
- Even with such extreme parameters and correspondingly extreme changes in stock price, the differences in optimal trading rates are minimal.



# Remarks

- For reasonable values of  $\sigma^2 T \ll 1$ , there is almost no difference in expected costs and risks between the optimal strategies under ABM and GBM assumptions.
- Intuitively, although the optimal strategy is stock price-dependent under GBM assumptions but not under ABM assumptions, when  $\sigma^2 T \ll 1$ , the difference in optimal frontiers is tiny because the stock-price  $S_t$  cannot diffuse very far away from  $S_0$  in the short time available.
- Equivalently, as in the plots, there can only be a small difference in optimal trading rates under the two assumptions.



# Practical comments

- It's not clear what the price of risk should be.
- More often than not, a trader wishes to complete an execution before some final time and otherwise just wants to minimize expected execution cost.
  - In Almgren-Chriss style models, the optimal strategy is just VWAP (trading at constant rate).
- From now on, we will drop the risk term and the dynamics we will consider will ensure that the statically optimal solution is dynamically optimal.

# Obizhaeva and Wang 2005

In the [Obizhaeva and Wang] model,

$$S_t = S_0 + \eta \int_0^t u_s e^{-\rho(t-s)} ds + \int_0^t \sigma dZ_s \quad (9)$$

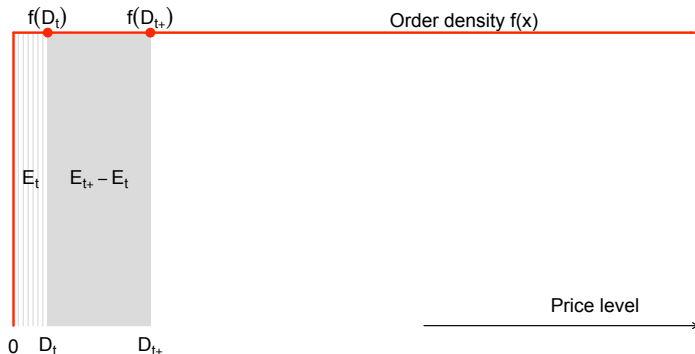
with  $u_t = -\dot{x}_t$ .

- Market impact is linear in the rate of trading but in contrast to Almgren and Chriss, market impact decays exponentially with some non-zero half-life.

The expected cost of trading becomes:

$$\mathcal{C} = \eta \int_0^T u_t dt \int_0^t u_s \exp\{-\rho(t-s)\} ds$$

# Obizhaeva Wang order book process



When a trade of size  $\xi$  is placed at time  $t$ ,

$$E_t \mapsto E_{t+} = E_t + \xi$$

$$D_t = \eta E_t \mapsto D_{t+} = \eta E_{t+} = \eta (E_t + \xi)$$

When the trading policy is statically optimal, the Euler-Lagrange equation applies:

$$\frac{\partial}{\partial t} \frac{\delta \mathcal{C}}{\delta u_t} = 0$$

where  $u_t = \dot{x}_t$ . Functionally differentiating  $\mathcal{C}$  with respect to  $u_t$  gives

$$\frac{\delta \mathcal{C}}{\delta u_t} = \int_0^t u_s e^{-\rho(t-s)} ds + \int_t^T u_s e^{-\rho(s-t)} ds = A \quad (10)$$

for some constant  $A$ . Equation (10) may be rewritten as

$$\int_0^T u_s e^{-\rho|t-s|} ds = A$$

which is a Fredholm integral equation of the first kind (see [Gatheral, Schied and Slynko]).

Now substitute

$$u_s = \delta(s) + \rho + \delta(s - T)$$

into (10) to obtain

$$\frac{\delta \mathcal{C}}{\delta u_t} = e^{-\rho t} + (1 - e^{-\rho t}) = 1$$

The optimal strategy consists of a block trade at time  $t = 0$ , continuous trading at the rate  $\rho$  over the interval  $(0, T)$  and another block trade at time  $t = T$ .

