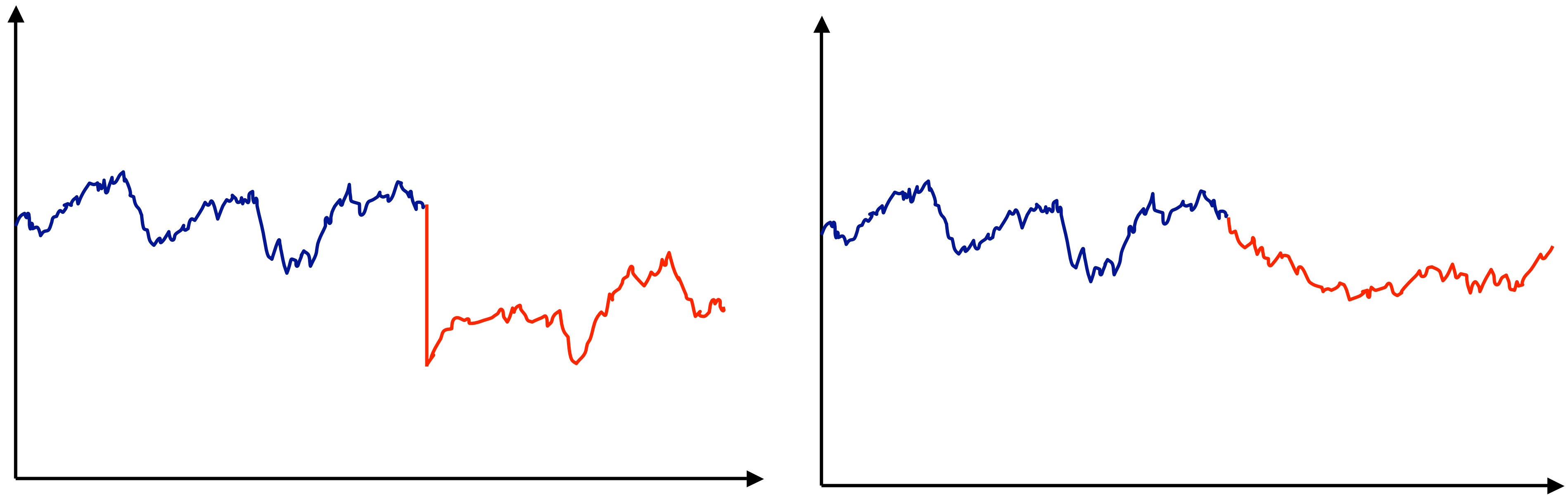
Market impact models and optimal trade execution

Alexander Schied
Mannheim University

9th Winter school on Mathematical Finance
Lunteren
January 18–20, 2010

Market impact: adverse feedback effect on the quoted price of a stock caused by one's own trading



Basic observation: liquidity costs of a large trade can be reduced signficantly by splitting the trade into a sequence of smaller trades, which are then spread out over a certain time interval.

Questions:

- Why is it better to spread out orders?
- What is an appropriate model for market impact?
- When is a model 'viable'? Can there be undesirable properties?
- What are the optimal trade execution strategies?
- Are strategies and models robust w.r.t. model parameters?

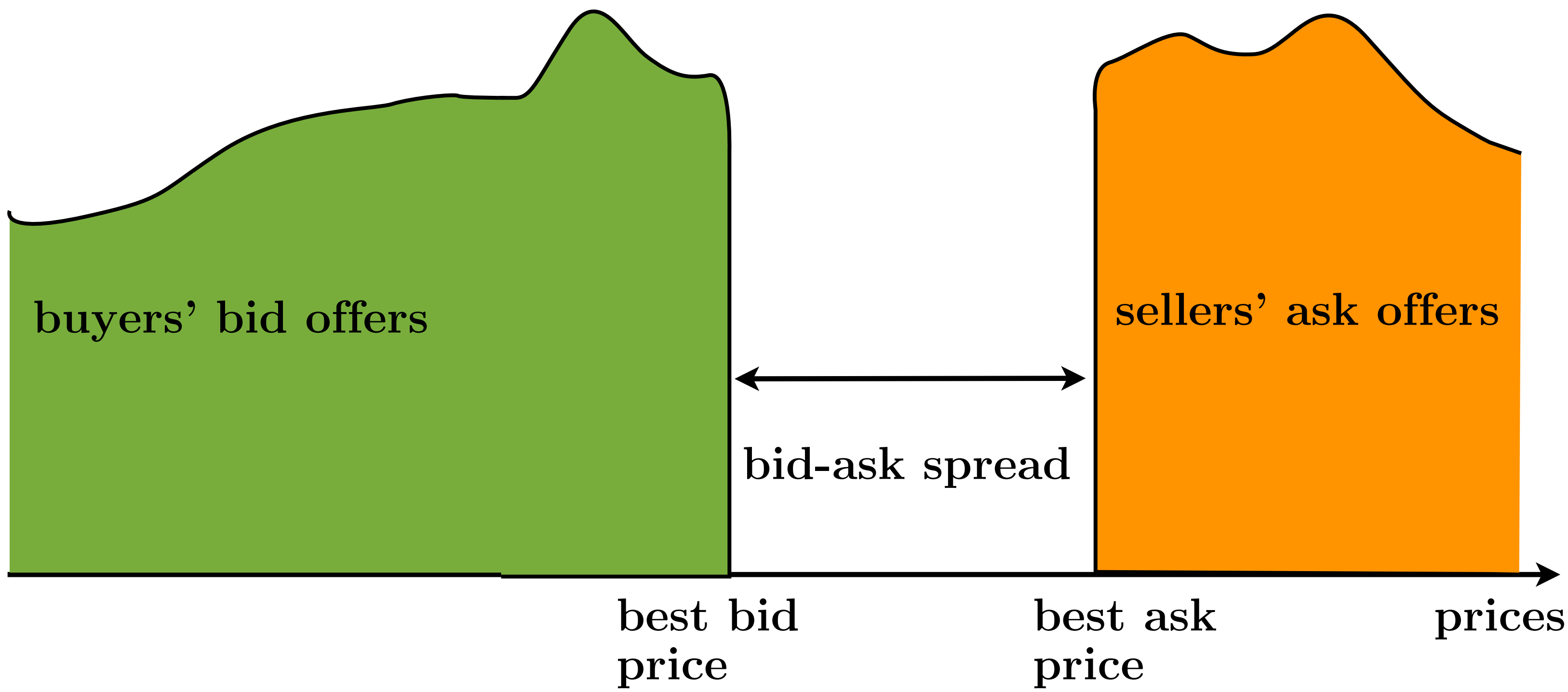
Interesting because:

- **Liquidity/market impact risk in its purest form**
  - development of realistic market impact models
  - checking viability of these models
  - building block for more complex problems

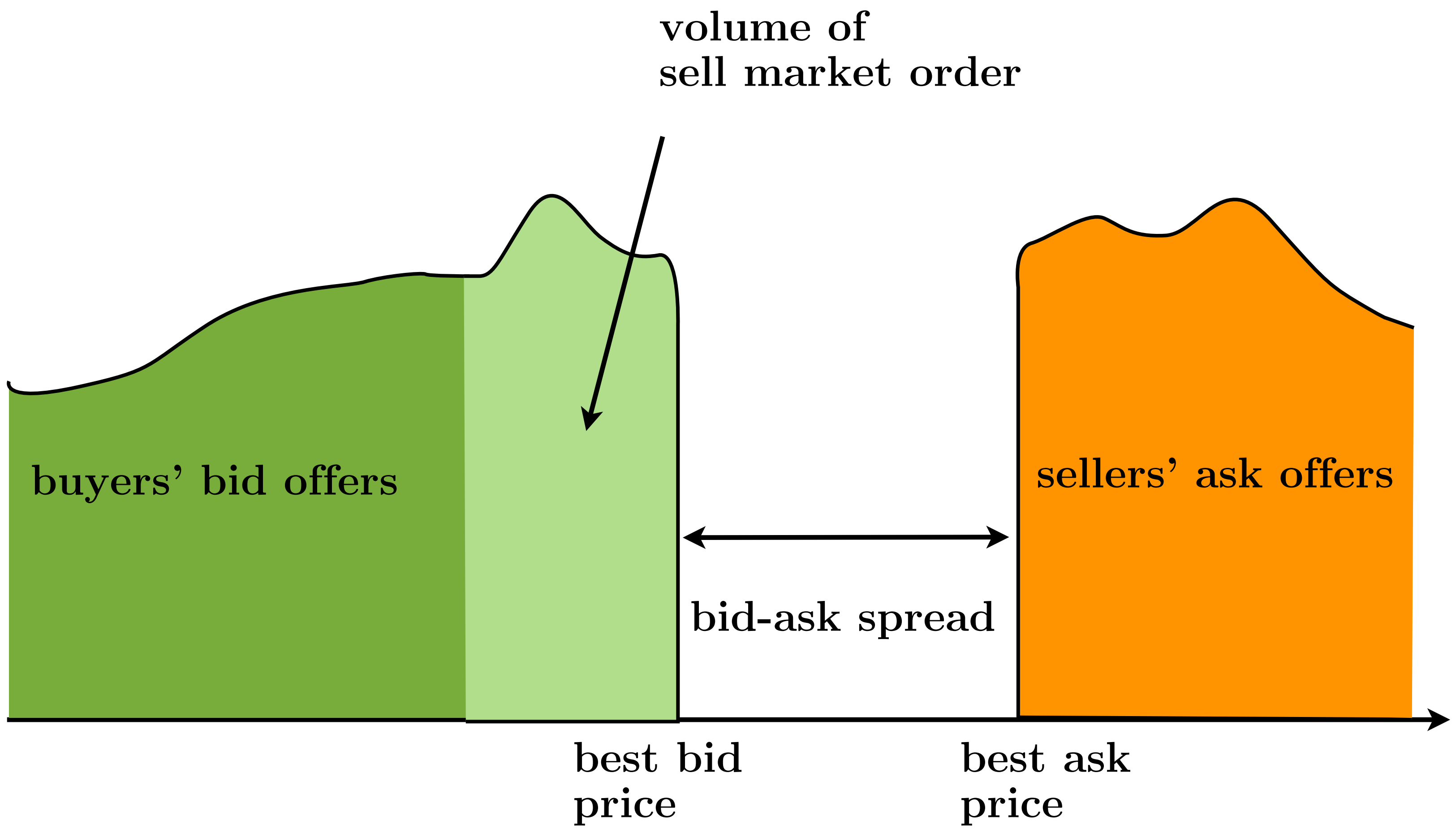
- **Relevant in applications**
  - real-world tests of new models

- **Interesting mathematics**

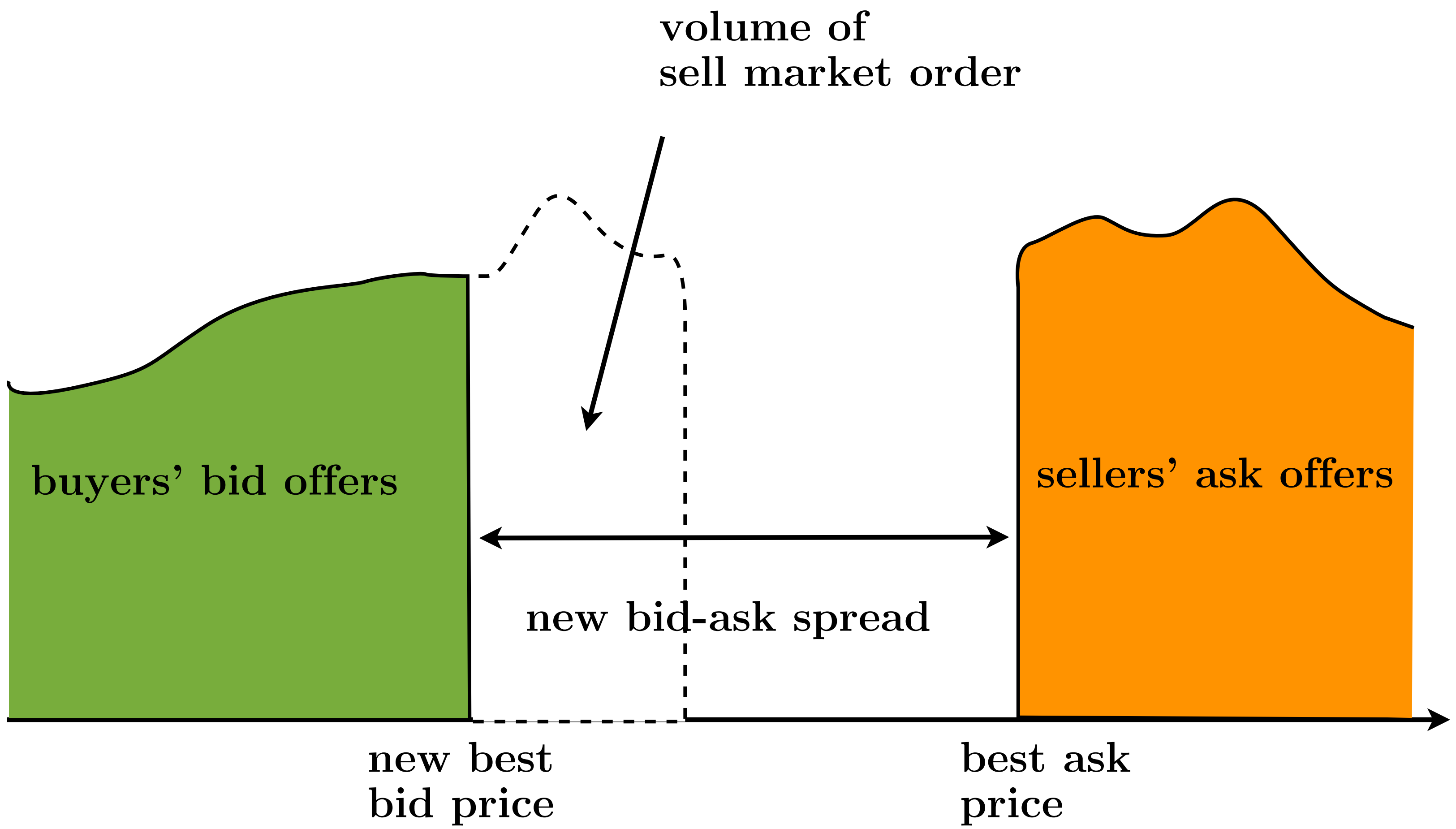
Limit order book before market order



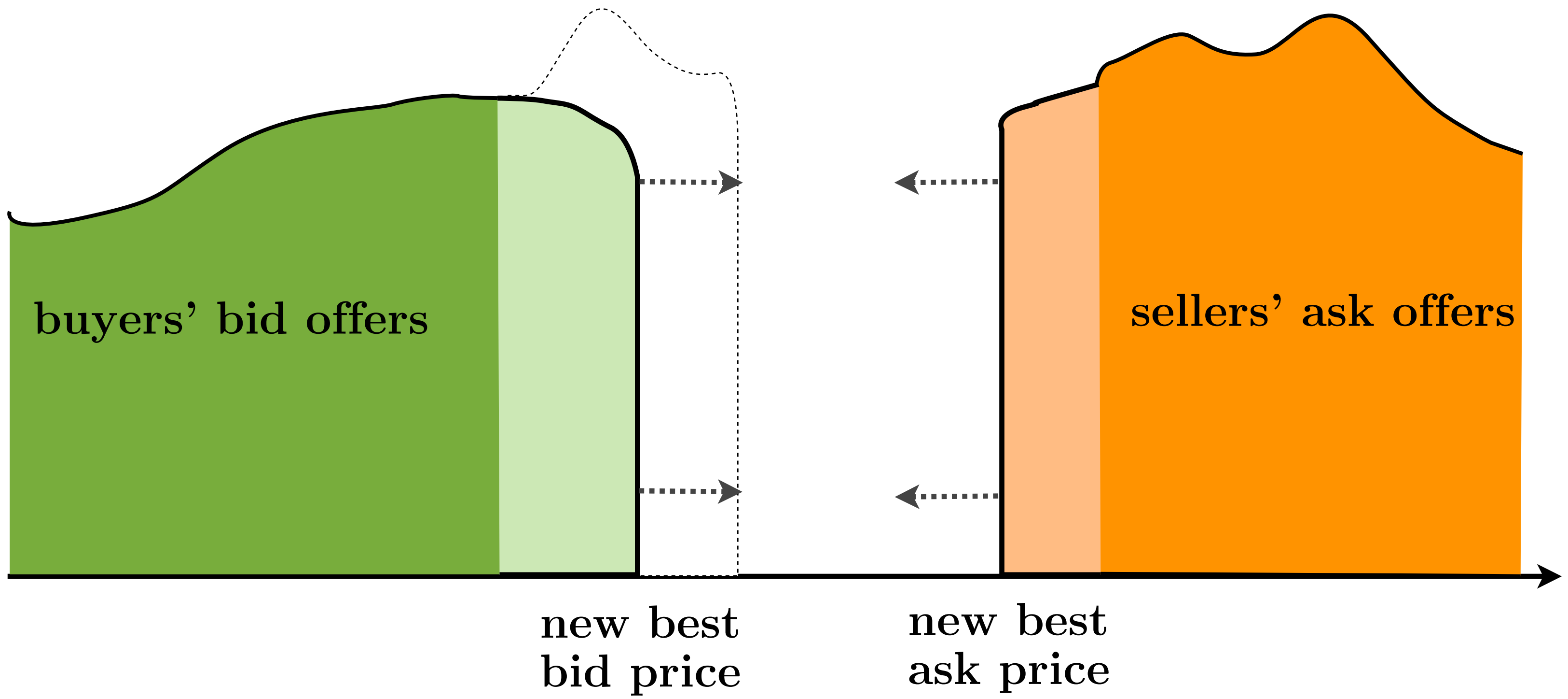
Limit order book before market order



Limit order book after market order



Resilience of the limit order book after market order



Overview:

I. Models based on order book dynamics

II. The qualitative effects of risk aversion

III. Multi-agent equilibrium

Overview:

I. Models based on order book dynamics

Microscopic: Emphasis on single trades

II. The qualitative effects of risk aversion

Mesoscopic: Emphasis on trajectory of trades

III. Multi-agent equilibrium

Macroscopic: Emphasis on interaction with competitors

Overview:

I. Models based on order book dynamics

Classical maths

II. The qualitative effects of risk aversion

Calculus of variations, stochastic control, and PDEs

III. Multi-agent equilibrium

Computer-aided proofs based on explicit computations

I. Order book models

1. Linear impact, general resilience

2. Nonlinear impact, exponential resilience

3. Gatheral's model

References

A. Alfonsi, A. Fruth, and A.S.: *Optimal execution strategies in limit order books with general shape functions.* To appear in Quantitative Finance

A. Alfonsi, A. Fruth, and A.S.: *Constrained portfolio liquidation in a limit order book model.* Banach Center Publ. 83, 9-25 (2008).

A. Alfonsi and A.S.: *Optimal execution and absence of price manipulations in limit order book models.* Preprint 2009.

A. Alfonsi, A.S., and A. Slynko: *Order book resilience, price manipulation, and the positive portfolio problem.* Preprint 2009.

Bertsimas, D., Lo, A. *Optimal control of execution costs.* Journal of Financial Markets, 1, 1-50 (1998).

Bouchaud, J. P., Gefen, Y., Potters, M., Wyart, M. *Fluctuations and response in financial markets: the subtle nature of 'random' price changes.* Quantitative Finance 4, 176-190 (2004).

J. Gatheral: *No-Dynamic-Arbitrage and Market Impact.* To appear in Quantitative Finance.

J. Gatheral, A.S., and Slynko, A. *Transient linear price impact and Fredholm integral equations.* Preprint (2010).

Huberman, G., Stanzl, W. *Price manipulation and quasi-arbitrage. Econometrica* **72**, 4, 1247-1275 (2004).

A. Obizhaeva and J. Wang: *Optimal Trading Strategy and Supply/Demand Dynamics.* Preprint (2005)

Potters, M., Bouchaud, J.-P. *More statistical properties of order books and price impact.* Physica A 324, No. 1-2, 133-140 (2003).

Weber, P., Rosenow, B. *Order book approach to price impact.* Quantitative Finance 5, no. 4, 357-364 (2005).

I. Order book models

1. Linear impact, general resilience

Unaffected price process: martingale S^0

Admissible trategy: predictable process $X = (X_t)$ that describes the number of shares held by the trader

- $t \to X_t$ is rightcontinuous with finite total variation
- the signed measure dX_t has compact support
- w.l.o.g. $X_t = 0$ for large enough t.

For instance, when $X_t = x$ for $t \leq t_0$ and $X_t = 0$ for $t > t_0$, then X describes a single trade of $|x|$ shares, executed at time t_0, which is a sell trade for $x > 0$ and a buy trade for $x < 0$.

Note: These strategies are of bounded variation.
So there will be no liquidation costs in models such as the Bank-Baum model, the Cetin-Jarrow-Protter model etc.

Impacted price process:

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

where

$$G : (0,\infty) \to [0,\infty)$$

is the decay kernel. It describes the resilience of price impact between trades; see Bouchaud et al. (2004), Obizhaeva and Wang (2005), Alfonsi et al. (2008, 2007), Gatheral (2008).

We first assume

(1) $\qquad G$ is bounded and $G(0) := \lim_{t \downarrow 0} G(t)$ exists.

Costs of a strategy X:

When X is continuous at t, then the infinitesimal order dX_t is executed at price S_t, so $S_t\,dX_t$ is the cost increment.
Thus, the total costs of a continuous strategy are

$$\int S_t\,dX_t = \int S_t^0\,dX_t + \int\int_{\{s<t\}} G(t-s)\,dX_s\,dX_t.$$

When X has a jump ΔX_t, then the price is moved from S_t to

$$S_{t+} = S_t + \Delta X_t G(0)$$

This linear price impact corresponds to a constant supply curve for which $G(0)^{-1}\,dy$ buy or sell orders are available at each price y. The trade ΔX_t is thus carried out at the following cost,

$$\int_{S_t}^{S_{t+}} yG(0)^{-1}\,dy = \frac{1}{2G(0)}\big(S_{t+}^2 - S_t^2\big) = \frac{G(0)}{2}(\Delta X_t)^2 + \Delta X_t S_t.$$

Hence, the total costs of an arbitrary admissible strategy X are given by

$$\begin{aligned}
&\int S_t\,dX_t + \frac{G(0)}{2}\sum(\Delta X_t)^2\\
&\quad= \int S^0_t\,dX_t + \int\int_{\{s<t\}} G(t-s)\,dX_s\,dX_t + \frac{G(0)}{2}\sum(\Delta X_t)^2\\
&\quad= \int S^0_t\,dX_t + \frac{1}{2}\int\int G(|t-s|)\,dX_s\,dX_t.
\end{aligned}$$

It therefore follows from the martingale property of S^0 that the **expected costs** of an admissible strategy are

$$\mathbb{E}\Big[\int S^0_t\,dX_t\Big] + \frac{1}{2}\mathbb{E}[\,\mathcal{C}(X)\,],$$

where

$$\mathcal{C}(X) := \int\int G(|t-s|)\,dX_s\,dX_t.$$

Next if, e.g., S^0 is continuous and T is such that $X_T = 0$, then

$$\int S_t^0 \, dX_t = X_0 S_0^0 - X_T S_T^0 - \int_0^T X_{t-} \, dS_t^0 .$$

Hence,

$$\mathbb{E}\Big[\int S_t^0 \, dX_t \Big] = X_0 S_0^0 ,$$

and the expected costs are

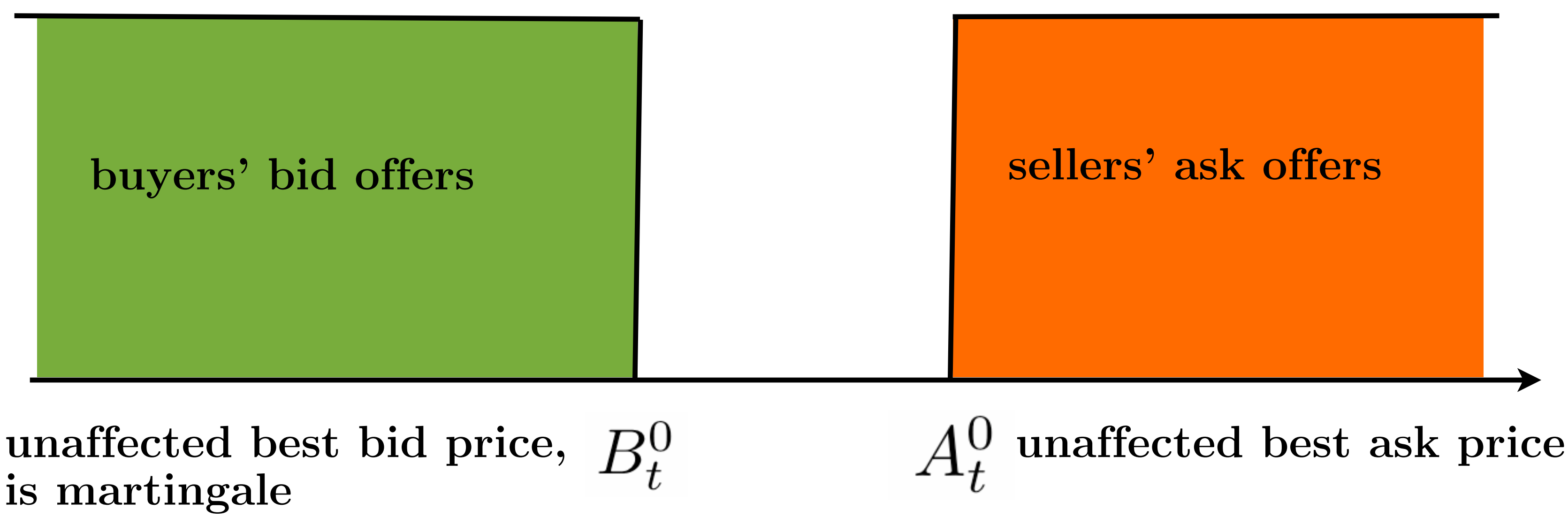
$$X_0 S_0^0 + \frac{1}{2} \mathbb{E}[\, \mathcal{C}(X)\,].$$

Remark: Instead of this simple market impact model, one can consider more complicated models for (block-shaped) electronic limit order books. In these models one can then show that

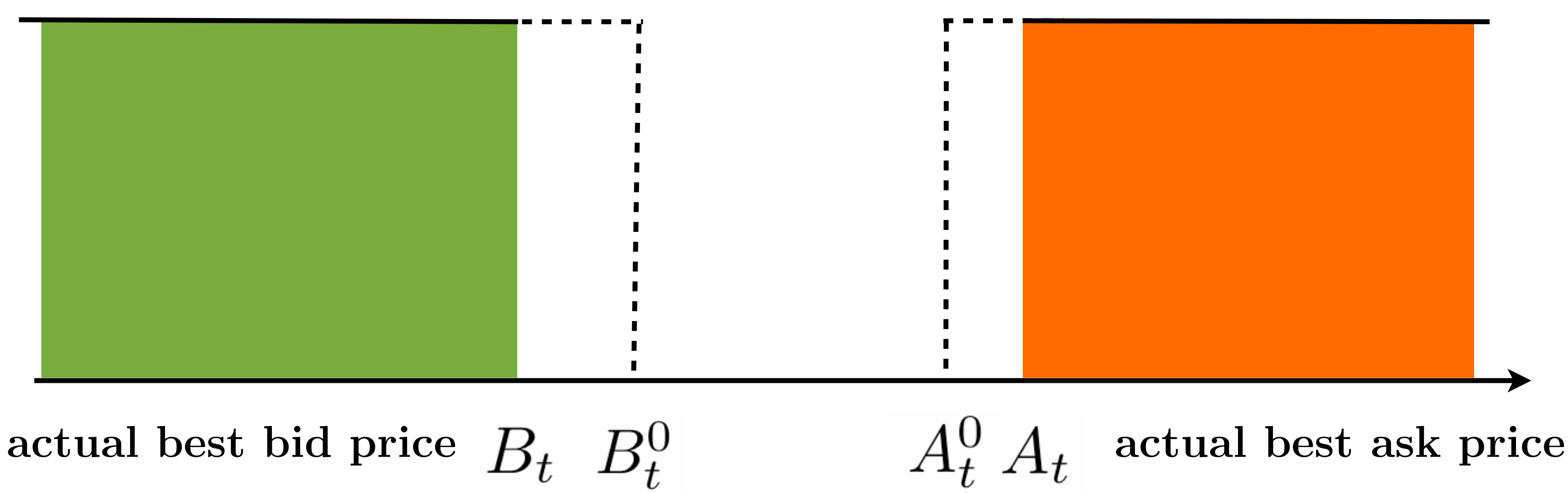
$$\text{Expected costs } \geq S_0^0 X_0 + \frac{1}{2}\mathbb{E}[\,\mathcal{C}(X)\,]$$

with equality for monotone strategies X.

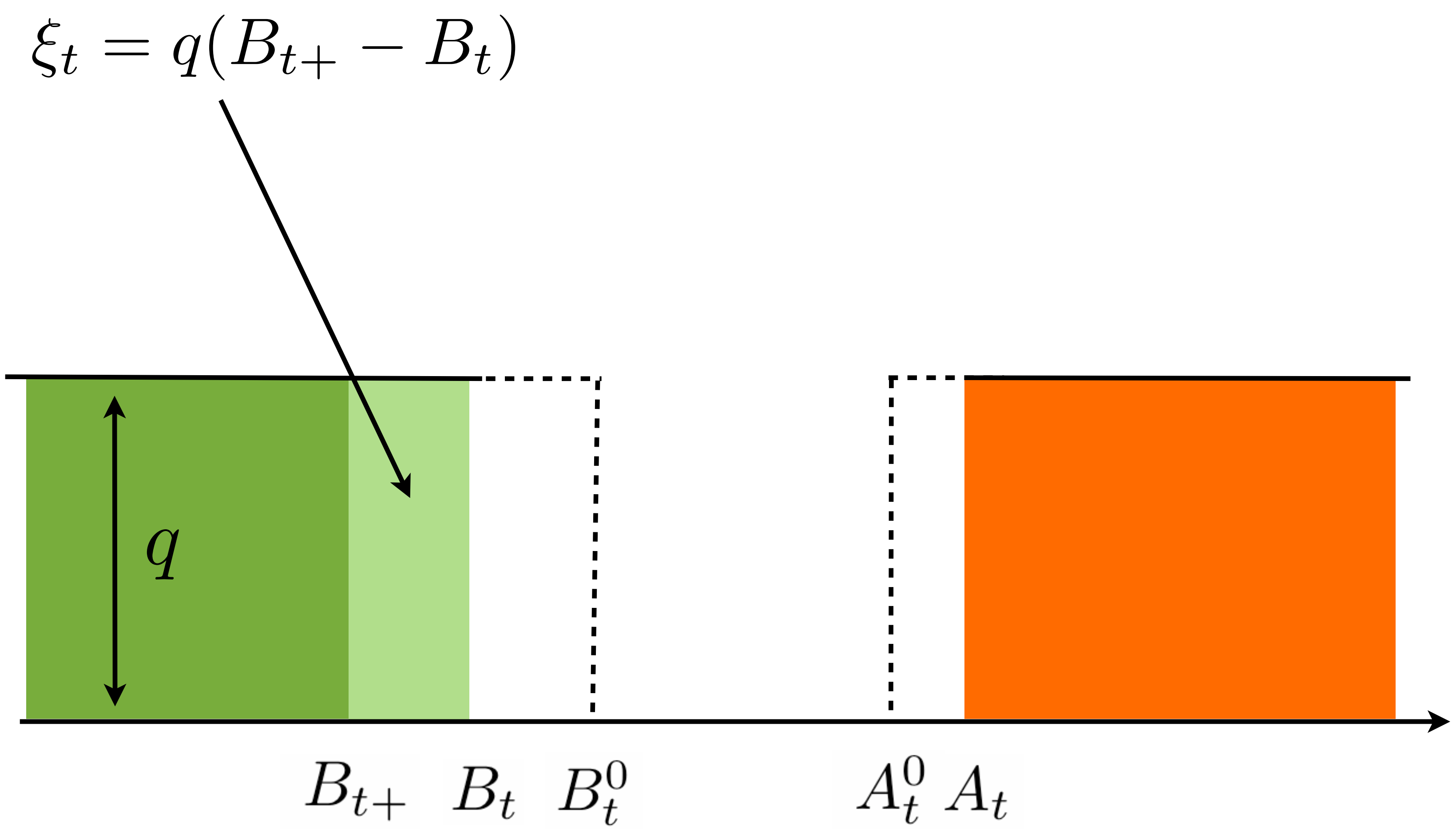
Limit order book model without large trader



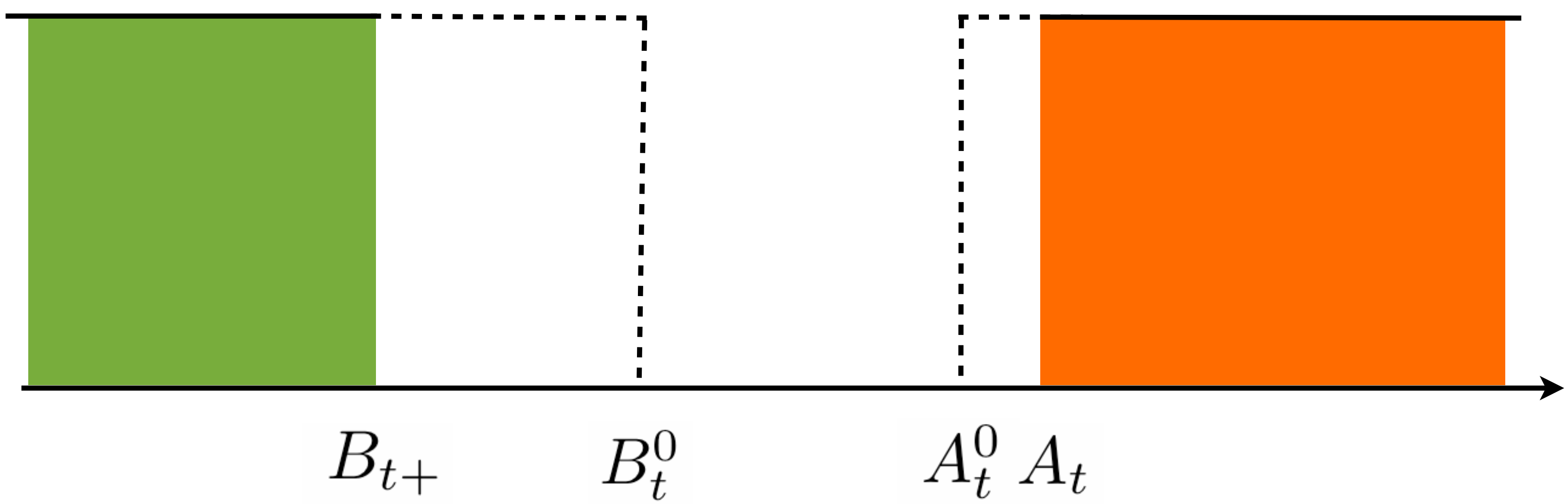
Limit order book model after large trades



Limit order book model at large trade



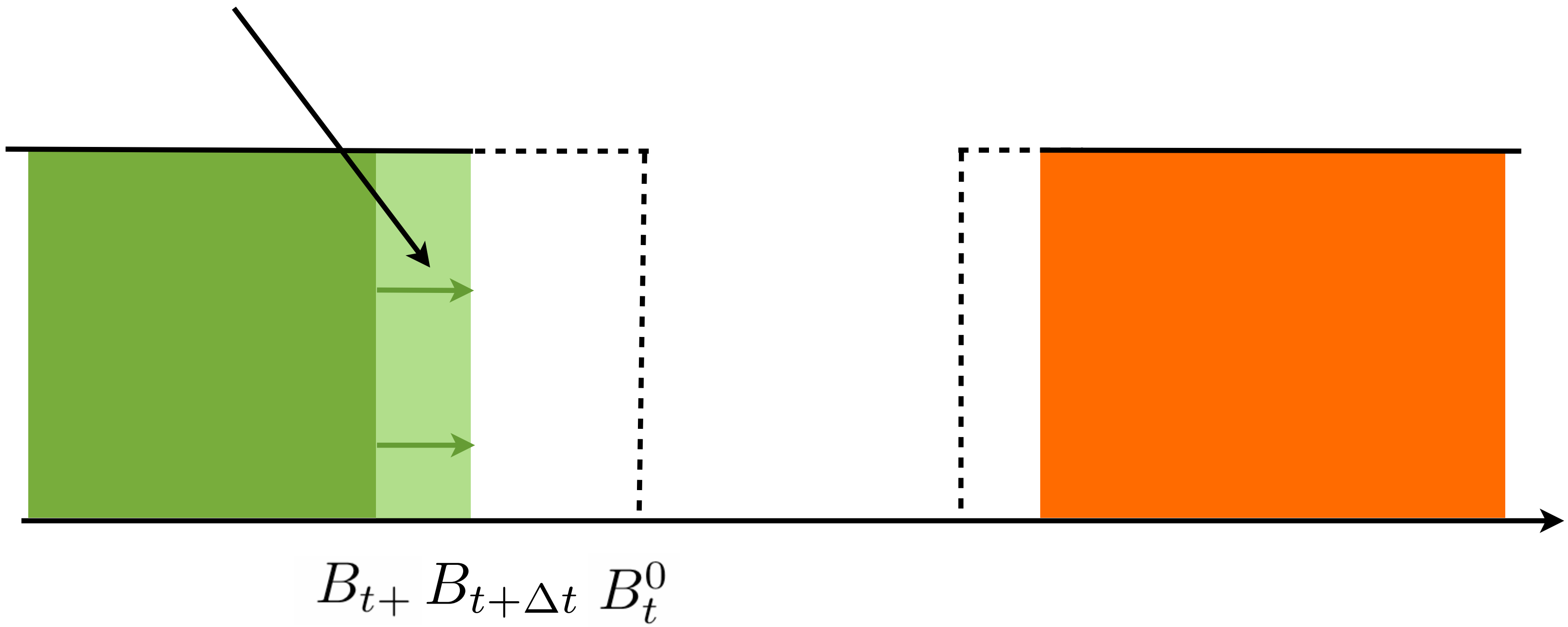
Limit order book model immediately after large trade



Resilience of the limit order book

$\psi : [0, \infty[\to [0, 1]$, $\psi(0) = 1$, decreasing

$\frac{\xi_t}{q} \cdot \psi(\Delta t) +$ decay of previous trades



Remark: Instead of this simple market impact model, one can consider more complicated models for (block-shaped) electronic limit order books. In these models one can then show that

$$\text{Expected costs } \geq S_0^0 X_0 + \frac{1}{2}\mathbb{E}[\,\mathcal{C}(X)\,]$$

with equality for monotone strategies X.

Two questions:

- Can there be model irregularities?
- Existence, uniqueness, and structure of strategies minimizing the expected costs?

Definition 1 (Huberman and Stanzl (2004)). A round trip is an admissible strategy with $X_0 = 0$. A price manipulation strategy is a round trip with strictly negative expected costs.

Clearly, there is no price manipulation when

$$\mathcal{C}(X) \geq 0 \qquad \text{for all strategies } X.$$

Proposition 1 (Straightforward extension of Bochner's thm). *$\mathcal{C}(X) \geq 0$ for all strategies $X \iff G(|\cdot|)$ can be represented as the Fourier transform of a positive finite Borel measure μ on $\mathbb{R}$, i.e.,*

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

(G is **positive definite***). If, in addition, the support of μ is not discrete, then $\mathcal{C}(X) > 0$ for every nonzero admissible strategy X (G is* **strictly positive definite***).*

Remark 1. Suppose that X is a step function with jumps at times $t_0, \ldots, t_N$, i.e.,

$$X_t = X_0 - \sum_{t_i < t} \xi_i.$$

Then

$$\mathcal{C}(X) = \sum \xi_i \xi_j G(|t_i - t_j|)$$

Proof of Proposition 1: Suppose first that $\mathcal{C}(X) \geq 0$ for all strategies X. When considering strategies with discrete support we are in the context of Bochner's theorem, and so $G(|\cdot|)$ must be the Fourier transform of a positive finite Borel measure μ on $\mathbb{R}$.

Conversely, suppose that $G(|x|) = \int_{\mathbb{R}} e^{ixz}\,\mu(dz)$. When X is an admissible strategy, then

$$\begin{aligned}\mathcal{C}(X) &= \int\int\int e^{iz(t-s)}\,\mu(dz)\,dX_s\,dX_t \\ &= \int\int e^{izt}\,dX_t\overline{\int e^{izs}\,dX_s}\,\mu(dz) = \int |\widehat{X}(z)|^2\,\mu(dz) \geq 0,\end{aligned}$$

where $\widehat{X}(z) = \int e^{itz}\,dX_t$ is the Fourier transform of X. It is well-defined due to our assumption that X has compact support.

Let us finally show that $\mathcal{C}$ is even positive definite when the support of μ is not discrete. Since X has compact support, the function $\widehat{X}(z)$ has a continuation to an entire analytic function on the complex plane. Indeed, one easily uses Lebesgue's theorem to see that

$$\widehat{X}(z) = \int e^{itz}\, dX_t$$

is finite and differentiable as a function of $z \in \mathbb{C}$.

Hence, for $X \neq 0$, the zero set of $\widehat{X}$ must be a discrete set. Thus, for the integral

$$\mathcal{C}(X) = \int |\widehat{X}(z)|^2\, \mu(dz)$$

to vanish, the measure μ needs to have discrete support. □

Optimal trade execution problem: Minimizing expected costs,

$$S_0^0 y + \frac{1}{2}\mathbb{E}[\,\mathcal{C}(X)\,]$$

for strategies that liquidate a given long or short position of y shares within a given time frame.