Consider the volume impact process  $E_t$ . The initial block-trade causes

$$0 = E_0 \mapsto E_{0+} = 1$$

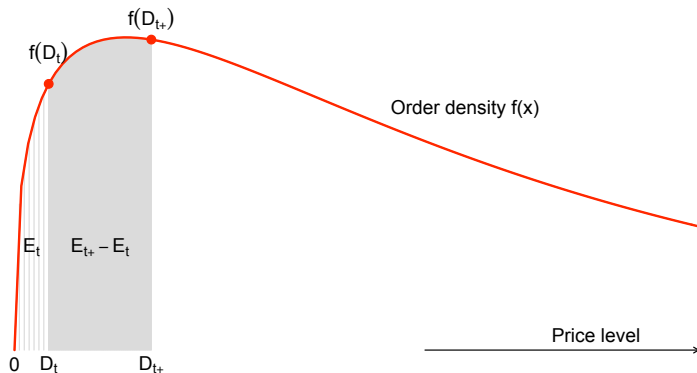
According to the assumptions of the model, the volume impact process reverts exponentially so

$$E_t = E_{0+} e^{-\rho t} + \rho \int_0^t e^{-\rho(t-s)} ds = 1$$

*i.e.* the volume impact process is constant when the trading strategy is optimal.



# Schematic of the model



When a trade of size  $\xi$  is placed at time  $t$ ,

$$E_t \mapsto E_{t+} = E_t + \xi$$

$$D_t = F^{-1}(E_t) \mapsto D_{t+} = F^{-1}(E_{t+}) = F^{-1}(E_t + \xi)$$



# Optimal liquidation strategy in the AS model

The cost of trade execution in the AS model is given by:

$$\mathcal{C} = \int_0^T v_t F^{-1}(E_t) dt + \sum_{t \leq T} [H(E_{t+}) - H(E_t)] \quad (11)$$

where

$$E_t = \int_0^t u_s e^{-\rho(t-s)} ds$$

is the volume impact process and

$$H(x) = \int_0^x F^{-1}(x) dx$$

gives the cost of executing an instantaneous block trade of size  $x$ .

Consider the ansatz  $u_t = \xi_0 \delta(t) + \xi_0 \rho + \xi_T \delta(T - t)$ . For  $t \in (0, T)$ , we have  $E_t = E_0 = \xi_0$ , a constant. With this choice of  $u_t$ , we would have

$$\begin{aligned} \mathcal{C}(X) &= F^{-1}(\xi_0) \int_0^T v_t dt + [H(E_{0+}) - H(E_0)] + [H(E_T) - H(E_{T-})] \\ &= F^{-1}(\xi_0) \xi_0 \rho T + H(\xi_0) + [H(\xi_0 + \xi_T) - H(\xi_0)] \\ &= F^{-1}(\xi_0) \xi_0 \rho T + H(X - \rho \xi_0 T) \end{aligned}$$

Differentiating this last expression gives us the condition satisfied by the optimal choice of  $\xi_0$ :

$$F^{-1}(X - \rho \xi_0 T) = F^{-1}(\xi_0) + F^{-1}'(\xi_0) \xi_0$$

or equivalently

$$F^{-1}(\xi_0 + \xi_T) = F^{-1}(\xi_0) + F^{-1}'(\xi_0) \xi_0$$

Functionally differentiating  $\mathcal{C}$  with respect to  $u_t$  gives

$$\begin{aligned}\frac{\delta \mathcal{C}}{\delta u_t} &= F^{-1}(E_t) + \int_t^T u_s F^{-1'}(E_s) \frac{\delta E_s}{\delta u_t} ds \\ &= F^{-1}(E_t) + \int_t^T u_s F^{-1'}(E_s) e^{-\rho(s-t)} ds \quad (12)\end{aligned}$$

The first term in (12) represents the marginal cost of new quantity at time  $t$  and the second term represents the marginal extra cost of future trading.

With our ansatz, and a careful limiting argument, we obtain

$$\begin{aligned}\frac{\delta \mathcal{C}}{\delta u_t} &= F^{-1}(\xi_0) + \xi_0 F^{-1'}(\xi_0) \left[ 1 - e^{-\rho(T-t)} \right] \\ &\quad + e^{-\rho(T-t)} \left[ F^{-1}(\xi_T + \xi_0) - F^{-1}(\xi_0) \right]\end{aligned}$$

Imposing our earlier condition on  $\xi_T$  gives

$$\begin{aligned}\frac{\delta \mathcal{C}}{\delta u_t} &= F^{-1}(\xi_0) + \xi_0 F^{-1'}(\xi_0) \left[ 1 - e^{-\rho(T-t)} \right] \\ &\quad + e^{-\rho(T-t)} \xi_0 F^{-1'}(\xi_0) \\ &= F^{-1}(\xi_0) + \xi_0 F^{-1'}(\xi_0)\end{aligned}$$

which is constant, demonstrating (static) optimality.