Time constraint: compact set $\mathbb{T} \subset [0, \infty)$.

Boils down to minimizing $\mathcal{C}(\cdot)$ over

$$\mathcal{X}(y, \mathbb{T}) := \Big\{ X \,\Big|\, \text{deterministic strategy with } X_0 = y \text{ and support in } \mathbb{T} \Big\}.$$

Suppose first that $\mathbb{T}$ is discrete, i.e., $\mathbb{T} = \{t_0, \dots, t_N\}$. Then the problem is equivalent to

$$\text{minimize} \ \sum_{i,j=0}^{N} x_i x_j G(|t_i - t_j|) \qquad \text{over } \boldsymbol{x} \in \mathbb{R} \text{ with } \boldsymbol{x}^\top \mathbf{1} = y$$

where

$$\mathbf{1} = (1, \dots, 1)^\top$$

Minimizers always exist when G is positive definite. When G is strictly positive definite, the optimal $\boldsymbol{x}^*$ is proportional to the solution of

$$M\boldsymbol{x} = \mathbf{1}, \text{ i.e., to } M^{-1}\mathbf{1}$$

where

$$M_{ij} = G(|t_i - t_j|)$$

Existence of minimizers not clear when $\mathbb{T}$ is not discrete.

Proposition 2. *When G is strictly positive definite there exists at most one optimal strategy for given y and $\mathbb{T}$.*

Proof: Let

$$\mathcal{C}(X,Y)=\frac{1}{2}\Big(\mathcal{C}(X+Y)-\mathcal{C}(X)-\mathcal{C}(Y)\Big)=\int\int G(|t-s|)\,dX_s\,dY_t$$

First, $X\neq Y$ implies that

$$0<\mathcal{C}(X-Y)=\mathcal{C}(X)+\mathcal{C}(Y)-2\mathcal{C}(X,Y).$$

Therefore,

$$\mathcal{C}\Big(\frac{1}{2}X+\frac{1}{2}Y\Big)=\frac{1}{4}\mathcal{C}(X)+\frac{1}{4}\mathcal{C}(Y)+\frac{1}{2}\mathcal{C}(X,Y)<\frac{1}{2}\mathcal{C}(X)+\frac{1}{2}\mathcal{C}(Y),$$

which implies the uniqueness of optimal execution strategies when they exist. □

Proposition 3. *Suppose that G is positive definite. Then $X^* \in \mathcal{X}(y,\mathbb{T})$ is optimal if and only if there is a constant λ such that X^* solves the generalized Fredholm integral equation*

$$\int G(|t-s|)\,dX^*_s = \lambda \qquad \text{for all } t \in \mathbb{T}. \tag{2}$$

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

Proof: To prove that (2) is necessary for optimality, fix $t_0, t \in \mathbb{T}$, and let Y be the round trip defined by $dY_u = \delta_{t_0}(ds) - \delta_t(ds)$. Then, for all $\alpha \in \mathbb{R}$,

$$\mathcal{C}(X^* + \alpha Y) = \mathcal{C}(X^*) + \alpha^2 \mathcal{C}(Y) + 2\alpha\mathcal{C}(X^*, Y).$$

By optimality, the righthand side must be $\geq \mathcal{C}(X^*)$ for all $\alpha \in \mathbb{R}$.

Taking the derivative with respect to α at $\alpha = 0$ it follows that

$$0 = \mathcal{C}(X^*, Y) = \int G(|t_0 - s|)\, dX^*_s - \int G(|t - s|)\, dX^*_s.$$

By varying t we see that (2) is necessary for optimality.

Conversely, suppose that $X^* \in \mathcal{X}(y, \mathbb{T})$ is a strategy satisfying (2).
Let $\widetilde{X}$ be any other strategy in $\mathcal{X}(y, \mathbb{T})$ and define $Z := \widetilde{X} - X^*$.
Then, for $T := \max \mathbb{T}$,

$$\mathcal{C}(X^*, Z) = \int\int G(|t - s|)\, dX^*_s\, dZ_t = \frac{\lambda}{2}(Z_T - Z_0) = 0$$

and hence

$$\mathcal{C}(\widetilde{X}) = \mathcal{C}(X^*+Z) = \mathcal{C}(X^*)+\mathcal{C}(Z)+2\mathcal{C}(X^*, Z) = \mathcal{C}(X^*)+\mathcal{C}(Z) \geq \mathcal{C}(X^*).$$

Hence, X^* is optimal. □

Examples

Example 1 (Exponential decay). For the exponential decay kernel

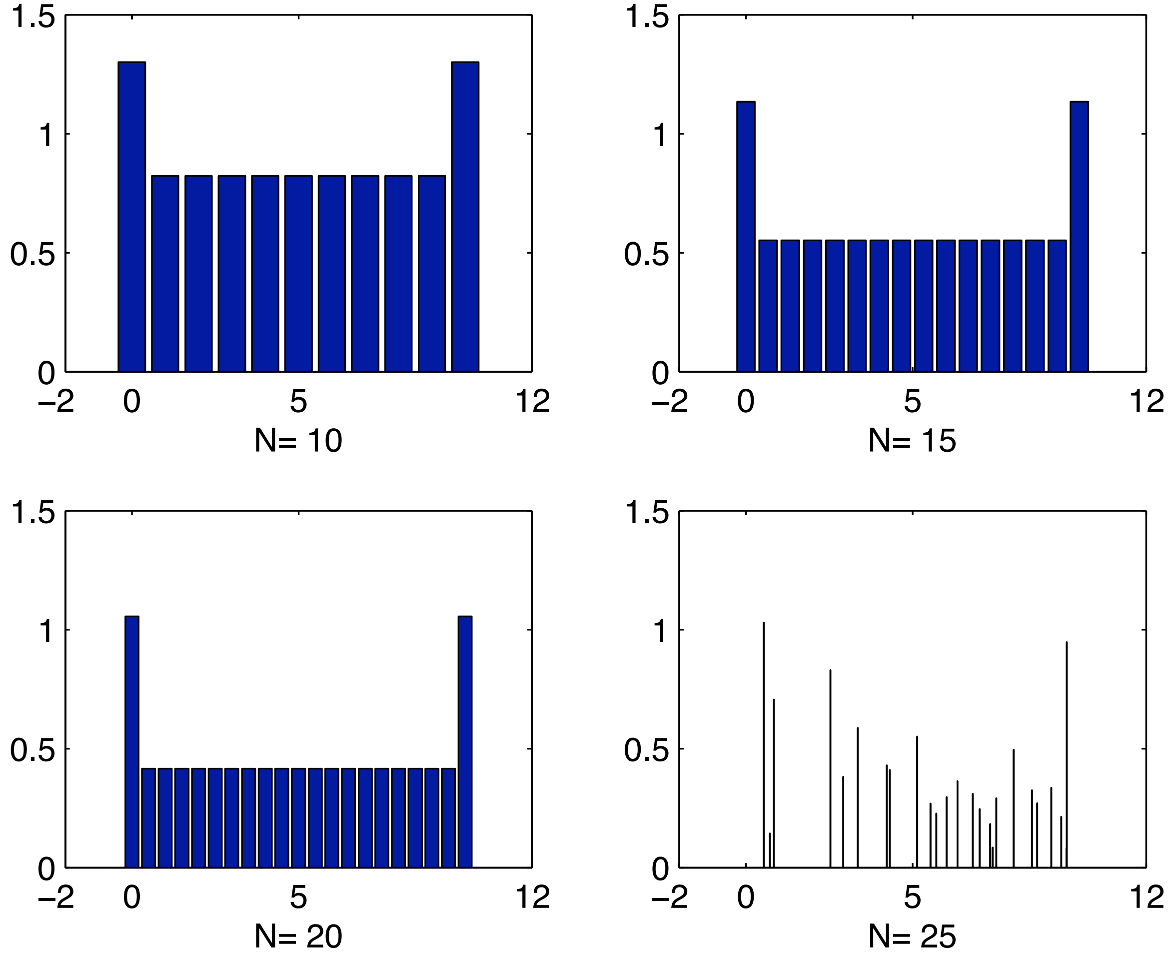
$$G(t) = e^{-\rho t},$$

$G(|\cdot|)$ is the Fourier transform of the positive measure

$$\mu(dt) = \frac{1}{\pi}\frac{\rho}{\rho^2 + t^2}\,dt$$

Hence, G is strictly positive definite.

Optimal strategies for $G(t) = e^{-\rho t}$ and discrete $\mathbb{T}$:



The optimal strategy can in fact be computed explicitly for any discrete time grid $\mathbb{T} = \{t_0, t_1, \dots, t_N\}$

Let $a_n := e^{-\rho(t_n - t_n - 1)}$ for $n = 1, \dots, N$. Then we can write

$$M = \begin{bmatrix} 1 & a_1 & a_1 a_2 & \cdots & \cdots & a_1 a_2 \cdots a_N \\ a_1 & 1 & a_2 & a_2 a_3 & \cdots & a_2 a_3 \cdots a_N \\ a_1 a_2 & a_2 & 1 & a_3 & \cdots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ a_2 \cdots a_N & & & a_{N-1} & 1 & a_N \\ a_1 a_2 \cdots a_N & \cdots & \cdots & a_{N-1} a_N & a_N & 1 \end{bmatrix}.$$

The inverse of M can be computed as the tridiagonal matrix

$$M^{-1} = \begin{bmatrix} \frac{1}{1-a_1^2} & \frac{-a_1}{1-a_1^2} & 0 & \cdots & 0 \\ \frac{-a_1}{1-a_1^2} & \left(\frac{1}{1-a_1^2} + \frac{a_2^2}{1-a_2^2}\right) & \frac{-a_2}{1-a_2^2} & 0 \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \frac{-a_{N-1}}{1-a_{N-1}^2} & \left(\frac{1}{1-a_{N-1}^2} + \frac{a_N^2}{1-a_N^2}\right) & \frac{-a_N}{1-a_N^2} \\ 0 & \cdots & 0 & \frac{-a_N}{1-a_N^2} & \frac{1}{1-a_N^2} \end{bmatrix}$$

From this formula, we get

$$M^{-1}\mathbf{1} = \begin{bmatrix} \frac{1}{1+a_1} \\ \frac{1}{1+a_1} - \frac{a_2}{1+a_2} \\ \vdots \\ \frac{1}{1+a_{N-1}} - \frac{a_N}{1+a_N} \\ \frac{1}{1+a_N} \end{bmatrix}$$

And hence

$$\boldsymbol{x}^* = \lambda_0 M^{-1}\mathbf{1}$$

for

$$\lambda_0 = \frac{y}{\mathbf{1}^\top M^{-1}\mathbf{1}} = \frac{y}{\frac{2}{1+a_1} + \sum_{n=2}^N \frac{1-a_n}{1+a_n}}.$$

The initial market order of the optimal strategy is hence

$$x_0^* = \frac{\lambda_0}{1+a_1},$$

the intermediate market orders are given by

$$x_n^* = \lambda_0\Big(\frac{1}{1+a_n} - \frac{a_{n+1}}{1+a_{n+1}}\Big), \qquad n = 1, \ldots, N-1,$$

and the final market order is

$$x_N^* = \frac{\lambda_0}{1+a_N}.$$

The initial market order of the optimal strategy is hence

$$x_0^* = \frac{\lambda_0}{1 + a_1},$$

the intermediate market orders are given by

$$x_n^* = \lambda_0 \Big(\frac{1}{1 + a_n} - \frac{a_{n+1}}{1 + a_{n+1}} \Big), \qquad n = 1, \dots, N-1,$$

and the final market order is

$$x_N^* = \frac{\lambda_0}{1 + a_N}.$$

It is clear that x_0^* and x_N^* are strictly positive. For $i = 1, \dots, N-1$ we have

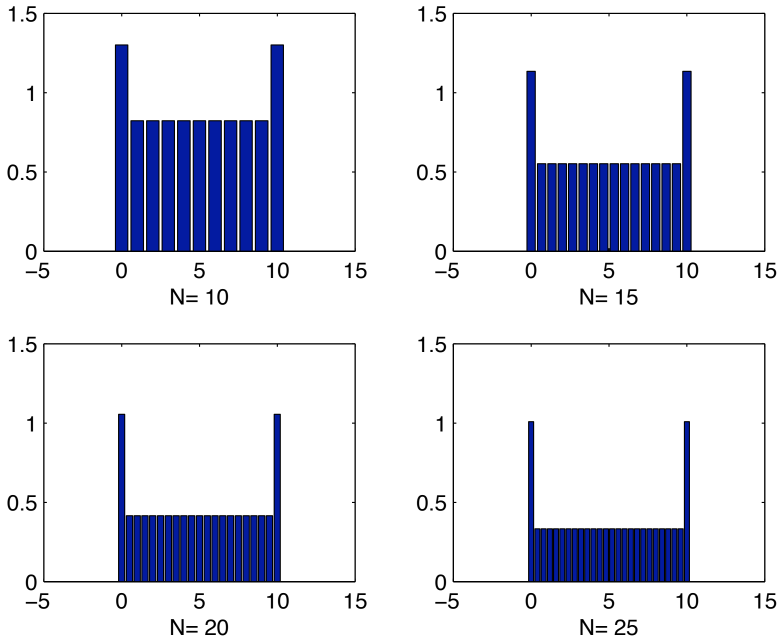
$$x_i^* = \lambda_0 \cdot \frac{(1 - a_i a_{i+1})}{(1 + a_i)(1 + a_{i+1})} > 0.$$

For the equidistant time grid $t_n = nT/N$ the solution simplifies:

$$x_0^* = x_N^* = \frac{y}{(N-1)(1-a)+2}$$

and

$$x_1^* = \cdots = x_{N-1}^* = \xi_0^*(1-a).$$

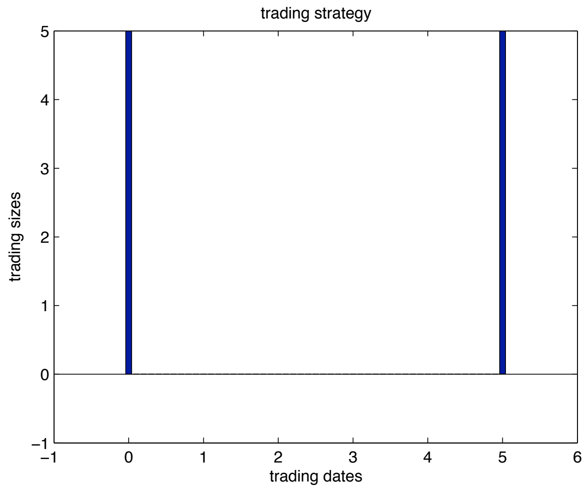


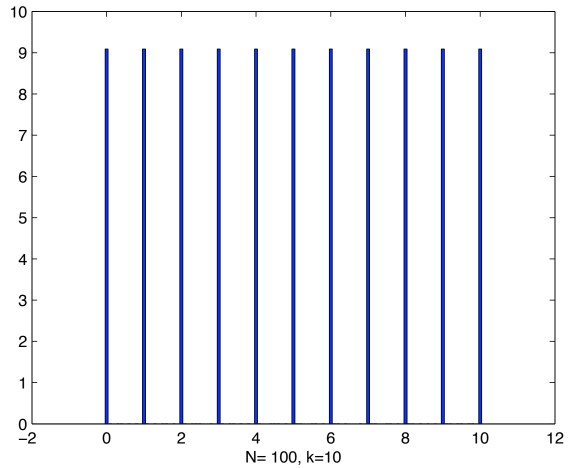
For $\mathbb{T} = [0, T]$:

$$dX_s^* = \frac{x}{\rho T + 2}\Big(\delta_0(ds) + \rho\, ds + \delta_T(ds)\Big).$$

Exercise: This strategy solves the generalized Fredholm integral equation.

Example 2 (Capped linear decay). $G(t) = (1 - \rho t)^+$



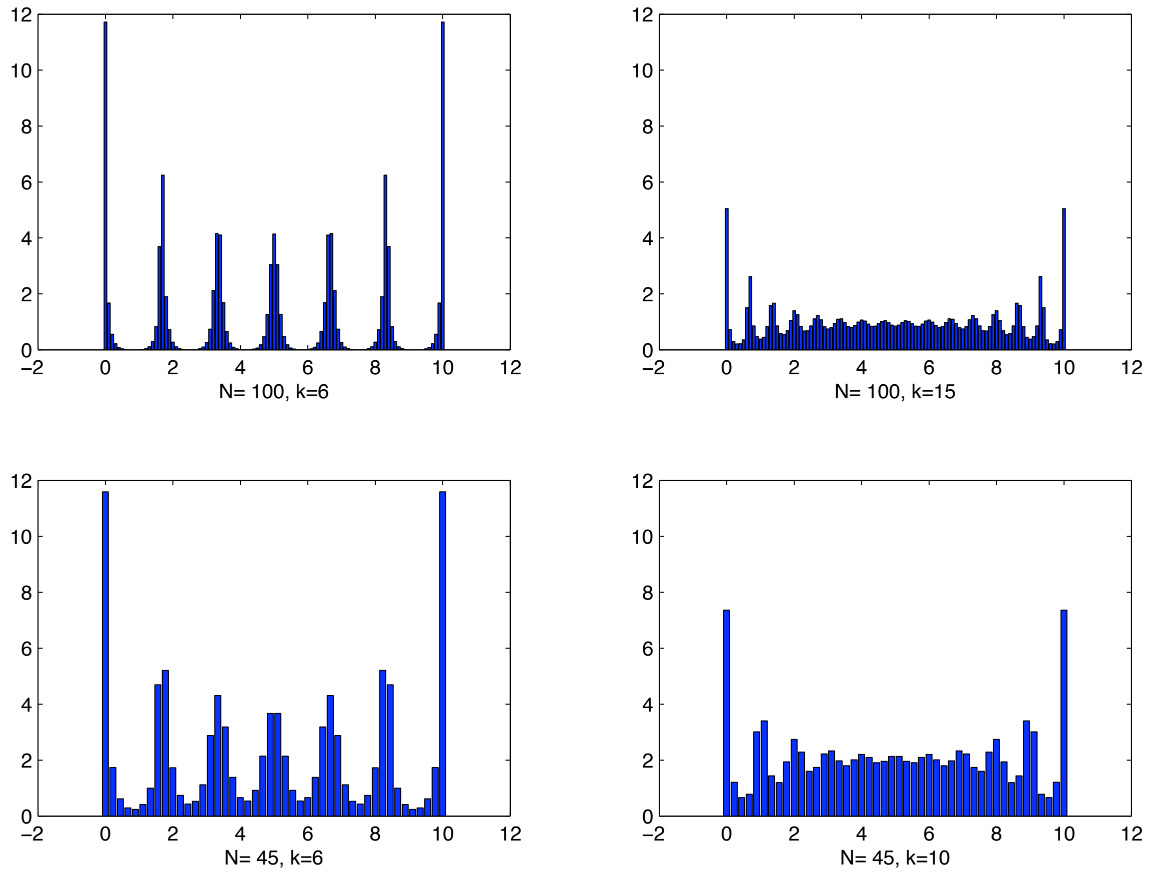


$\rho \leq 1/T$ and arbitrary $\mathbb{T}$

$\rho = N/T$, $\mathbb{T} = [0, T]$ or equisitant

Exercise: For $\mathbb{T} = [0, T]$, these strategies satisfy the corresponding Fredholm integral equations.

Otherwise, for equistant grid $\mathbb{T}$,



More generally: Convex decay

Theorem [Carathéodory (1907), Toeplitz (1911), Young (1912)]
G is convex, decreasing, nonnegative, and nonconstant $\Longrightarrow$
$G(|\cdot|)$ *is strictly positive definite.*

More generally: Convex decay

Theorem [Carathéodory (1907), Toeplitz (1911), Young (1912)]
G is convex, decreasing, nonnegative, and nonconstant $\Longrightarrow$ $G(|\cdot|)$ is strictly positive definite.

Proof: W.l.o.g.: G is continuous (exercise).
G' = right-hand derivative.
$G''(dx)$ = second derivative (= Borel measure on $[0, \infty]$).

For $\varepsilon > 0$ let $G_\varepsilon(x) := e^{-\varepsilon x} G(x)$ (is again convex and decreasing).