### Example

With  $F^{-1}(x) = \sqrt{x}$ ,

$$\sqrt{\xi_0 + \xi_T} = F^{-1}(\xi_0 + \xi_T) = F^{-1}(\xi_0) + F^{-1'}(\xi_0) \xi_0 = \sqrt{\xi_0} + \frac{1}{2} \sqrt{\xi_0}$$

which has the solution  $\xi_T = \frac{5}{4} \xi_0$ .

# Generalization

Alexander Weiss [Weiss] and then Predoiu, Shaikhet and Shreve [Predoiu, Shaikhet and Shreve] have shown that the bucket-shaped strategy is optimal under more general conditions than exponential resiliency. Specifically, if resiliency is a function of  $E_t$  (or equivalently  $D_t$ ) only, the optimal strategy has a block trades at inception and completion and continuous trading at a constant rate in-between.

# Optimality and price manipulation

- For all of the models considered so far, there was an optimal strategy.
- The optimal strategy always involved trades of the same sign. So no sells in a buy program, no buys in a sell program.
- It turns out (see [Gatheral]) that we can write down models for which price manipulation is possible.
- In such cases, a round-trip trade can generate cash on average.
  - You would want to repeat such a trade over and over.
  - There would be no optimal strategy.

# Linear transient market impact

The price process assumed in [Gatheral] is

$$S_t = S_0 + \int_0^t f(v_s) G(t-s) ds + \text{noise}$$

In [Gatheral, Schied and Slynko], this model is on the one hand extended to explicitly include discrete optimal strategies and on the other hand restricted to the case of linear market impact. When the admissible strategy  $X$  is used, the price  $S_t$  is given by

$$S_t = S_t^0 + \int_{\{s < t\}} G(t-s) dX_s, \quad (13)$$

and the expected cost of liquidation is given by

$$\mathcal{C}(X) := \frac{1}{2} \int \int G(|t-s|) dX_s dX_t. \quad (14)$$

# Condition for no price manipulation

## Definition (Huberman and Stanzl)

A *round trip* is an admissible strategy with  $X_0 = 0$ . A *price manipulation strategy* is a round trip with strictly negative expected costs.

## Proposition (Bochner)

$G(X) \geq 0$  for all admissible strategies  $X$  if and only if  $G(|\cdot|)$  can be represented as the Fourier transform of a positive finite Borel measure  $\mu$  on  $\mathbb{R}$ , i.e.,

$$G(|x|) = \int e^{ixz} \mu(dz).$$



# First order condition

## Theorem

*Suppose that  $G$  is positive definite. Then  $X^*$  minimizes  $\mathcal{C}(\cdot)$  if and only if there is a constant  $\lambda$  such that  $X^*$  solves the generalized Fredholm integral equation*

$$\int G(|t - s|) dX_s^* = \lambda \quad \text{for all } t \in \mathbb{T}. \quad (15)$$

*In this case,  $\mathcal{C}(X^*) = \frac{1}{2} \lambda x$ . In particular,  $\lambda$  must be nonzero as soon as  $G$  is strictly positive definite and  $x \neq 0$ .*

# Transaction-triggered price manipulation

## Definition (Alfonsi, Schied, Slynko (2009))

A market impact model admits *transaction-triggered price manipulation* if the expected costs of a sell (buy) program can be decreased by intermediate buy (sell) trades.

As discussed in [Alfonsi, Schied and Slynko], transaction-triggered price manipulation can be regarded as an additional model irregularity that should be excluded. Transaction-triggered price manipulation can exist in models that do not admit standard price manipulation in the sense of Huberman and Stanzl definition.

# Condition for no transaction-triggered price manipulation

## Theorem

*Suppose that the decay kernel  $G(\cdot)$  is convex, satisfies  $\int_0^1 G(t) dt < \infty$  and that the set of admissible strategies is nonempty. Then there exists a unique admissible optimal strategy  $X^*$ . Moreover,  $X_t^*$  is a monotone function of  $t$ , and so there is no transaction-triggered price manipulation.*

## Remark

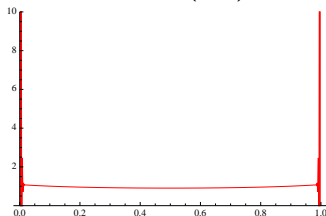
If  $G$  is not convex in a neighborhood of zero, there is transaction-triggered price manipulation.