The inverse Fourier transform of $G_\varepsilon(|\cdot|)$ is proportional to

$$\begin{aligned}
\int_{-\infty}^{\infty} G_\varepsilon(|x|)e^{-ixz}\,dx &= 2\int_0^\infty G_\varepsilon(x)\cos xz\,dx \\
&= -2\int_0^\infty G'_\varepsilon(x)\int_0^x \cos zt\,dt\,dx \\
&= 2\int_0^\infty \int_0^x \int_0^t \cos sz\,ds\,dt\,G''_\varepsilon(dx) \\
&= 2\int_0^\infty \frac{1-\cos xz}{z^2}\,G''_\varepsilon(dx)
\end{aligned}$$

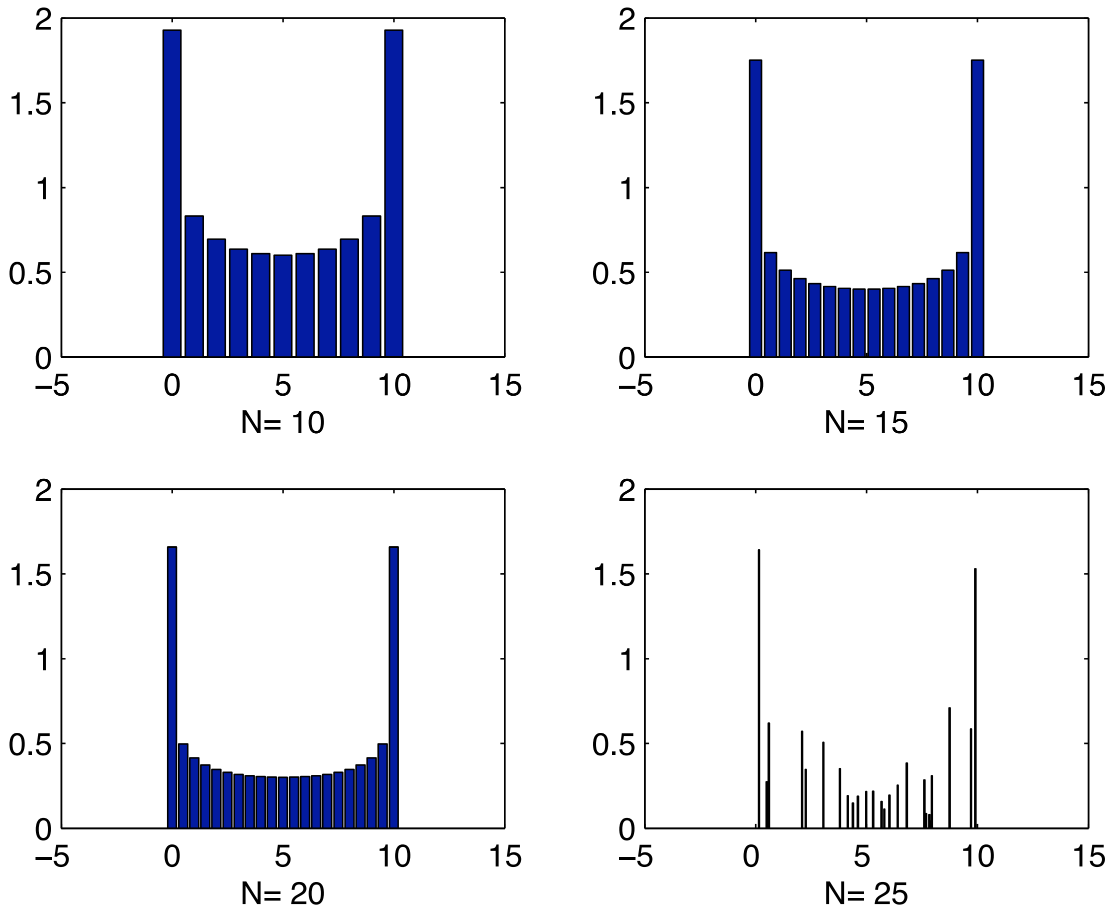
As a function of z, the right-hand side is the density of a positive finite Borel measure μ_ε. It follows that G_ε, and hence G, are positive definite functions.

Since $G_\varepsilon \to G$ pointwise, Lévy's theorem entails that μ_ε converges weakly to the measure μ, the inverse Fourier transform of G modulo a proportionality factor. By the portmanteau theorem:

$$\mu([a,b]) \ge \limsup_{\varepsilon\downarrow 0} \mu_\varepsilon([a,b]) \ge 2\int_0^\infty \int_a^b \frac{1-\cos xz}{z^2}\, dz\, G''(dx) > 0$$

for all $0 < a < b$. Hence, μ has full support, and so G is strictly positive definite. □

Example 3 (Power law decay). $G(t) = (1+t)^{-\alpha}$ and equidistant grid $\mathbb{T}$,



So everything looks nice for

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

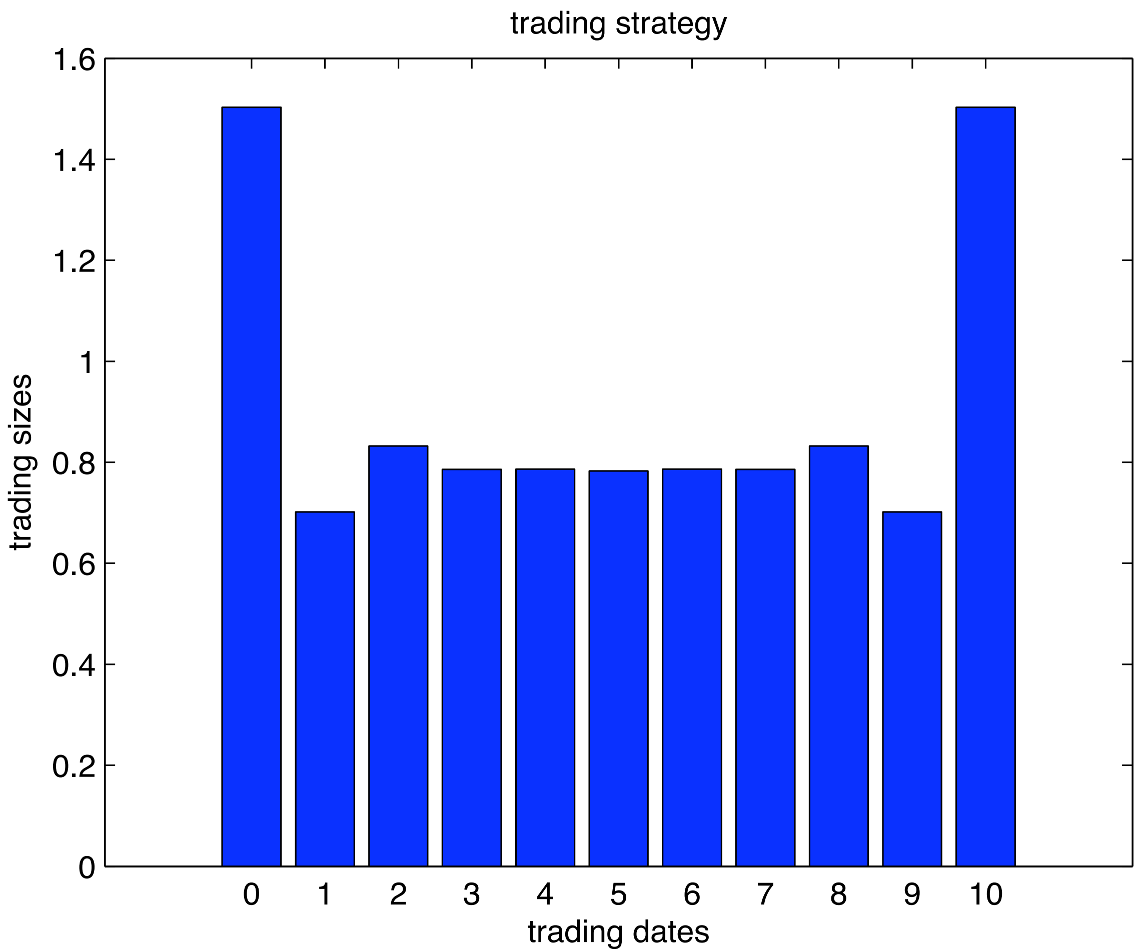
Let's look at:

Example 4 (Modified power-law decay). The decay kernel

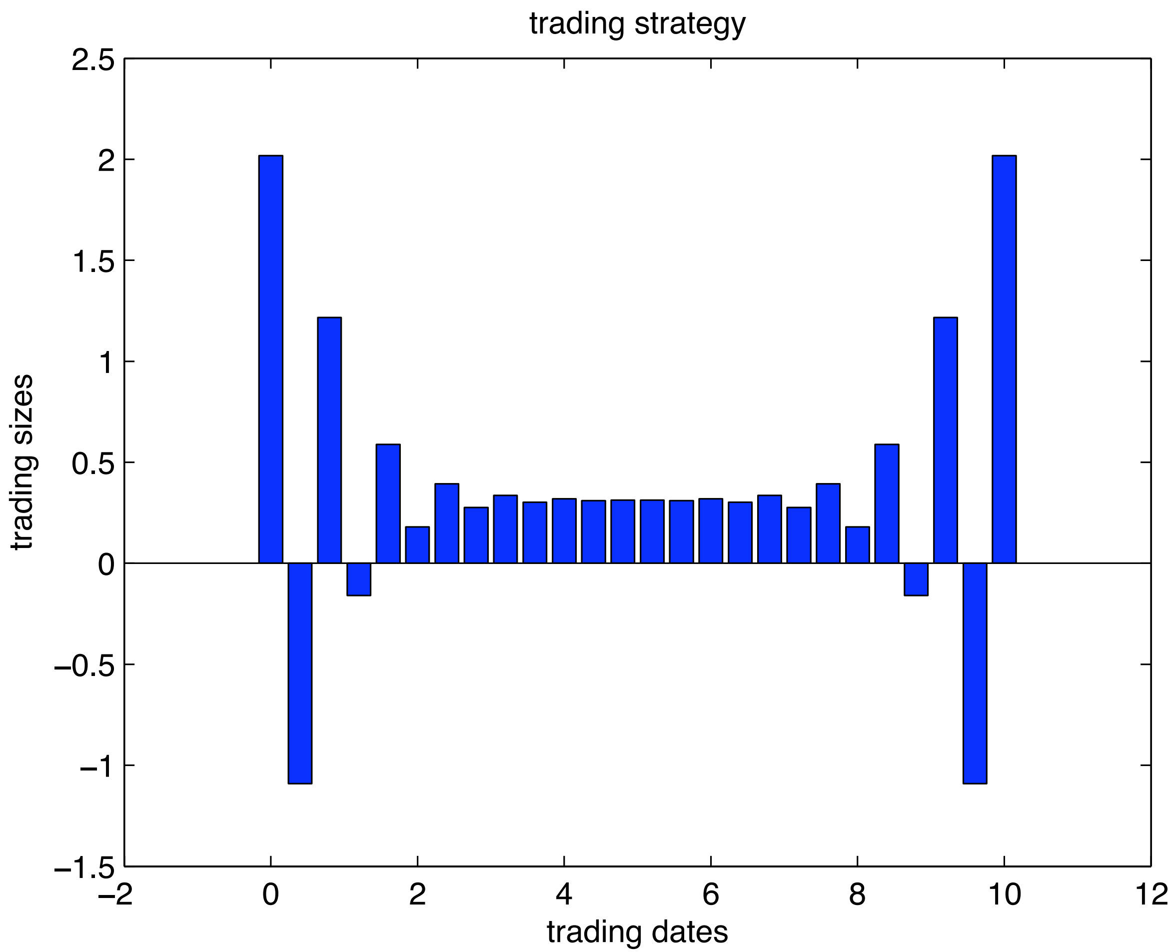
$$G(t) = \frac{1}{1+t^2}$$

is the Fourier transform of the function $\frac{1}{2}e^{-|x|}$. So it is strictly positive definite.

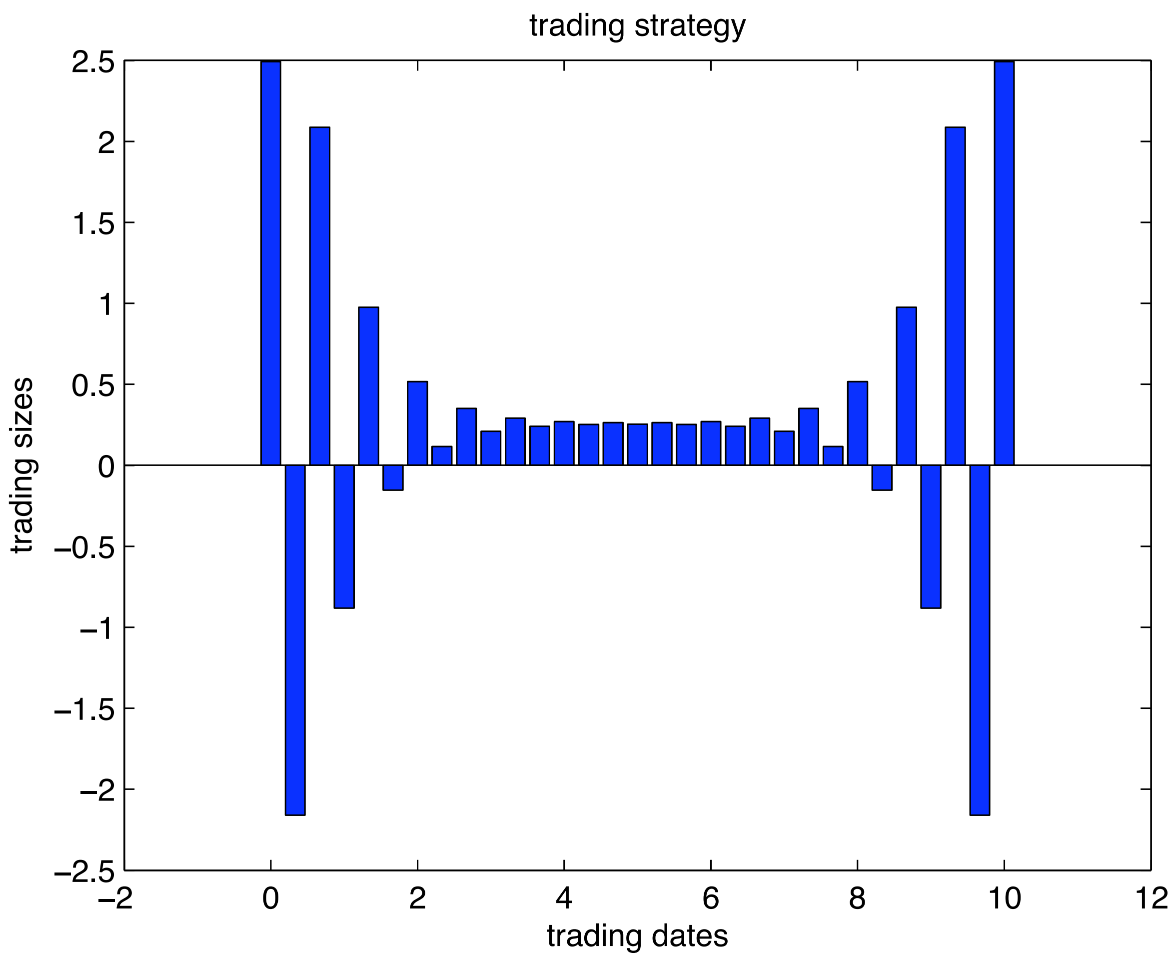
Modified power-law decay $G(t) = 1/(1+t^2)$, $N = 10$



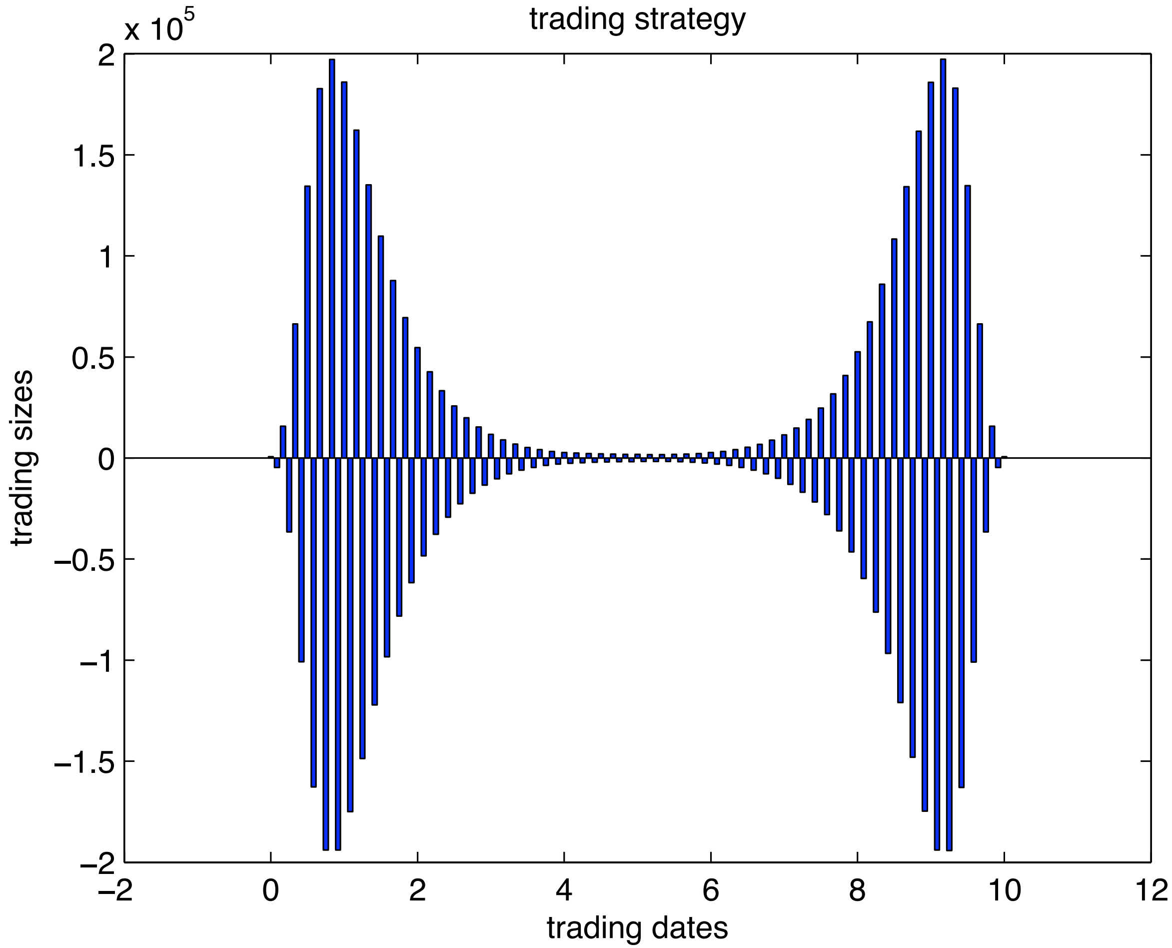
Modified power-law decay $G(t) = 1/(1+t^2)$, $N = 25$