# An instructive example

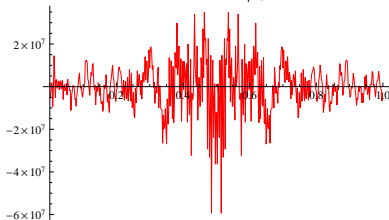
We solve a discretized version of the Fredholm equation (with 512 time points) for two similar decay kernels:

$$G_1(\tau) = \frac{1}{(1+t)^2}; \quad G_2(\tau) = \frac{1}{1+t^2}$$

$G_1(\cdot)$  is convex, but  $G_2(\cdot)$  is concave near  $\tau = 0$  so there should be a unique optimal strategy with  $G_1(\cdot)$  as a choice of kernel but there should be transaction-triggered price manipulation with  $G_2(\cdot)$  as the choice of decay kernel.

$$G_1(\tau) = \frac{1}{(1+t)^2}$$


$$G_2(\tau) = \frac{1}{1+t^2}$$



In the left hand figure, we observe block trades at  $t = 0$  and  $t = 1$  with continuous (nonconstant) trading in  $(0, 1)$ . In the right hand figure, we see numerical evidence that the optimal strategy does not exist.

Now we give some examples of the optimal strategy under linear transient market impact with choices of kernel that preclude transaction-triggered price manipulation.

## Example I: Linear market impact with exponential decay

$G(\tau) = e^{-\rho\tau}$  and the optimal strategy  $u(s)$  solves

$$\int_0^T u(s) e^{-\rho|t-s|} ds = \text{const.}$$

We already derived the solution which is

$$u(s) = A \{ \delta(t) + \rho + \delta(T - t) \}$$

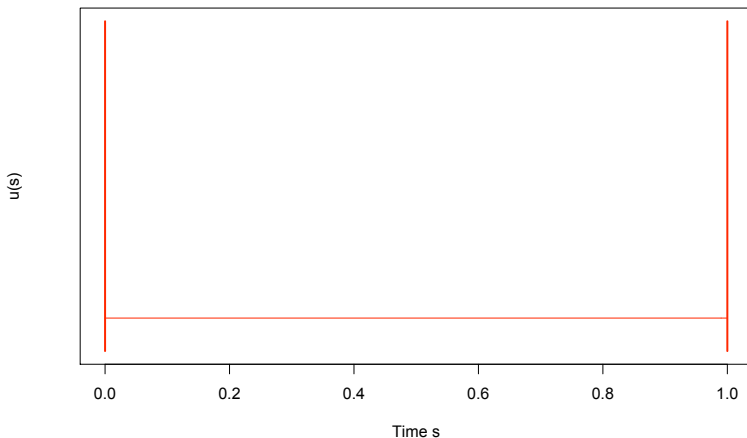
The normalizing factor  $A$  is given by

$$\int_0^T u(t) dt = X = A (2 + \rho T)$$

The optimal strategy consists of block trades at  $t = 0$  and  $t = T$  and continuous trading at the constant rate  $\rho$  between these two times.

# Schematic of optimal strategy

The optimal strategy with  $\rho = 0.1$  and  $T = 1$





## Example II: Linear market impact with power-law decay

$G(\tau) = \tau^{-\gamma}$  and the optimal strategy  $u(s)$  solves

$$\int_0^T \frac{u(s)}{|t-s|^\gamma} ds = \text{const.}$$

The solution is

$$u(s) = \frac{A}{[s(T-s)]^{(1-\gamma)/2}}$$

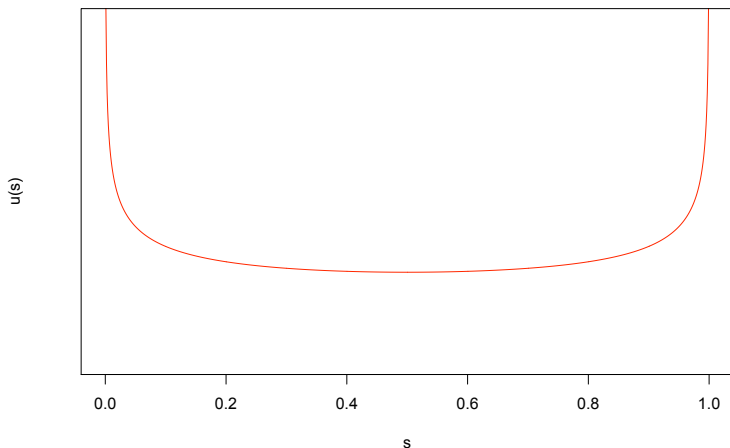
The normalizing factor  $A$  is given by

$$\int_0^T u(t) dt = X = A \sqrt{\pi} \left(\frac{T}{2}\right)^\gamma \frac{\Gamma\left(\frac{1+\gamma}{2}\right)}{\Gamma\left(1+\frac{\gamma}{2}\right)}$$