Modified power-law decay $G(t) = 1/(1 + t^2)$, $N = 30$



Modified power-law decay $G(t) = 1/(1+t^2)$, $N = 120$



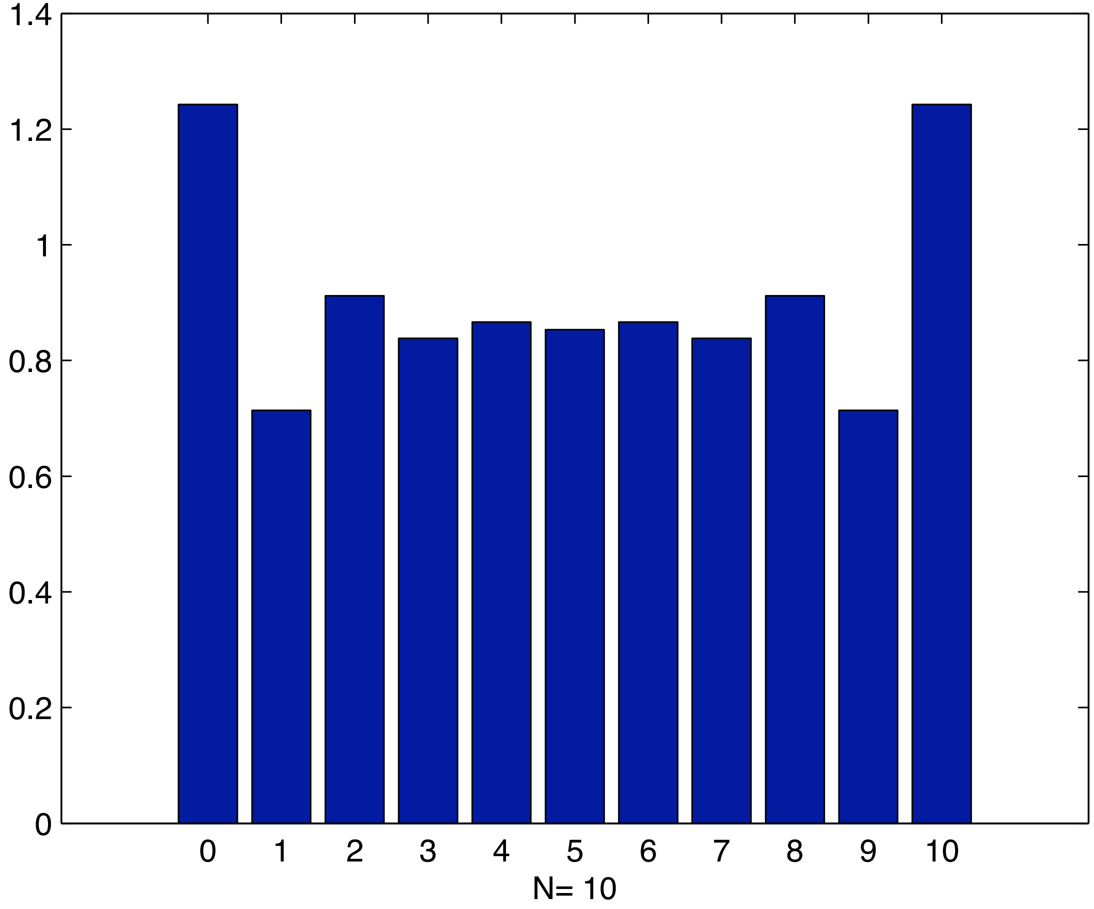
Example 4: Gaussian decay

The Gaussian decay function

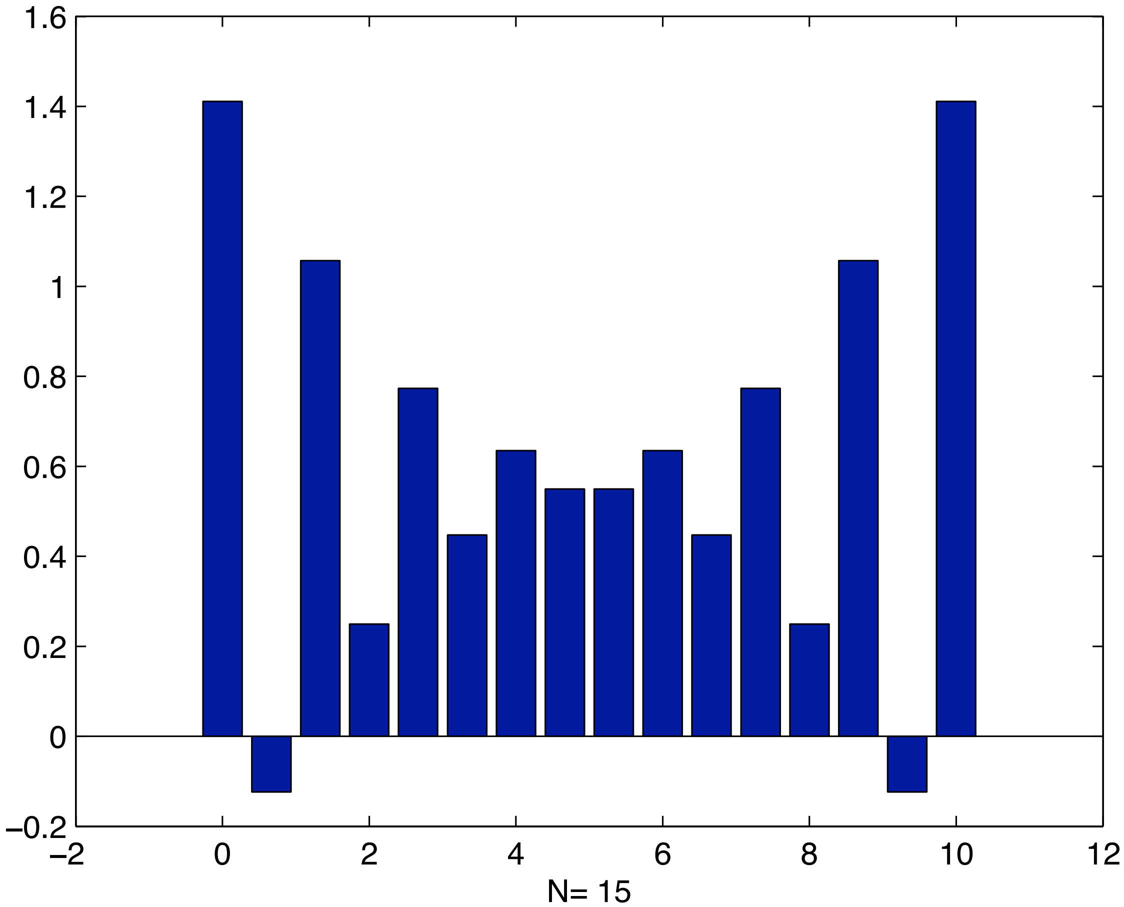
$$G(t) = e^{-t^2}$$

is its own Fourier transform (modulo constants) and hence strictly positive definite.

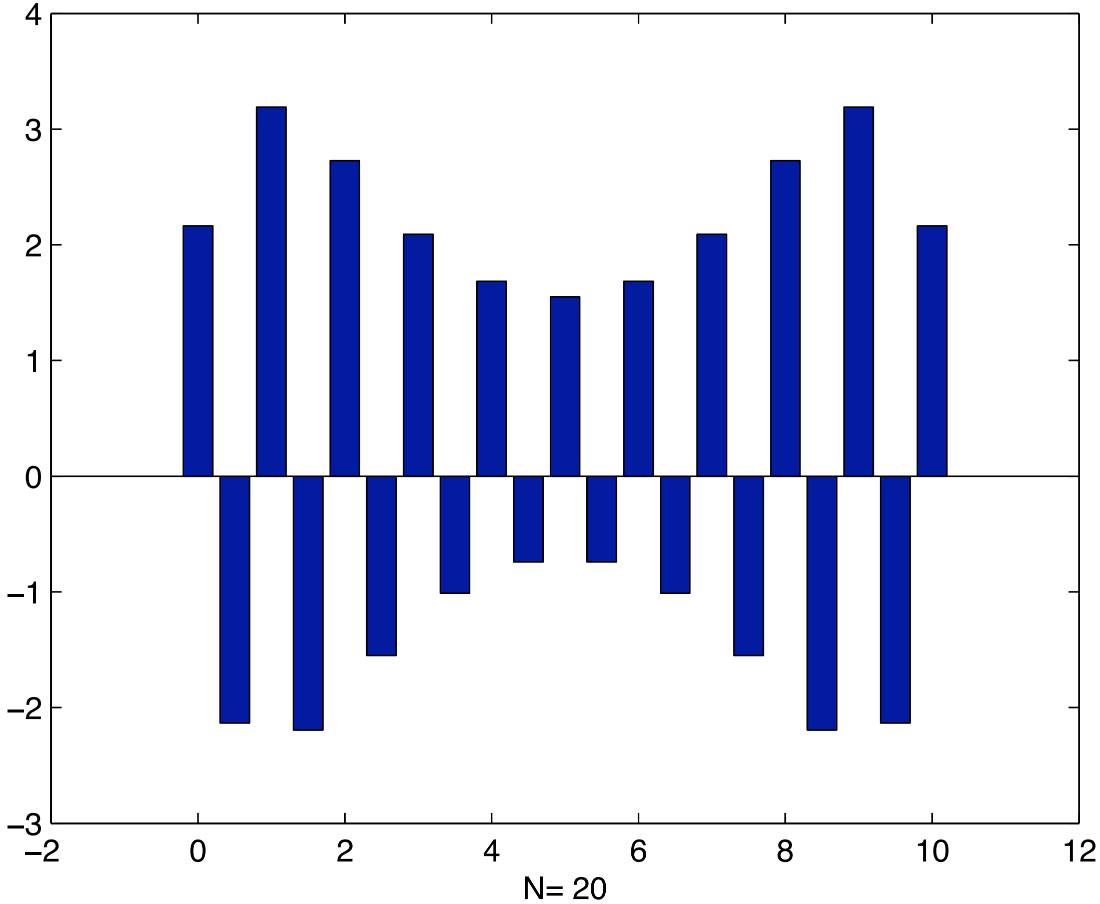
Gaussian decay $G(t) = e^{-t^2}$, $N = 10$



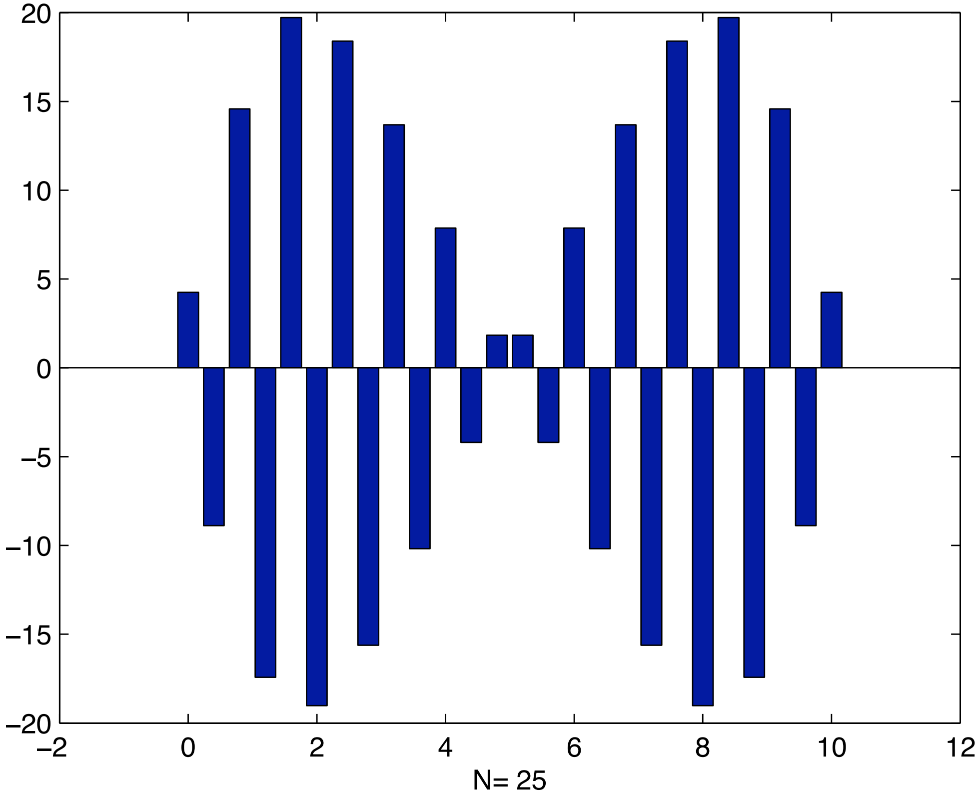
Gaussian decay $G(t) = e^{-t^2}$, $N = 15$



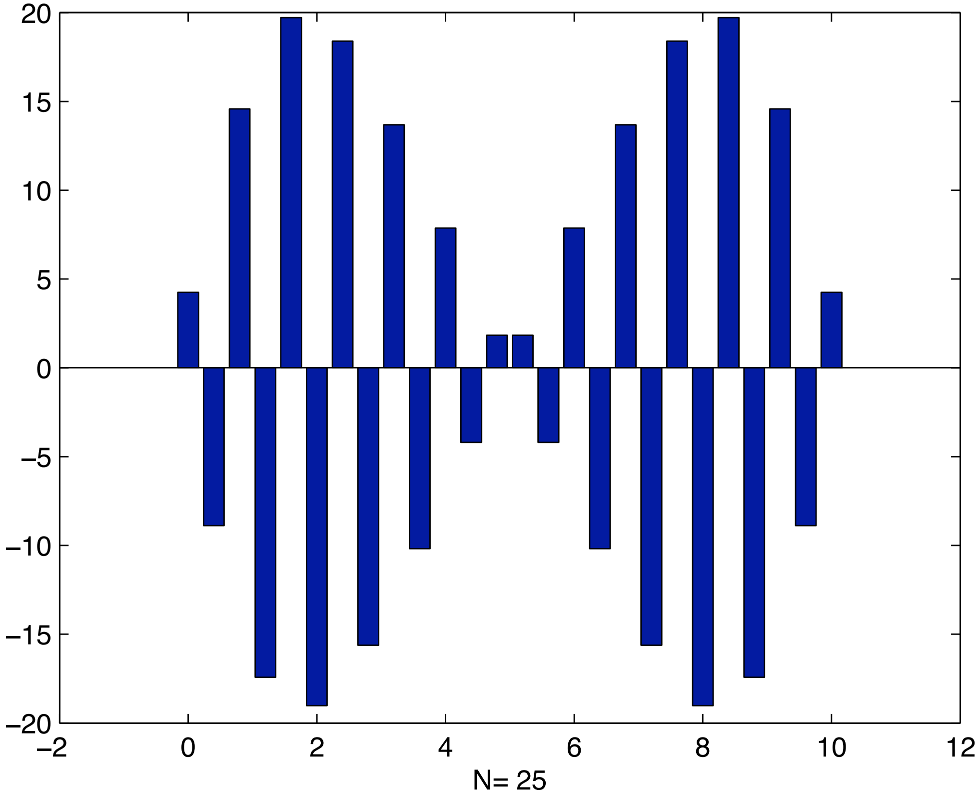
Gaussian decay $G(t) = e^{-t^2}$, $N = 20$



Gaussian decay $G(t) = e^{-t^2}$, $N = 25$



Gaussian decay $G(t) = e^{-t^2}$, $N = 25$



$\Rightarrow$ **absence of price manipulation strategies is not enough**

Definition [Hubermann & Stanzl (2004)]
A market impact model admits

price manipulation

if there is a round trip with negative expected liquidation costs.

Definition: [Alfonsi, A.S., & Slynko (2009)]
A market impact model admits

transaction-triggered price manipulation

if the expected liquidation costs of a sell (buy) program can be decreased by intermediate buy (sell) trades.

Situation for non-discrete $\mathbb{T}$:

Theorem 1. *Suppose that $G(|\cdot|)$ is the Fourier transform of a finite Borel measure μ for which*

$$\int e^{\varepsilon x}\,\mu(dx) < \infty \qquad \text{for some } \varepsilon > 0. \tag{3}$$

Suppose furthermore that the support of μ is not discrete. Then there are no optimal strategies in $\mathcal{X}(y,\mathbb{T})$ when $x \neq 0$ and $\mathbb{T}$ is not discrete.

Examples:

$$G(t) = e^{-t^2} \qquad \text{or} \qquad G(t) := \frac{1}{1+t^2}$$

$$\text{or} \qquad G(t) = 2\frac{1-\cos t}{t^2} \qquad \text{or} \qquad G(t) = 1 + \frac{\sin t}{t}$$

Sketch of proof: Suppose that X^* would be an optimal strategy. Due to the exponential moment condition,

$$h(t) := \int G(|t-s|)\, dX^*_s = \int\int e^{i(s-t)y}\,\mu(dy)\,dX^*_s = \int e^{-ity}\widehat{X}^*(y)\,\mu(dy)$$

admits an holomorphic continuation to the strip

$$S := \big\{ z \in \mathbb{C} \,\big|\, -\varepsilon < \Im(z) < \varepsilon \big\}$$

which is given by

$$h(z) = \int e^{-izy}\widehat{X}^*(y)\,\mu(dy), \qquad z \in S.$$

Next, $h(-t)$ is the Fourier transform of the complex-valued measure $\nu(dy) = \widehat{X}^*(y)\,\mu(dy)$, which is nontrivial. Hence, h is not constant, and so the zero set of $h(t) - \lambda$ must be discrete for any $\lambda \in \mathbb{R}$. $\square$

Theorem 2. *If G is nonconstant, nonincreasing, and convex, then there exists a unique optimal strategy X^* within each class $\mathcal{X}(y, \mathbb{T})$. Moreover, X^*_t is a monotone function of t.*

Theorem 2. *If G is nonconstant, nonincreasing, and convex, then there exists a unique optimal strategy X^* within each class $\mathcal{X}(y, \mathbb{T})$. Moreover, X_t^* is a monotone function of t.*

Proposition 4. *Suppose that there are s, $t > 0$, $s \neq t$, such that*

$$G(0) - G(s) < G(t) - G(t+s). \tag{4}$$

Then there is transaction-triggered price manipulation for the choice $\mathbb{T} := \{0, s, t+s\}$.

Condition (4) is satisfied, e.g., when $G(t)$ is strictly concave in a neighborhood of zero

and also implied by condition (3),

For discrete $\mathbb{T} = \{t_0, \ldots, t_N\}$:

Question: When does the minimizer $\boldsymbol{x}^*$ of

$$\sum_{i,j} x_i x_j G(|t_i - t_j|) \qquad \text{with} \quad \sum_i x_i = y$$

have only nonnegative components?

For discrete $\mathbb{T} = \{t_0, \dots, t_N\}$:

Question: When does the minimizer $\boldsymbol{x}^*$ of

$$\sum_{i,j} x_i x_j G(|t_i - t_j|) \qquad \text{with} \quad \sum_i x_i = y$$

have only nonnegative components?

Related to the **positive portfolio problem** in finance:

When are there no short sales in a Markowitz portfolio?

I.e. when is the solution of the following problem nonnegative

$$\boldsymbol{x}^\top M \boldsymbol{x} - \boldsymbol{m}^\top \boldsymbol{x} \to \min \qquad \text{for } \boldsymbol{x}^\top \mathbf{1} = y,$$

where M is a covariance matrix of assets and $\boldsymbol{m}$ is the returns vector?

Partial results, e.g., by Green (1986), Nielsen (1987)

Theorem 3. [Alfonsi, A.S., Slynko (2009)]

- *If G is convex then all components of $\boldsymbol{x}^*$ are nonnegative.*
- *If G is strictly convex, then all components are strictly positive.*

Theorem 3. [Alfonsi, A.S., Slynko (2009)]

- *If G is convex then all components of $\boldsymbol{x}^*$ are nonnegative.*
- *If G is strictly convex, then all components are strictly positive.*

Proof of first two assertions needs the following duality result:

Lemma 1. *Let M be an symmetric invertible matrix. Then*

$$M^{-1}\mathbf{1} \geq \mathbf{0} \qquad \textit{or} \qquad M^{-1}\mathbf{1} \leq \mathbf{0}$$

if and only if there is no vector $\boldsymbol{z}$ such that

$$\boldsymbol{z}^\top \mathbf{1} = 0 \qquad \textit{and} \qquad M\boldsymbol{z} > \mathbf{0}$$

Proof of Lemma 1. First suppose that $M^{-1}\mathbf{1} \geq 0$ or $M^{-1}\mathbf{1} \leq 0$. Assume by way of contradiction that there exists $\boldsymbol{z}$ with $\boldsymbol{z}^\top \mathbf{1} = 0$ and $M\boldsymbol{z} > \mathbf{0}$. Since $M^{-1}\mathbf{1} \neq 0$ we must have that $0 < (M^{-1}\mathbf{1})^\top M\boldsymbol{z}$ or $0 < (M^{-1}\mathbf{1})^\top M\boldsymbol{z}$. On the other hand

$$(M^{-1}\mathbf{1})^\top M\boldsymbol{z} = \mathbf{1}^\top M^{-1} M\boldsymbol{z} = \mathbf{1}^\top \boldsymbol{z} = 0,$$

which is a contradiction.

Conversely, suppose that neither $M^{-1}\mathbf{1} \geq 0$ nor $M^{-1}\mathbf{1} \leq 0$. Then the vector $\boldsymbol{x} := M^{-1}\mathbf{1}$ has two components $x_i < 0$ and $x_j > 0$. Hence there exists $\varepsilon > 0$ and a vector $\boldsymbol{y}$ with $y_i > 0$, $y_j > 0$, and $y_k = \varepsilon$ for all other components such that $\boldsymbol{y}^\top \boldsymbol{x} = 0$. It follows that $\boldsymbol{z} := M^{-1}\boldsymbol{y}$ satisfies $M\boldsymbol{z} = \boldsymbol{y} > 0$, $\boldsymbol{z} \neq 0$, and $\boldsymbol{z}^\top \mathbf{1} = \boldsymbol{y}^\top M^{-1}\mathbf{1} = \boldsymbol{y}^\top \boldsymbol{x} = 0$. $\square$

Proof of Theorem 3. Use induction on N to exclude the existence of $\boldsymbol{z} = (z_0, \dots, z_N)^\top$ such that $\boldsymbol{z}^\top \mathbf{1} = 0$ and $M\boldsymbol{z} > \mathbf{0}$ with $M_{ij} = G(|t_i - t_j|)$. For $N = 0$ the result is evident.

Suppose now that the assertion has already been proved for $N - 1$. Since $\boldsymbol{z}$ must satisfy $\boldsymbol{z}^\top \mathbf{1}_N = 0$ as well as $\boldsymbol{z} \neq 0$, there must be some $k \in \{0, 1, \dots, N-1\}$ such that $z_k > 0$.

If $k = N$, then the fact that G is decreasing yields

$$G(|t_N - t_m|) z_N \leq G(|t_{N-1} - t_m|) z_N \quad \text{for} \quad m = 0, 1, \dots, N-1.$$

Hence, the N-dimensional vector

$$\tilde{\boldsymbol{z}} := (z_0, z_1, \dots, z_{N-2}, z_{N-1} + z_N)^\top$$

satisfies both $\tilde{\boldsymbol{z}}^\top \mathbf{1} = 0$ and $\tilde{M}\tilde{\boldsymbol{z}} > 0$, with $\tilde{M}$ corresponding to the time grid $\{t_0, t_1, \dots, t_{N-1}\}$. But by induction hypothesis this is impossible.

Next, if $k = 0$, then

$$G(t_m)z_0 \leq G(|t_m - t_1|)z_0 \quad \text{for} \quad m = 1, 2, \ldots, N.$$

Hence,

$$\hat{\boldsymbol{z}} := \big(z_0 + z_1, z_2, \ldots, z_N\big)$$

satisfies both $\hat{\boldsymbol{z}}^\top \mathbf{1} = 0$ and $\hat{M}\hat{\boldsymbol{z}} > 0$, with $\hat{M}$ corresponding to the time grid $\{t_1 - t_1, t_2 - t_1, \ldots, t_N - t_1\}$, which is again impossible due to the induction hypothesis.

Finally, let us suppose that $1 \leq k \leq N-1$. Let $\alpha \in [0,1]$ be such that $t_k = \alpha t_{k-1} + (1-\alpha)t_{k+1}$. We then have

$$G(|t_k - t_l|)z_k \leq \alpha G(|t_{k-1} - t_l|)z_k + (1-\alpha)G(|t_{k+1} - t_l|)z_k \quad \text{for } l \neq k.$$

Hence, the vector

$$\bar{\boldsymbol{z}} := (z_0, z_1, \dots, z_{k-2}, z_{k-1} + \alpha z_k, z_{k+1} + (1-\alpha)z_k, z_{k+2}, \dots, z_N)$$

satisfies both $\bar{\boldsymbol{z}}^\top \mathbf{1} = 0$ and $\bar{M}\bar{\boldsymbol{z}} > 0$, with $\bar{M}$ corresponding to the time grid

$$\{t_0, t_1, \dots, t_{k-1}, t_{k+1}, t_{k+2}, \dots, t_N\}.$$

This is again impossible due to the induction hypothesis $\square$

Sketch of proof of Theorem 2: $\mathbb{T}$ admits a countable dense subset $\{t_0, t_1, \dots\}$. For $N \in \mathbb{N}$ we define the finite set $\mathbb{T}_N := \{t_0, t_1, \dots, t_N\}$.

It follows from Theorem 3 that for each N there exists a unique optimal strategy X^N within each class $\mathcal{X}(y, \mathbb{T}_N)$, and X_t^N is a nondecreasing or nonincreasing function of $t \in \mathbb{T}_N$, depending on the sign of x. It thus follows that $\frac{1}{x}\,dX^N$ is a Borel probability measure on $\mathbb{T}$. Since the space of all Borel probability measures on $\mathbb{T}$ is compact with respect to the weak topology, there is a subsequence (X^{N_k}) that converges toward a strategy X^* in the sense of weak convergence of the associated probability measures.

Then show $\mathcal{C}(X^{(N_k)}) \to \mathcal{C}(X^*)$ as $k \uparrow \infty$ via continuity arguments.

Finally show that X^* is indeed optimal by proving that it solves the generalized Fredholm integral equation. $\square$

Qualitative properties of optimal strategies

Remark 2. (Time reversal)
Suppose for simplicity that $0 = \min \mathbb{T}$ and let $T := \max \mathbb{T}$. The time-reversed set $\check{\mathbb{T}}$ is defined by

$$\check{\mathbb{T}} := \{T - t \,|\, t \in \mathbb{T}\}$$

Similarly, the time reversal of a strategy $X \in \mathcal{X}(y, \mathbb{T})$ is defined as

$$\check{X}_t := \begin{cases} x - X_{(T-t)-} & \text{for } t < T \\ \check{X}_t := 0 & \text{for } t \geq T. \end{cases}$$

Clearly, $\check{X} \in \mathcal{X}(y, \check{\mathbb{T}})$ and $\mathcal{C}(\check{X}) = \mathcal{C}(X)$. It follows that $\check{X}^*$ is optimal in $\mathcal{X}(y, \check{\mathbb{T}})$ iff X^* is optimal in $\mathcal{X}(y, \mathbb{T})$. When $\check{\mathbb{T}} = \mathbb{T}$ (e.g. for $\mathbb{T} = [0, T]$), then $\check{X}^*$ is again optimal. When in addition G is strictly positive definite, Proposition 2 thus implies $\check{X}^* = X^*$. ◇

Theorem 4. *Let G be nonconstant, nonincreasing, and convex and suppose $x \neq 0$. Then the optimal strategy X^* in $\mathcal{X}(y, \mathbb{T})$ has impulse trades at $t_{\min} := \min \mathbb{T}$ and $t_{\max} := \max \mathbb{T}$, that is*

$$\Delta X^*_{t_{\min}} \neq 0 \text{ and } \Delta X^*_{t_{\max}} \neq 0.$$

Proof: Remark 2: enough to prove the assertion for $t_{\min}$. Moreover, w.l.o.g. $t_{\min} = 0$. Wewrite $T := t_{\max}$.

We claim that $\operatorname{supp} X^*$ must contain at least two points. Indeed, by Remark 2 the unique optimal strategy X^0 in $\mathcal{X}(y, \{0, T\})$ is given by $dX^0_t = \frac{x}{2}(\delta_0 + \delta_T)(dt)$, and so its cost is strictly smaller than the cost of any strategy whose support consists of a single point. But since $\{0, T\} \subset \mathbb{T}$ it follows that $\mathcal{C}(X^*) \leq \mathcal{C}(X^0)$, which proves our claim. Therefore

$$t_0 := \inf \big\{t \in \operatorname{supp} X \,|\, t > 0\} \in \mathbb{T}$$

By Theorem 3 we have

$$(5) \qquad \int G(|t-s|)\,dX_s^* = \int G(|u-s|)\,dX_s^* \qquad \text{for all } t,u \in \mathbb{T}$$

Let us first consider the case $t_0 > 0.$ When taking $t := 0$ and $u := t_0$ in (5), we obtain

$$(6) \qquad (G(0) - G(t_0))\Delta X_0^* = \int_{\{s \geq t_0\}} \big[G(|t_0 - s|) - G(s)\big]\,dX_s^*$$

Since G is convex, nonincreasing, and nonconstant, we have $G(0) - G(t_0) > 0$. Moreover, there must be $\varepsilon > 0$ such that $G(|t_0 - s|) - G(s) > 0$ for all $s \in [t_0, t_0 + \varepsilon]$. Since by construction $[t_0, t_0 + \varepsilon] \cap \operatorname{supp} X \neq \emptyset$, we conclude that the righthand side of (6) is nonzero. Thus, $\Delta X_0^* \neq 0$.

Now we consider the case $t_0 = 0$. We take $u > t$ and rewrite (6) as

$$
\begin{aligned}
0 &= \int \frac{G(|u-s|) - G(|t-s|)}{u-t}\, dX_s^* \\
&= \int_{\{s \le t\}} \frac{G(u-s) - G(t-s)}{u-t}\, dX_s^* \\
&\qquad + \int_{\{t < s \le u\}} \frac{G(u-s) - G(s-t)}{u-t}\, dX_s^* \\
&\qquad + \int_{\{s > u\}} \frac{G(s-u) - G(s-t)}{u-t}\, dX_s^* .
\end{aligned}
$$

When sending $u \downarrow t$, the convexity of G, monotone integration, and Lebesgue's theorem yield that each integral in the preceding sum converges.

More precisely,

$$\int_{\{s\leq t\}} \frac{G(u-s)-G(t-s)}{u-t}\,dX_s^* \longrightarrow \int_{\{s\leq t\}} G_+'(t-s)\,dX_s^*,$$

$$\int_{\{t<s\leq u\}} \frac{G(u-s)-G(s-t)}{u-t}\,dX_s^* \longrightarrow 0,$$

$$\int_{\{s>u\}} \frac{G(s-u)-G(s-t)}{u-t}\,dX_s^* \longrightarrow -\int_{\{s>t\}} G_-'(s-t)\,dX_s^*,$$

where G_+' and G_-' are the respective right- and lefthand derivatives of G.

We thus arrive at

$$\int_{\{s \le t\}} G'_+(t-s)\,dX^*_s = \int_{\{s>t\}} G'_-(s-t)\,dX^*_s.$$

Sending $t \downarrow 0$ thus yields that

$$G'_+(0)\Delta X^*_0 = \int_{\{s>0\}} G'_-(s)\,dX^*_s.$$

As in the case $t_0 > 0$ one argues that both the righthand side of this equation and the coefficient $G'_+(0)$ must be nonzero, so that $\Delta X^*_0 \neq 0$. $\square$

Now we relax the boundedness of G and assume instead

$$G \text{ is nonconstant, nonincreasing, convex, and } \int_0^1 G(t)\,dt < \infty.$$

E.g.,

$$G(t) = t^{-\gamma} \qquad \text{for } 0 < \gamma < 1, \text{ or}$$
$$G(t) = \log^-(t).$$

Let

$$\mathcal{X}_G(y, \mathbb{T}) := \Big\{ X \in \mathcal{X}(y, \mathbb{T}) \,\Big|\, \int\!\!\int G(|t-s|)\,d|X|_s\,d|X|_t < \infty \Big\}$$

Note: $\mathcal{X}_G(y, \mathbb{T})$ can be empty, e.g., for discrete $\mathbb{T}$.

Theorem 5. *When $\mathcal{X}_G(y, \mathbb{T}) \neq \emptyset$, there exists a unique optimal strategy X^* in $\mathcal{X}_G(y, \mathbb{T})$. Moreover, X^*_t is a monotone function of t.*

Sketch of proof: Show first that there exists a positive Radon measure η on $(0,\infty)$ such that

$$G(x) = G(\infty-) + \int_{(0,\infty)} (y-x)^+ \,\eta(dy) \qquad \text{for } x > 0.$$

Moreover,

$$\int_{(0,\infty)} y \wedge y^2 \,\eta(dy) < \infty \tag{7}$$

When $G(0+) = \infty$, G will not be the Fourier transform of a finite but of an infinite Radon measure μ. When $\mu([-x,x])$ grows at most polynomially, μ gives rise to a continuous linear functional $f \mapsto \int f\,du$ defined on to the Schwartz space $\mathcal{S}(\mathbb{R})$. The Fourier transform of μ is defined as the linear functional $\widehat{\mu}$ on $\mathcal{S}(\mathbb{R})$ given by

$$\widehat{\mu}(f) = \int \widehat{f}\,d\mu, \qquad f \in \mathcal{S}(\mathbb{R}).$$

Show then that G is the Fourier transform of the positive Radon measure

$$\mu(dx) = G(\infty-)\delta_0(dx) + \varphi(x)\,dx,$$

on $\mathbb{R}$, where

$$\varphi(x) = \frac{1}{\pi}\int_{(0,\infty)} \frac{1-\cos xy}{x^2}\,\eta(dy)$$

Then approximate G monotonically by the convex functions

$$G_n(x) := G(\infty-) + \int_{(0,\infty)} (y-x)^+ \mathrm{I}_{(1/n,\infty)}(y)\,\eta(dy)$$

To conclude

$$\mathcal{C}(X) = \int |\widehat{X}(z)|^2\,\mu(dz).$$

Use this approximation also to obtain existence and monotonicity of optimal strategies (as in the proof of Theorem 2). □

A set $A \subset \mathbb{R}$ will be called exceptional when there exists a G_δ-set $G \supset A$ that is a nullset for every finite Borel measure ν on $\mathbb{R}$ for which $\int\int G(|t-s|)\,\nu(ds)\,\nu(dt) < \infty$.

Clearly: $\mathcal{X}_G(y, \mathbb{T})$ is empty for $x \neq 0$ iff $\mathbb{T}$ is exceptional.

Theorem 6. *A strategy $X^* \in \mathcal{X}_G(y, \mathbb{T})$ is optimal if and only if there is a constant λ such that X^* solves the generalized Fredholm integral equation*

$$\int G(|t-s|)\,dX^*_s = \lambda \qquad \textit{for quasi every } t \in \mathbb{T}. \tag{8}$$

Moreover, λ must be nonzero as soon as $x \neq 0$.

Example 5 (Power-law decay kernel). $G(t) = t^{-\gamma}$ with $0 < \gamma < 1$

$$\int_0^1 \frac{u(s)}{|t-s|^\gamma}\,ds = 1 \qquad \text{for } 0 < t < 1,$$

is solved by

$$u^*(s) = \frac{c}{(s(1-s))^{\frac{1-\gamma}{2}}},$$

where c is a suitable constant. Thus, the unique optimal strategy in $\mathcal{X}_G(y,[0,1])$ is

$$X_t^* = x\left(1 - \frac{\Gamma(3-\gamma)}{\Gamma\big(\frac{3-\gamma}{2}\big)^2}\int_0^t \frac{1}{(s(1-s))^{\frac{1-\gamma}{2}}}\,ds\right).$$

Example 6 (Logarithmic decay kernel). $G(t) = \log^-(t)$

$$\int_0^1 u(s)G(|t-s|)\,ds = -\int_0^1 u(s)\log|t-s|\,ds = 1 \qquad \text{for } 0 < t < 1$$

solved by

$$u^*(s) = \frac{ds}{2\pi\log 2\sqrt{s(1-s)}}.$$

This fact was discovered by Carleman (1922). The unique optimal strategy in $\mathcal{X}_G(y,[0,1])$ is thus given by

$$X_t^* = y\Big(1 - \frac{1}{\pi}\int_0^t \frac{1}{\sqrt{s(1-s)}}\,ds\Big) = \frac{2y}{\pi}\arccos\sqrt{t}.$$

Conclusion:

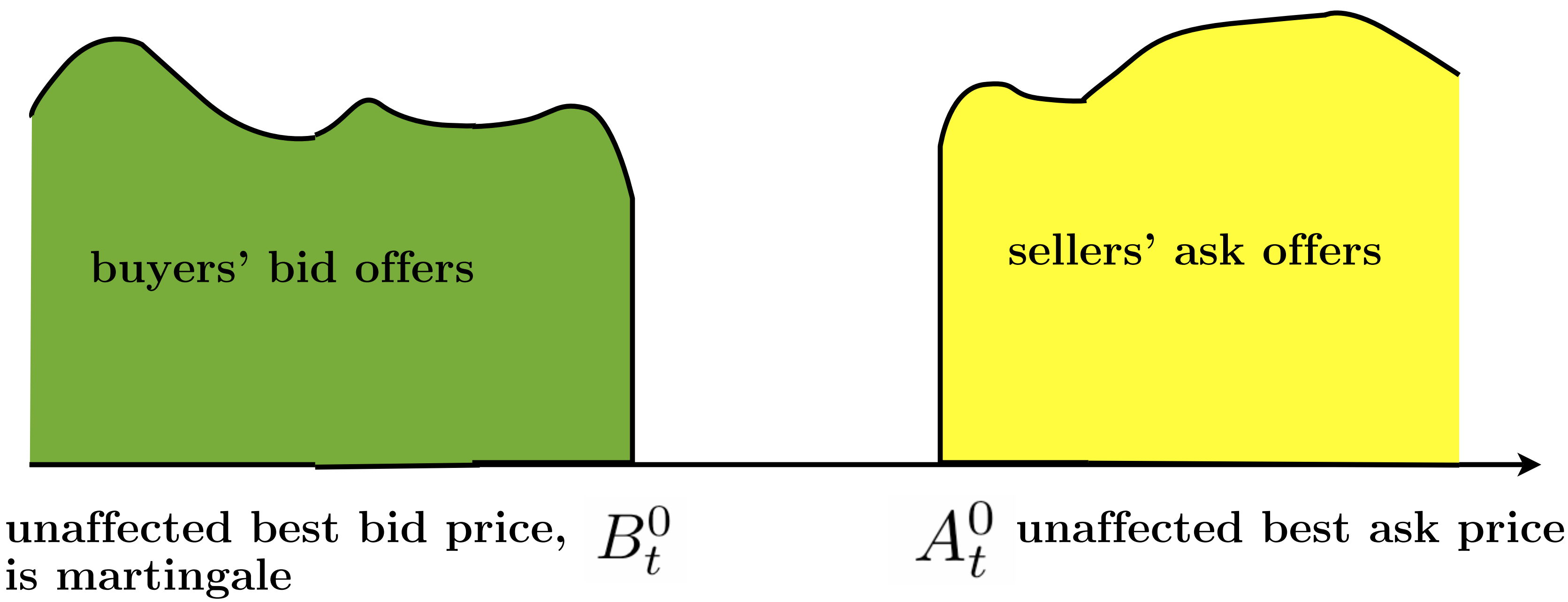
- Transient market impact can create new types of irregularities: price manipulation, transaction-triggered price manipulation
- The irregularities do not occut for convex decay of price impact
- Non-robustness with respect to G

I. Order book models

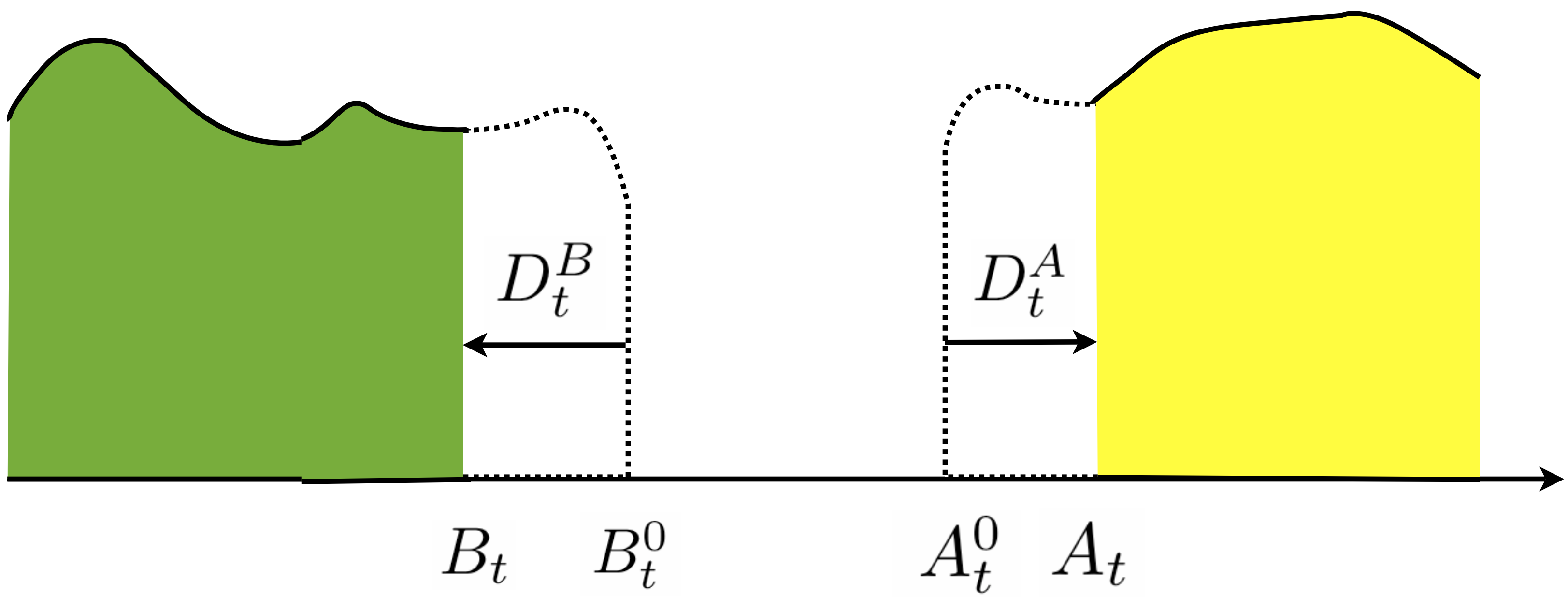
1. Linear impact, general resilience

2. Nonlinear impact, exponential resilience

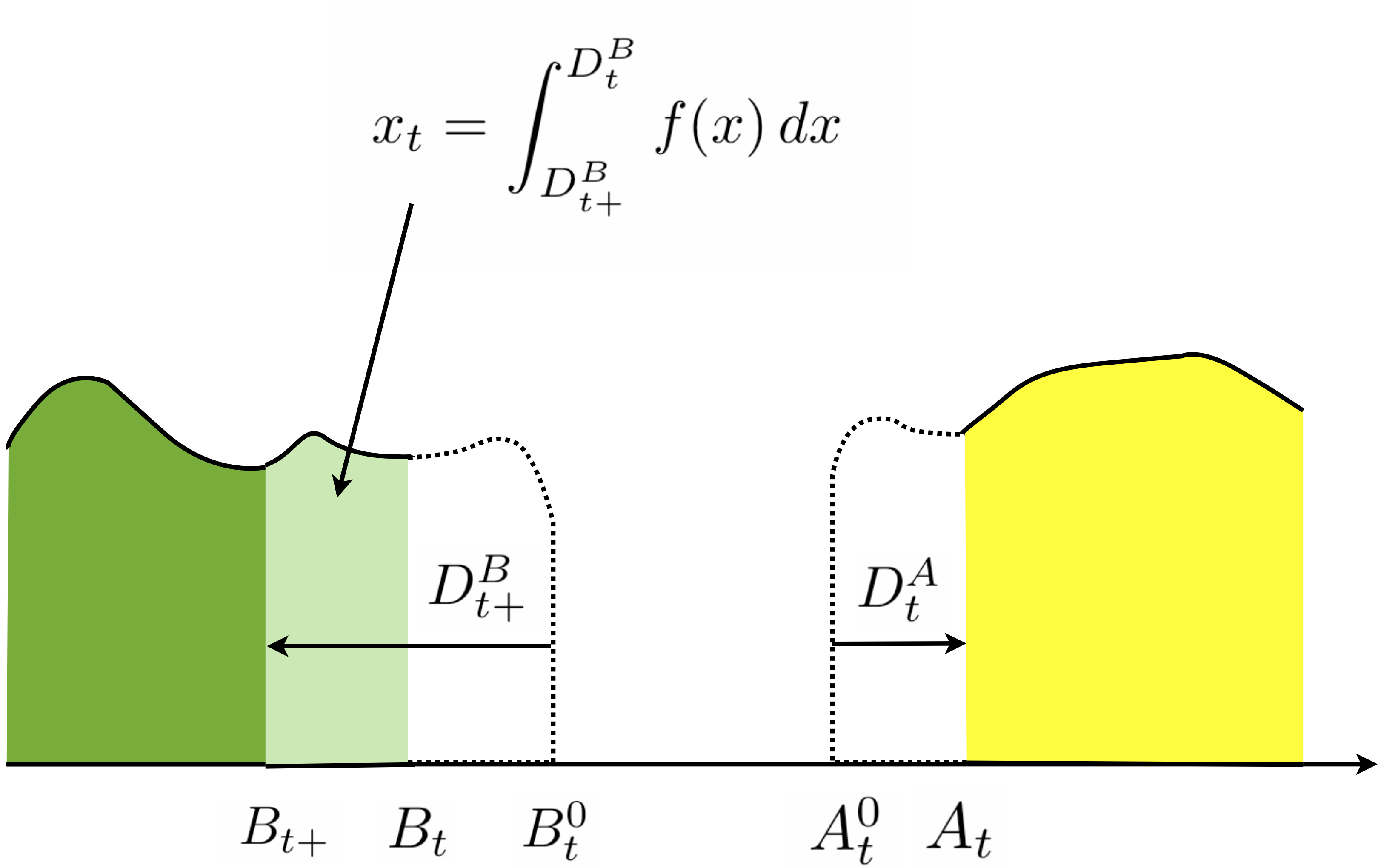
Limit order book model without large trader