The optimal strategy is absolutely continuous with no block trades. However, it is singular at  $t = 0$  and  $t = T$ .

## Schematic of optimal strategy

The red line is a plot of the optimal strategy with  $T = 1$  and  $\gamma = 1/2$ .



## Example III: Linear market impact with linear decay

$G(\tau) = (1 - \rho \tau)^+$  and the optimal strategy  $u(s)$  solves

$$\int_0^T u(s) (1 - \rho |t - s|)^+ ds = \text{const.}$$

Let  $N := \lfloor \rho T \rfloor$ , the largest integer less than or equal to  $\rho T$ . Then

$$u(s) = A \sum_{i=0}^N \left(1 - \frac{i}{N+1}\right) \left\{ \delta\left(s - \frac{i}{\rho}\right) + \delta\left(T - s - \frac{i}{\rho}\right) \right\}$$

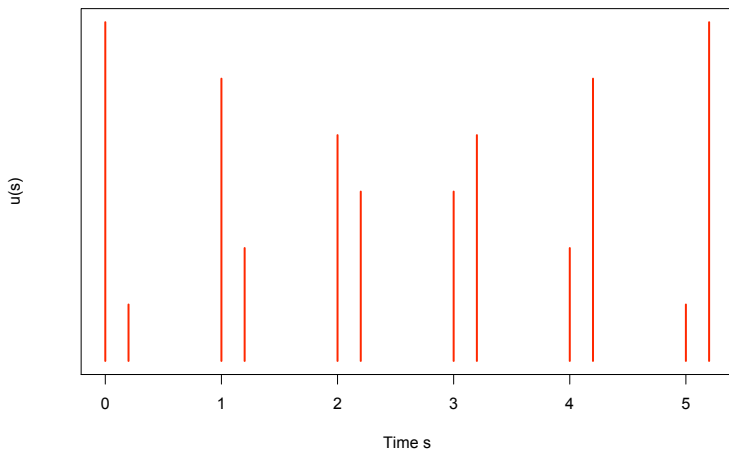
The normalizing factor  $A$  is given by

$$\int_0^T u(t) dt = X = A \sum_{i=0}^N 2 \left(1 - \frac{i}{N+1}\right) = A(2 + N)$$

The optimal strategy consists only of block trades with no trading between blocks.

# Schematic of optimal strategy

Positions and relative sizes of the block trades in the optimal strategy with  $\rho = 1$  and  $T = 5.2$  (so  $N = \lfloor \rho T \rfloor = 5$ ).



# Nonlinear transient market impact

- We know that the reaction of market price to quantity is in general nonlinear.
  - Concave for small quantity and convex for large quantity.
- We also know that the market price reverts after completion of a meta-order (using VWAP say).
- What is the optimal strategy under nonlinear transient market impact?

# The square-root formula for market impact

- For many years, traders have used the simple sigma-root-liquidity model described for example by Grinold and Kahn in 1994.
- Software incorporating this model includes:
  - Salomon Brothers, StockFacts Pro since around 1991
  - Barra, Market Impact Model since around 1998
  - Bloomberg, TCA function since 2005
- The model is always of the rough form

$$\Delta P = \text{Spread cost} + \alpha \sigma \sqrt{\frac{Q}{V}}$$

where  $\sigma$  is daily volatility,  $V$  is daily volume,  $Q$  is the number of shares to be traded and  $\alpha$  is a constant pre-factor of order one.

# Empirical question

So traders and trading software have been using the square-root formula to provide a pre-trade estimate of market impact for a long time.

## Empirical question

Is the square-root formula empirically verified?