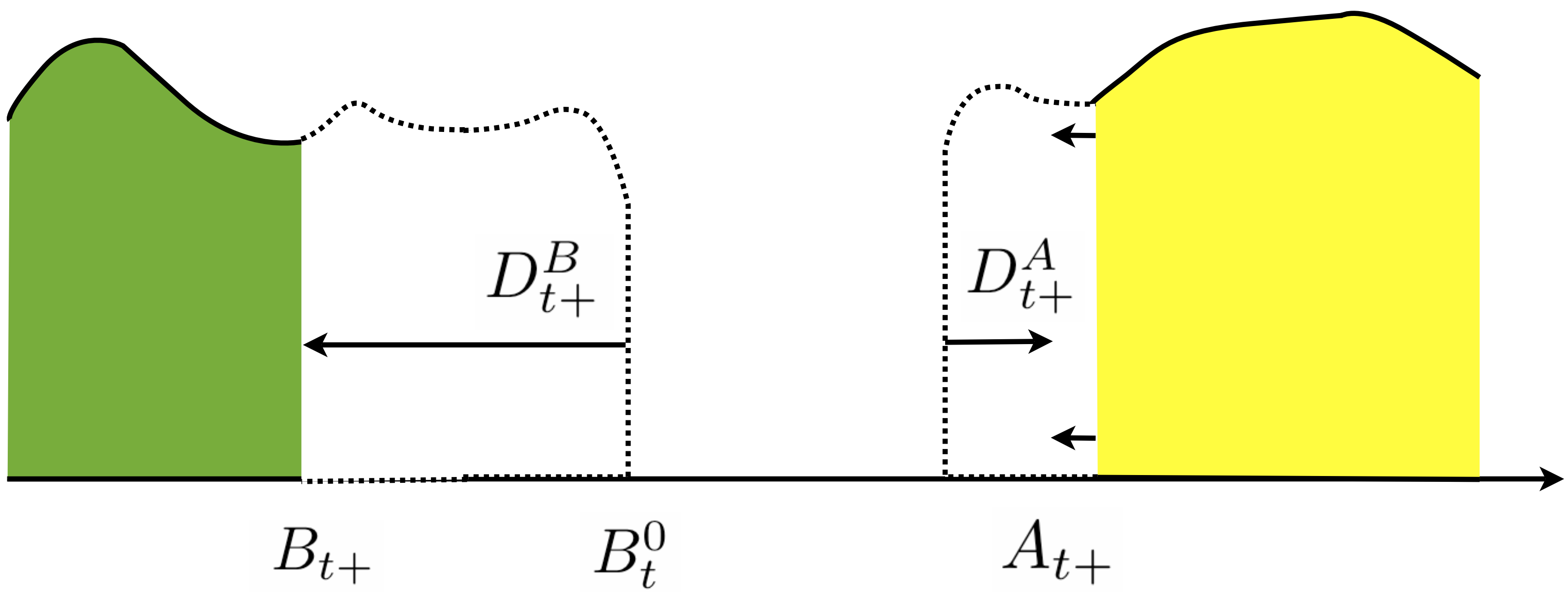
Limit order book model after large trades



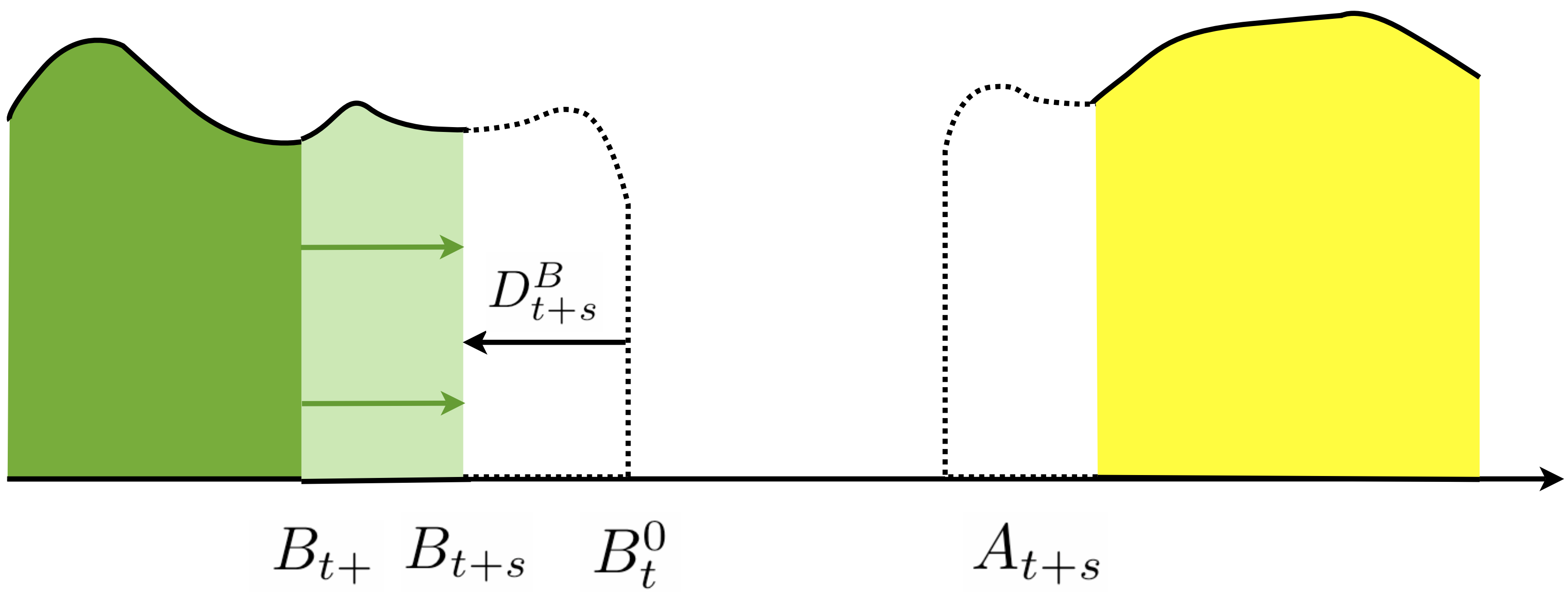
Limit order book model at large trade



Limit order book model immediately after large trade



Limit order book model with resilience



$f(x)$ = shape function = densities of bids for $x < 0$, asks for $x > 0$

B_t^0 = 'unaffected' bid price at time t, is martingale

B_t = bid price after market orders before time t

$D_t^B = B_t - B_t^0$

If sell order of $\xi_t \geq 0$ shares is placed at time t:

D_t^B changes to D_{t+}^B, where

$$\int_{D_t^B}^{D_{t+}^B} f(x)dx = -\xi_t$$

and

$$B_{t+} := B_t + D_{t+}^B - D_t^B = B_t^0 + D_{t+}^B,$$

$\Longrightarrow$ nonlinear price impact

A_t^0 = 'unaffected' ask price at time t, satisfies $B_t^0 \leq A_t^0$

A_t = bid price after market orders before time t

$D_t^A = A_t - A_t^0$

If buy order of $\xi_t \leq 0$ shares is placed at time t:

D_t^A changes to D_{t+}^A, where

$$\int_{D_t^A}^{D_{t+}^A} f(x)dx = -\xi_t$$

and

$$A_{t+} := A_t + D_{t+}^A - D_t^A = A_t^0 + D_{t+}^A,$$

For simplicity, we assume that the LOB has infinite depth, i.e., $|F(x)| \to \infty$ as $|x| \to \infty$, where

$$F(x) := \int_0^x f(y)\,dy.$$

If the large investor is inactive during the time interval $[t, t+s[$, there are *two* possibilities:

- **Exponential recovery of the extra spread**

$$D_t^B = e^{-\int_s^t \rho_r \, dr} D_s^B \qquad \text{for } s < t.$$

- **Exponential recovery of the order book volume**

$$E_t^B = e^{-\int_s^t \rho_r \, dr} E_s^B \qquad \text{for } s < t,$$

where

$$E_t^B = \int_{D_t^B}^0 f(x) \, dx =: F(D_t^B).$$

In both cases: analogous dynamics for D^A or E^A

Strategy:

$N+1$ market orders: ξ_n shares placed at time τ_n s.th.

a) the (τ_n) are stopping times s.th. $0=\tau_0 \leq \tau_1 \leq \cdots \leq \tau_N = T$

b) ξ_n is $\mathcal{F}_{\tau_n}$-measurable and bounded from below,

c) we have $\displaystyle\sum_{n=0}^{N} \xi_n = X_0$

Will write

$$(\boldsymbol{\tau}, \boldsymbol{\xi})$$

and optimize jointly over $\boldsymbol{\tau}$ and $\boldsymbol{\xi}$.

- When selling $\xi_n > 0$ shares, we sell $f(x)\,dx$ shares at price $B^0_{\tau_n} + x$ with x ranging from $D^B_{\tau_n}$ to $D^B_{\tau_n+} < D^B_{\tau_n}$, i.e., the costs are negative:

$$c_n(\boldsymbol{\tau},\boldsymbol{\xi}) := \int_{D^B_{\tau_n}}^{D^B_{\tau_n+}} (B^0_{\tau_n} + x) f(x)\,dx = -\xi_n B^0_{\tau_n} + \int_{D^B_{\tau_n}}^{D^B_{\tau_n+}} x f(x)\,dx$$

- When buying shares ($\xi_n < 0$), the costs are positive:

$$c_n(\boldsymbol{\tau},\boldsymbol{\xi}) := -\xi_n A^0_{\tau_n} + \int_{D^A_{\tau_n}}^{D^A_{\tau_n+}} x f(x)\,dx$$

- The expected costs for the strategy $(\boldsymbol{\tau},\boldsymbol{\xi})$ are

$$\mathcal{C}(\boldsymbol{\tau},\boldsymbol{\xi}) = \mathbb{E}\Big[\sum_{n=0}^{N} c_n(\boldsymbol{\tau},\boldsymbol{\xi})\Big]$$

Instead of the τ_k, we will use

$$\alpha_k := \int_{\tau_{k-1}}^{\tau_k} \rho_s ds, \qquad k = 1, \ldots, N. \tag{9}$$

The condition $0 = \tau_0 \le \tau_1 \le \cdots \le \tau_N = T$ is equivalent to $\boldsymbol{\alpha} := (\alpha_1, \ldots, \alpha_N)$ belonging to

$$\mathcal{A} := \Big\{ \boldsymbol{\alpha} := (\alpha_1, \ldots, \alpha_N) \in \mathbb{R}_+^N \,\Big|\, \sum_{k=1}^N \alpha_k = \int_0^T \rho_s \, ds \Big\}.$$

A simplified model without bid-ask spread

S_t^0 = unaffected price, is (continuous) martingale.

$$S_{t_n} = S_{t_n}^0 + D_n$$

where D and E are defined as follows:

$$E_0 = D_0 = 0, \qquad E_n = F(D_n) \qquad \text{and} \qquad D_n = F^{-1}(E_n).$$

For $n = 0, \dots, N$, regardless of the sign of ξ_n,

$$E_{n+} = E_n - \xi_n \qquad \text{and} \qquad D_{n+} = F^{-1}(E_{n+}) = F^{-1}\left(F\left(D_n\right) - \xi_n\right).$$

For $k = 0, \dots, N-1$,

$$E_{k+1} = e^{-\alpha_{k+1}} E_{k+} = e^{-\alpha_{k+1}}(E_k - \xi_k)$$

The costs are

$$\overline{c}_n(\boldsymbol{\tau}, \boldsymbol{\xi}) = -\xi_n S_{\tau_n}^0 + \int_{D_{\tau_n}}^{D_{\tau_n+}} x f(x)\, dx$$

Lemma 2. *Suppose that* $S^0 = B^0$. *Then, for any strategy* $\boldsymbol{\xi}$,

$$\overline{c}_n(\boldsymbol{\xi}) \le c_n(\boldsymbol{\xi}) \qquad \text{with equality if } \xi_k \ge 0 \text{ for all } k.$$

Moreover,

$$\overline{\mathcal{C}}(\boldsymbol{\tau},\boldsymbol{\xi}) := \mathbb{E}\Big[\sum_{n=0}^{N} \overline{c}_n(\boldsymbol{\tau},\boldsymbol{\xi})\Big] = \mathbb{E}\Big[C(\boldsymbol{\alpha},\boldsymbol{\xi})\Big] - X_0 S_0^0$$

where

$$C(\boldsymbol{\alpha},\boldsymbol{\xi}) := \sum_{n=0}^{N} \int_{D_n}^{D_{n+}} x f(x)\, dx$$

is a deterministic function of $\boldsymbol{\alpha} \in \mathcal{A}$ *and* $\boldsymbol{\xi} \in \mathbb{R}^{N+1}$.

Implies that is is enough to minimize $C(\boldsymbol{\alpha},\boldsymbol{\xi})$ over $\boldsymbol{\alpha} \in \mathcal{A}$ and

$$\boldsymbol{\xi} \in \Big\{\boldsymbol{x} = (x_0,\dots,x_N) \in \mathbb{R}^{N+1} \,\Big|\, \sum_{n=0}^{N} x_n = X_0\Big\}.$$

Theorem 7. *Suppose f is increasing on $\mathbb{R}_-$ and decreasing on $\mathbb{R}_+$. Then there is a unique optimal strategy $(\boldsymbol{\xi}^*, \boldsymbol{\tau}^*)$ consisting of homogeneously spaced trading times,*

$$\int_{\tau_i^*}^{\tau_{i+1}^*} \rho_r\, dr = \frac{1}{N}\int_0^T \rho_r\, dr =: -\log a,$$

and trades defined via

$$F^{-1}\left(X_0 - N\xi_0^*\left(1-a\right)\right) = \frac{F^{-1}(\xi_0^*) - aF^{-1}(a\xi_0^*)}{1-a},$$

and

$$\xi_1^* = \cdots = \xi_{N-1}^* = \xi_0^*\left(1-a\right),$$

as well as

$$\xi_N^* = X_0 - \xi_0^* - (N-1)\xi_0^*\left(1-a\right).$$

Moreover, $\xi_i^* > 0$ for all i.

Taking $X_0 \downarrow 0$ yields:

Corollary 1. *Both the original and simplified models admit neither ordinary nor transaction-triggered price manipulation strategies.*

$$f(x) = \frac{1}{1 + |x|}$$

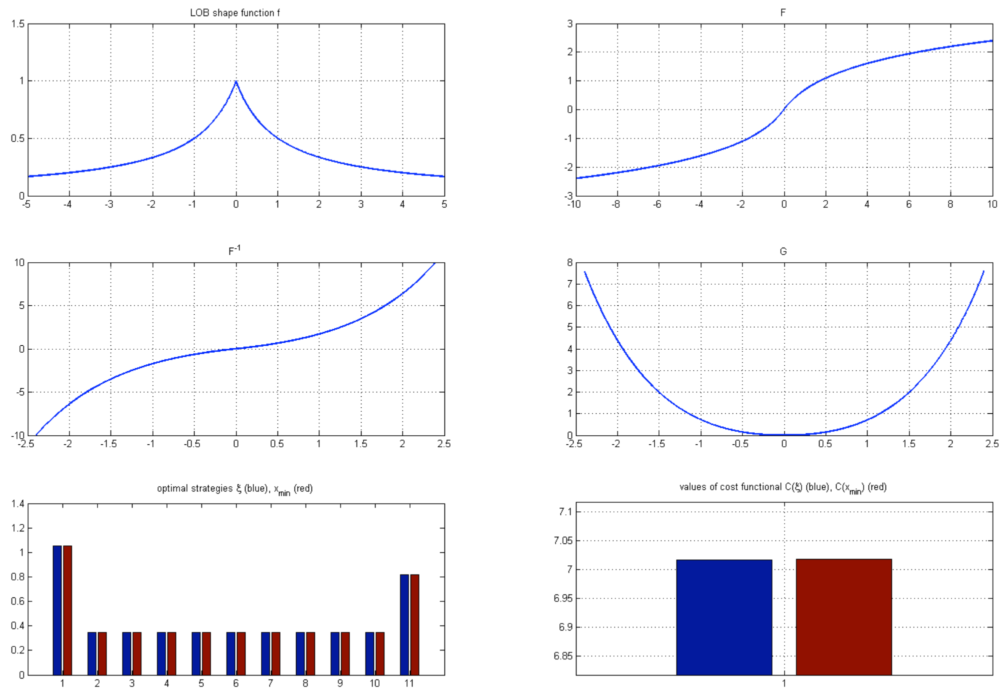


Figure 1: f, F, F^{-1}, G and optimal strategy

Strategy of proving Theorem 7:

(a) Show that there exists a (unique) minimizer $\boldsymbol{x}^*(\boldsymbol{\alpha})$ for each $\boldsymbol{\alpha}$.
(Prove that $C(\boldsymbol{\alpha},\boldsymbol{x}) \to \infty$ for $|\boldsymbol{x}| \to \infty$)

(b) Show that all components of $\boldsymbol{x}^*(\boldsymbol{\alpha})$ are positive
(Use that $\boldsymbol{x}^*(\boldsymbol{\alpha})$ must be a critical point of $\boldsymbol{x} \to C(\boldsymbol{\alpha},\boldsymbol{x}) - \nu\boldsymbol{x}^\top\mathbf{1}$ for some Lagrange multiplier ν. Then compute gradient of $C(\boldsymbol{\alpha},\cdot)$ and use explicit estimates....)

(c) By (a) and (b) we can restrict the optimization of $C(\boldsymbol{\alpha},\boldsymbol{x})$ to $(\boldsymbol{\alpha},\boldsymbol{x})$ belonging to the compact simplex

$$\mathcal{A} \times \Big\{\boldsymbol{x} \in \mathbb{R}^{N+1} \,\Big|\, x_i \geq 0 \text{ and } \sum_{n=0}^{N} x_n = X_0\Big\}.$$

Hence a minimizer $(\boldsymbol{\alpha}^*,\boldsymbol{x}^*)$ exists.