# Impact of proprietary metaorders (from Tóth et al.)

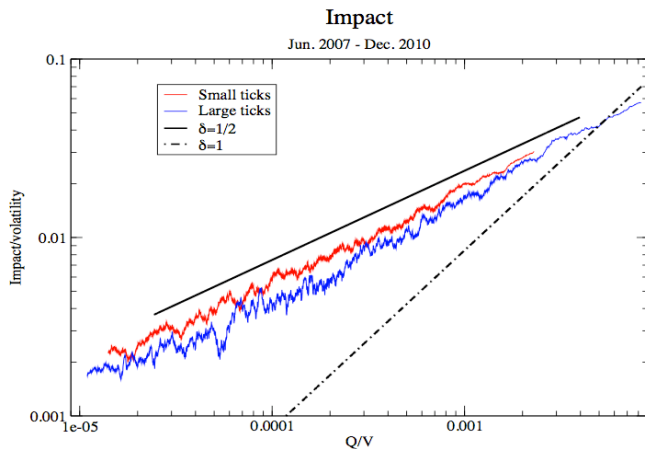


Figure 1: Log-log plot of the volatility-adjusted price impact vs the ratio  $Q/V$



# Notes on Figure 1

- In Figure 1 which is taken from [Tóth et al.], we see the impact of metaorders for CFM<sup>1</sup> proprietary trades on futures markets, in the period June 2007 to December 2010.
  - Impact is measured as the average execution shortfall of a meta-order of size  $Q$ .
  - The sample studied contained nearly 500,000 trades.
- We see that the square-root market impact formula is verified empirically for meta-orders with a range of sizes spanning two to three orders of magnitude!

---

<sup>1</sup>Capital Fund Management (CFM) is a large Paris-based hedge fund.

# An explanation for the square-root formula

- In [Tóth et al.], the authors present an argument which says that if latent supply and demand is linear in price over some reasonable range of prices, market impact should be square-root.
- The condition for linearity of latent supply and demand over a range of prices is simply that submitters of buy and sell meta-orders should be insensitive to price over this range.
  - We need also to assume that high frequency traders have no net effect on the latent supply and demand schedule.
- It would then seem reasonable to suppose that latent supply and demand should be linear over a range of prices  $\sim \sigma \sqrt{T}$  where  $T$  is the average life of a meta-order.
  - A distribution of meta-order durations would give rise to a concave latent supply/demand function and a market impact function with an exponent greater than  $\frac{1}{2}$  as is indeed observed empirically.

# Some implications of the square-root formula

- The square-root formula refers only to the size of the trade relative to daily volume.
- It does not refer to for example:
  - The rate of trading
  - How the trade is executed
  - The capitalization of the stock
- Surely impact must be higher if trading is very aggressive?
  - The database of trades only contains sensible trades with reasonable volume fractions.
  - Were we to look at very aggressive trades, we would indeed find that the square-root formula breaks down.

# A nonlinear transient market impact model

Once again, the price process assumed in [Gatheral] is

$$S_t = S_0 + \int_0^t f(v_s) G(t-s) ds + \text{noise} \quad (16)$$

- The instantaneous impact of a trade at time  $s$  is given by  $f(v_s)$  – some function of the rate of trading.
- A proportion  $G(t-s)$  of this initial impact is still felt at time  $t > s$ .

# The square-root model

Consider the following special case of (16) with  $f(v) = \frac{3}{4}\sigma \sqrt{v/V}$  and  $G(\tau) = 1/\sqrt{\tau}$ :

$$S_t = S_0 + \frac{3}{4} \sigma \int_0^t \sqrt{\frac{v_s}{V}} \frac{ds}{\sqrt{t-s}} + \text{noise} \quad (17)$$

which we will call the *square-root process*.

It is easy to verify that under the square-root process, the expected cost of a VWAP execution is given by the square-root formula for market impact:

$$\frac{\mathcal{C}}{X} = \sigma \sqrt{\frac{X}{V}} \quad (18)$$

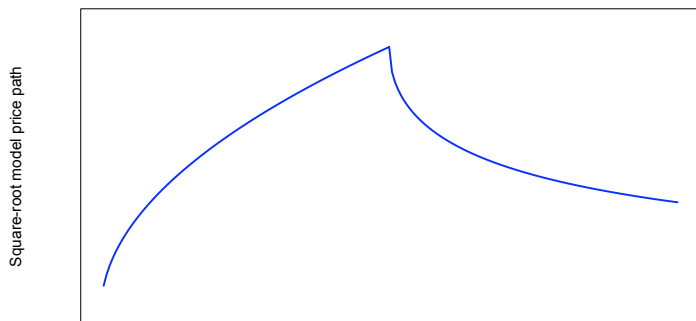
- Of course, that doesn't mean that the square-root process is the true underlying process!

# The optimal strategy under the square-root process

- Because  $f(\cdot)$  is concave, an optimal strategy does not exist in this case.
  - It is possible to drive the expected cost of trading to zero by increasing the number of slices and decreasing the duration of each slice.
  - To be more realistic,  $f(v)$  must be convex for large  $v$  and in this case, an optimal strategy does exist that involves trading in bursts, usually more than two.

# Price path in the square-root model

**Figure 2:** The expected path of the market price during and after execution of a VWAP order in the square-root model.



- The optimal strategy does not exist in this model.

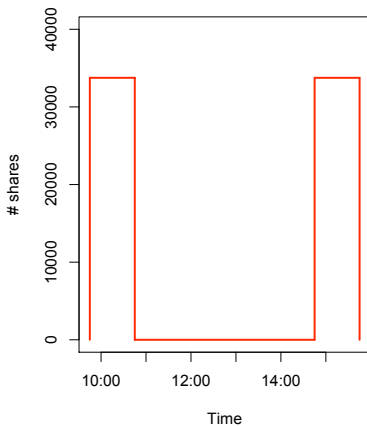
# Potential cost savings from optimal scheduling

- To estimate potential savings from optimal scheduling, assume that the square-root process (17) is correct and consider a one-day order to sell 540,000 shares of IBM.
  - Daily volatility is assumed to be 2% and daily volume to be 6 million shares.
  - We consider liquidation starting at 09:45 and ending at 15:45 with child orders lasting 15 minutes.
- Because we are not confident in the square-root model for high volume fractions, we constrain volume fraction to be no greater than 25%.
- We compare the costs of VWAP, a two-slice bucket-like strategy and a quasi-optimal strategy that consists of seven roughly equal slices.
  - The quasi-optimal strategy consists of bursts of trading separated by periods of non-trading.

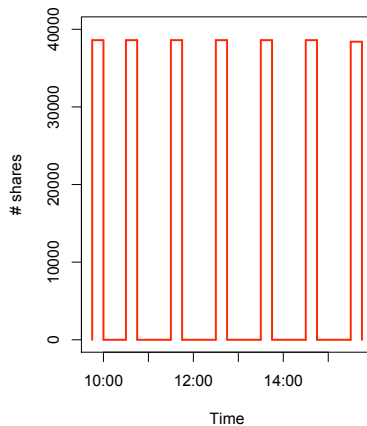


# Stock trading schedules

## Bucket-like schedule



## Quasi-optimal schedule



# Comparison of results

In the square-root model (17), the cost of a VWAP execution is given exactly by the square-root formula:

$$\sigma \sqrt{\frac{Q}{V}} = 0.02 \times \sqrt{\frac{540}{6000}} = 0.02 \times 0.3 = 60 \text{ bp}$$

Table 1: Cost comparison

Strategy	Cost	Saving
VWAP	60.0 bp	
Bucket-like	49.6 bp	17%
Quasi-optimal	40.8 bp	32%

# Summary I

- The optimal trading strategy depends on the model.
  - For Almgren-Chriss style models, if the price of risk is zero, the minimal cost strategy is VWAP.
  - In Alfonsi-Schied style models with resiliency that depends only on the current spread, the minimal cost strategy is to trade a block at inception, a block at completion and at a constant rate in between.
  - More generally, if market impact is transient, the optimal strategy involves bursts of trading; VWAP is never optimal.
- In most conventional models, the optimal liquidation strategy is independent of the stock price.
  - However, for each such model, it is straightforward to specify a similar model in which the optimal strategy does depend on the stock price.
  - With reasonable parameters and timescales, the optimal strategy is close to the static one.

## Summary II

- In some models, price manipulation is possible and there is no optimal strategy.
- It turns out that we also need to exclude *transaction-triggered price manipulation*.
  - We presented example of models for which price manipulation is possible.
  - In the case of linear transient impact, we provided conditions under which transaction-triggered price manipulation is precluded.
- Empirically, the simple square-root model of market impact turns out to be a remarkably accurate description for reasonably sized meta-orders.
  - Assuming square-root dynamics consistent with this model, we showed that large cost savings are possible by optimizing the scheduling strategy.