(d) Use again Lagrange multipliers to identify $(\boldsymbol{\alpha}^*,\boldsymbol{x}^*)$:

Let us introduce the functions

$$\tilde{F}(x) := \int_0^x z f(z)\, dz \qquad \text{and} \qquad G = \tilde{F} \circ F^{-1}.$$

Then, since $D_n = F^{-1}(E_n)$ and $D_{n+} = F^{-1}(E_{n+})$

$$\begin{aligned} C(\boldsymbol{\alpha}, \boldsymbol{x}) &= \sum_{n=0}^{N} \int_{D_n}^{D_{n+}} x f(x)\, dx = \sum_{n=0}^{N} \Big[\widetilde{F}(D_{n+}) - \widetilde{F}(D_n) \Big] \\ &= \sum_{n=0}^{N} \Big[G(E_{n+}) - G(E_n) \Big] = \sum_{n=0}^{N} \Big[G(E_n - x_n) - G(E_n) \Big] \end{aligned}$$

where

$$E_0 = 0 \qquad \text{and} \qquad E_n = -\sum_{i=0}^{n-1} x_i e^{-\sum_{k=i+1}^{n} \alpha_k}, \qquad 1 \le n \le N.$$

Lemma 3. *For $i = 0, \dots, N-1$, we have the following recursive formula,*

$$\frac{\partial C}{\partial x_i} = e^{-\alpha_{i+1}} F^{-1}(E_{i+1}) - F^{-1}(E_i - x_i) + e^{-\alpha_{i+1}} \frac{\partial C}{\partial x_{i+1}}. \tag{10}$$

Moreover, for $i = 1, \dots, N$,

$$\frac{\partial C}{\partial \alpha_i} = E_i \sum_{n=i}^{N} \big[F^{-1}(E_n - x_n) - F^{-1}(E_n)\big] e^{-\sum_{k=i+1}^{n} \alpha_k}. \tag{11}$$

When $(\boldsymbol{\alpha}, \boldsymbol{x})$ is a minimizer, then it is a critical point of

$$(\boldsymbol{\beta}, \boldsymbol{y}) \longmapsto C(\boldsymbol{\beta}, \boldsymbol{y}) - \nu \boldsymbol{y}^\top \mathbf{1} - \lambda \boldsymbol{\beta}^\top \mathbf{1}.$$

Hence

$$\frac{\partial C}{\partial x_i} = \nu \qquad \text{and} \qquad \frac{\partial C}{\partial \alpha_j} = \lambda \qquad \text{for all } i, j$$

Plugging this into (10) yields $\nu = -F^{-1}(E_N - x_N)$ and

$$\nu = e^{-\alpha_{i+1}}F^{-1}(E_{i+1}) - F^{-1}(E_i - x_i) + e^{-\alpha_{i+1}}\nu$$

or, since $E_{i+1} = e^{-\alpha_{i+1}}(E_i - x_i)$,

$$\nu = -\frac{F^{-1}(E_i - x_i) - a_{i+1}F^{-1}(a_{i+1}(E_i - x_i))}{1 - a_{i+1}}$$

where $a_{i+1} = e^{-\alpha_{i+1}}$.

Plugging this into (10) yields $\nu = -F^{-1}(E_N - x_N)$ and

$$\nu = e^{-\alpha_{i+1}} F^{-1}(E_{i+1}) - F^{-1}(E_i - x_i) + e^{-\alpha_{i+1}}\nu$$

or, since $E_{i+1} = e^{-\alpha_{i+1}}(E_i - x_i)$,

$$\nu = -\frac{F^{-1}(E_i - x_i) - a_{i+1}F^{-1}(a_{i+1}(E_i - x_i))}{1 - a_{i+1}}$$

where $a_{i+1} = e^{-\alpha_{i+1}}$.

Similarly,

$$\begin{aligned} \frac{\lambda}{E_j} &= \sum_{n=j}^{N} \big[F^{-1}(E_n - x_n) - F^{-1}(E_n)\big] e^{-\sum_{k=j+1}^{n} \alpha_k} \\ &= -F^{-1}(E_j) + \big[F^{-1}(E_j - x_j) - F^{-1}(E_{j+1})e^{-\alpha_{j+1}}\big] + \dots \\ &\quad + \big[F^{-1}(E_{N-1} - x_{N-1}) - F^{-1}(E_N)e^{-\alpha_N}\big] e^{-\sum_{k=j+1}^{N-1} \alpha_k} \\ &\quad + F^{-1}(E_N - x_N) e^{-\sum_{k=j+1}^{N} \alpha_k} \end{aligned}$$

Plugging this into (10) yields $\nu = -F^{-1}(E_N - x_N)$ and

$$\nu = e^{-\alpha_{i+1}} F^{-1}(E_{i+1}) - F^{-1}(E_i - x_i) + e^{-\alpha_{i+1}}\nu$$

or, since $E_{i+1} = e^{-\alpha_{i+1}}(E_i - x_i)$,

$$\nu = -\frac{F^{-1}(E_i - x_i) - a_{i+1}F^{-1}(a_{i+1}(E_i - x_i))}{1 - a_{i+1}}$$

where $a_{i+1} = e^{-\alpha_{i+1}}$.

Similarly,

$$\begin{aligned}
\frac{\lambda}{E_j} &= \sum_{n=j}^{N} \big[F^{-1}(E_n - x_n) - F^{-1}(E_n)\big] e^{-\sum_{k=j+1}^{n} \alpha_k} \\
&= -F^{-1}(E_j) + \big[F^{-1}(E_j - x_j) - F^{-1}(E_{j+1})e^{-\alpha_{j+1}}\big] + \dots \\
&\quad + \big[F^{-1}(E_{N-1} - x_{N-1}) - F^{-1}(E_N)e^{-\alpha_N}\big] e^{-\sum_{k=j+1}^{N-1} \alpha_k} \\
&\quad + F^{-1}(E_N - x_N) e^{-\sum_{k=j+1}^{N} \alpha_k}
\end{aligned}$$

$$
\begin{aligned}
&= -F^{-1}(E_j) - (1 - e^{-\alpha_{j+1}})\nu - \cdots - (1 - e^{-\alpha_N})\nu e^{-\sum_{k=j+1}^{N-1}\alpha_k} \\
&\qquad -\nu e^{-\sum_{k=j+1}^{N}\alpha_k} \\
&= -F^{-1}(E_j) - \nu
\end{aligned}
$$

Hence

$$
\begin{aligned}
\lambda &= -E_j(F^{-1}(E_j) + \nu) \\
&= E_j\left[\frac{F^{-1}(E_j - x_j) - a_{j+1}F^{-1}(a_{j+1}(E_j - x_j))}{1 - a_{j+1}} - F^{-1}(E_j)\right]
\end{aligned}
$$

Altogether:

$$\nu = -\frac{F^{-1}(E_{i-1}-x_{i-1}) - e^{-\alpha_i}F^{-1}(e^{-\alpha_i}(E_{i-1}-x_{i-1}))}{1-e^{-\alpha_i}},$$

$$\lambda = e^{-\alpha_i}(E_{i-1}-x_{i-1})\frac{F^{-1}(E_{i-1}-x_{i-1}) - F^{-1}(e^{-\alpha_i}(E_{i-1}-x_{i-1}))}{1-e^{-\alpha_i}},$$

for $i = 1, \ldots, N$.

$$
\begin{aligned}
\nu &= -\frac{F^{-1}(E_{i-1}-x_{i-1}) - e^{-\alpha_i}F^{-1}(e^{-\alpha_i}(E_{i-1}-x_{i-1}))}{1-e^{-\alpha_i}},\\
\lambda &= e^{-\alpha_i}(E_{i-1}-x_{i-1})\frac{F^{-1}(E_{i-1}-x_{i-1}) - F^{-1}(e^{-\alpha_i}(E_{i-1}-x_{i-1}))}{1-e^{-\alpha_i}},
\end{aligned}
$$

for $i = 1, \dots, N$.

Lemma 4. *Given ν and λ, these equations uniquely determine α_i and $E_{i-1} - x_{i-1}$*

It follows that

$$\alpha_1 = \cdots = \alpha_N \qquad \text{and} \qquad -x_0 = E_1 - x_1 = \cdots = E_{N-1} - x_{N-1}.$$

The theorem now follows easily. □

Robustness of the optimal strategy
[Plots by C. Lorenz (2009)]

First figure:

$$f(x) = \frac{1}{1 + |x|}$$

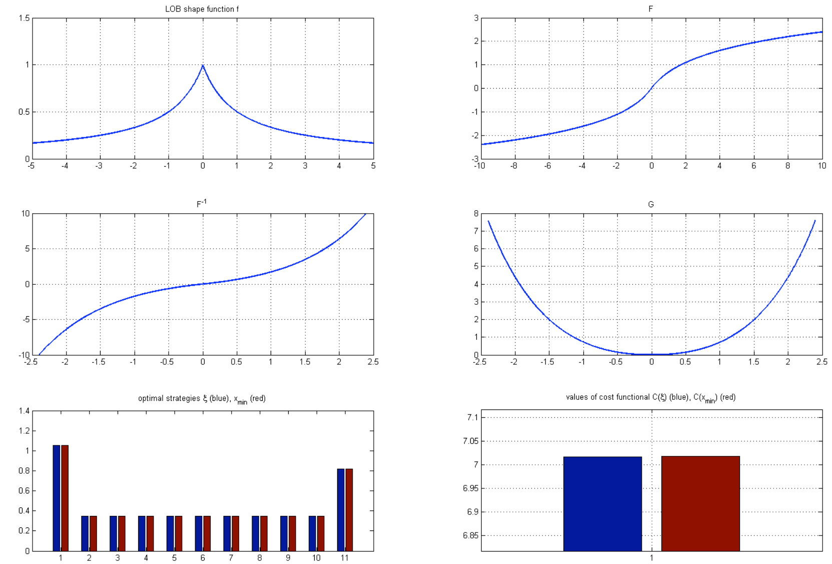


Figure 2: f, F, F^{-1}, G and optimal strategy

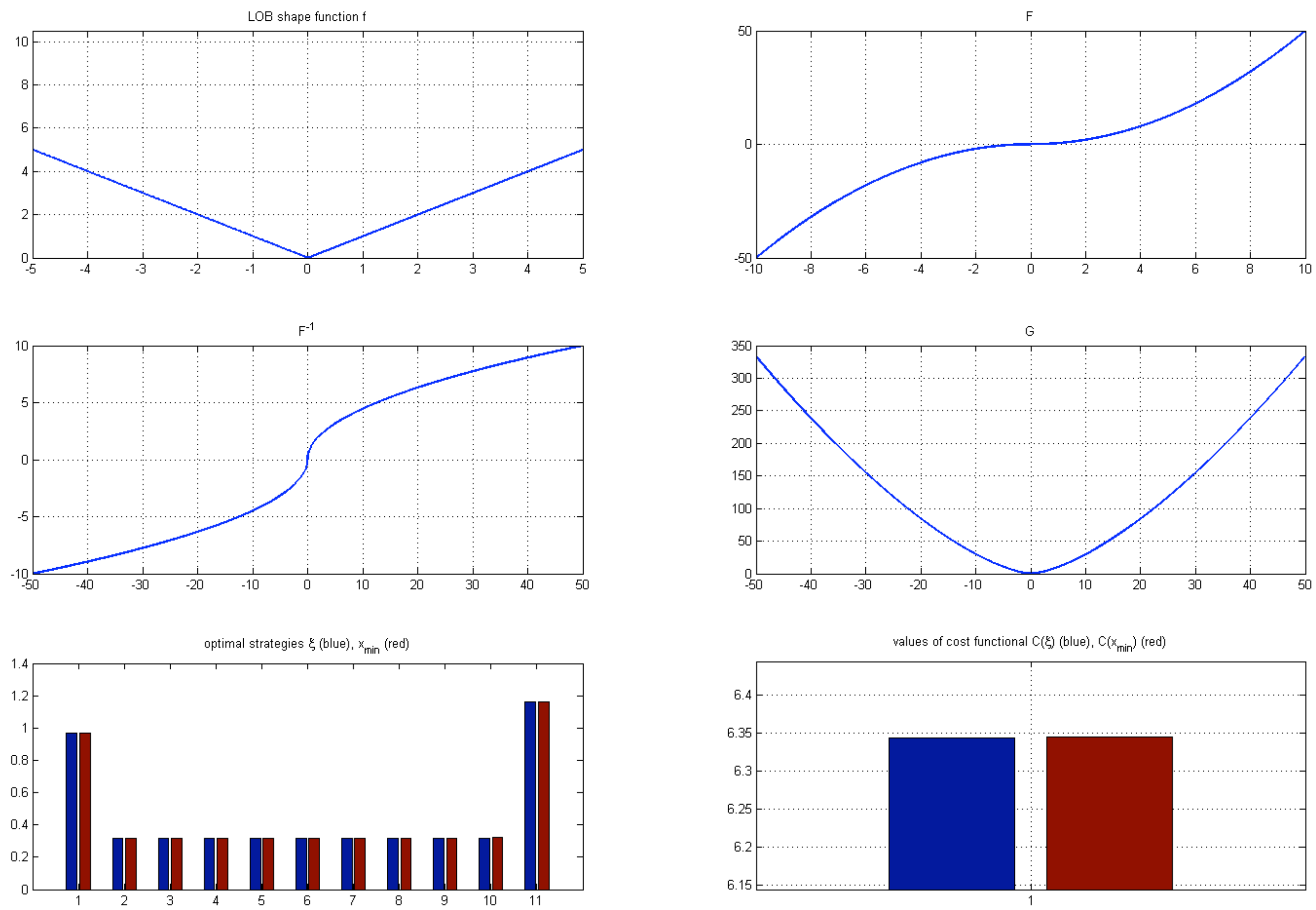

LOB shape function f
F
F^{-1}
G
optimal strategies ξ (blue), x_{min} (red)
values of cost functional C(ξ) (blue), C(x_{min}) (red)

Figure 3: $f(x) = |x|$

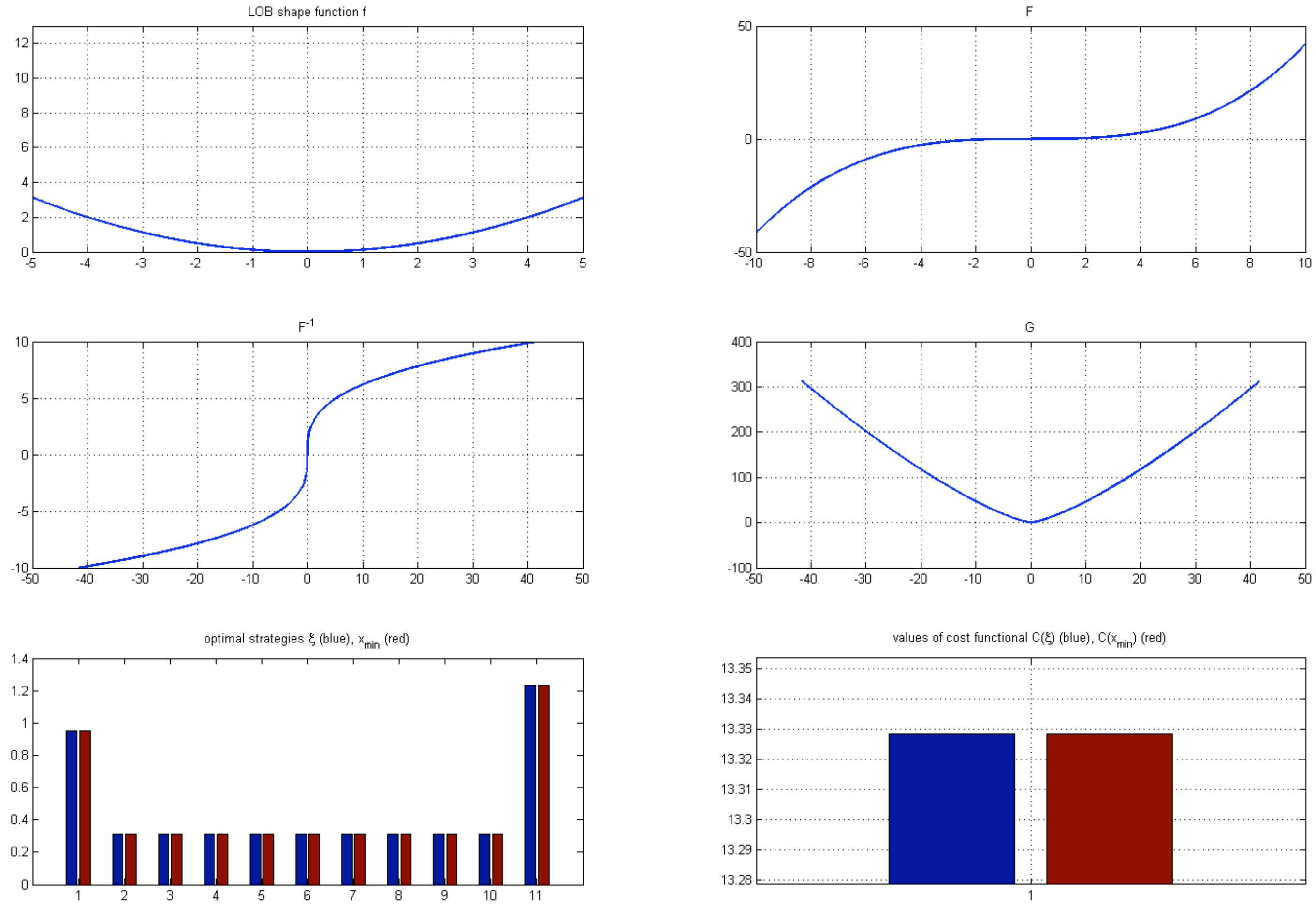


Figure 4: $f(x) = \frac{1}{8}x^2$

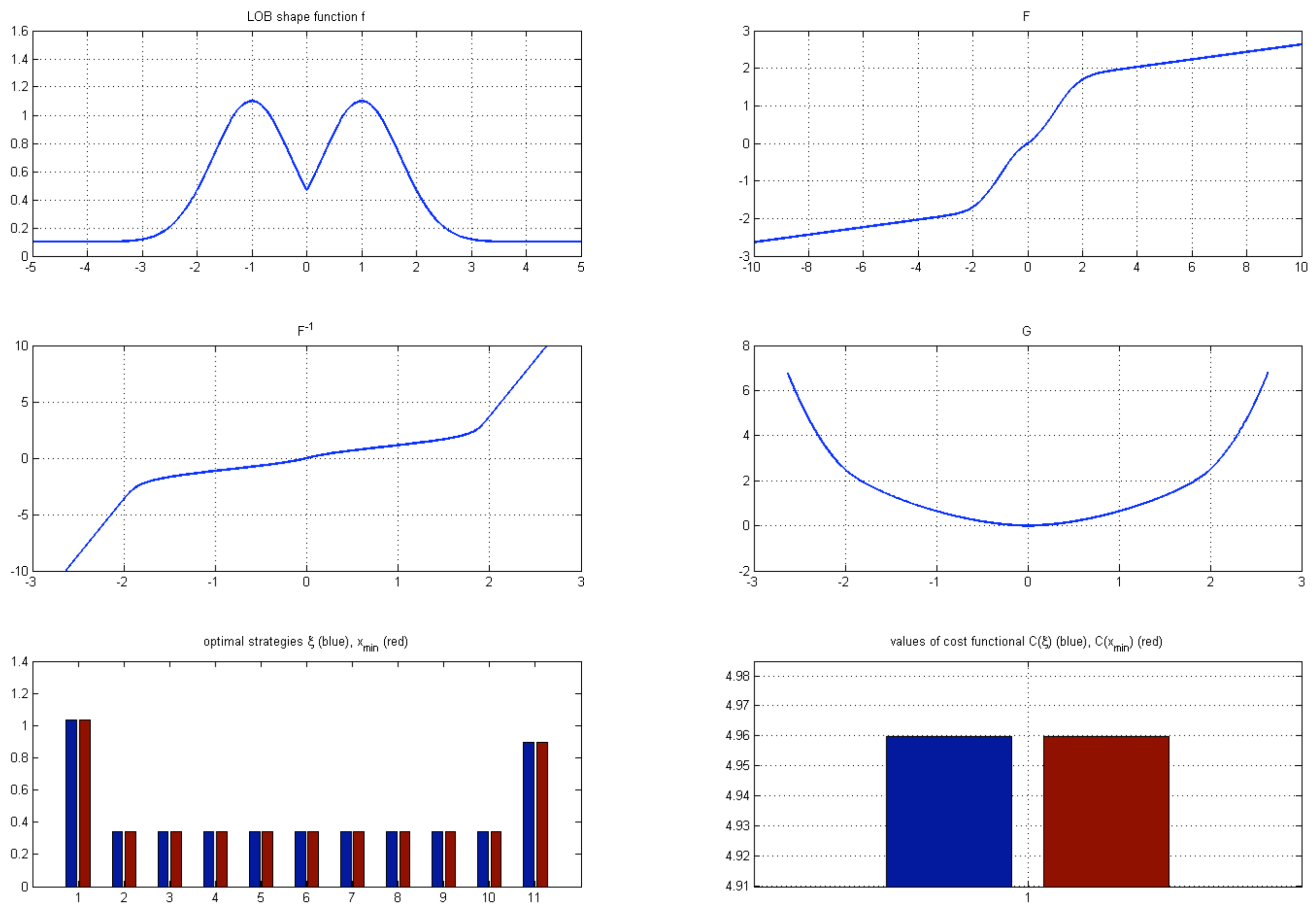


Figure 5: $f(x) = \exp(-(|x|-1)^2) + 0.1$

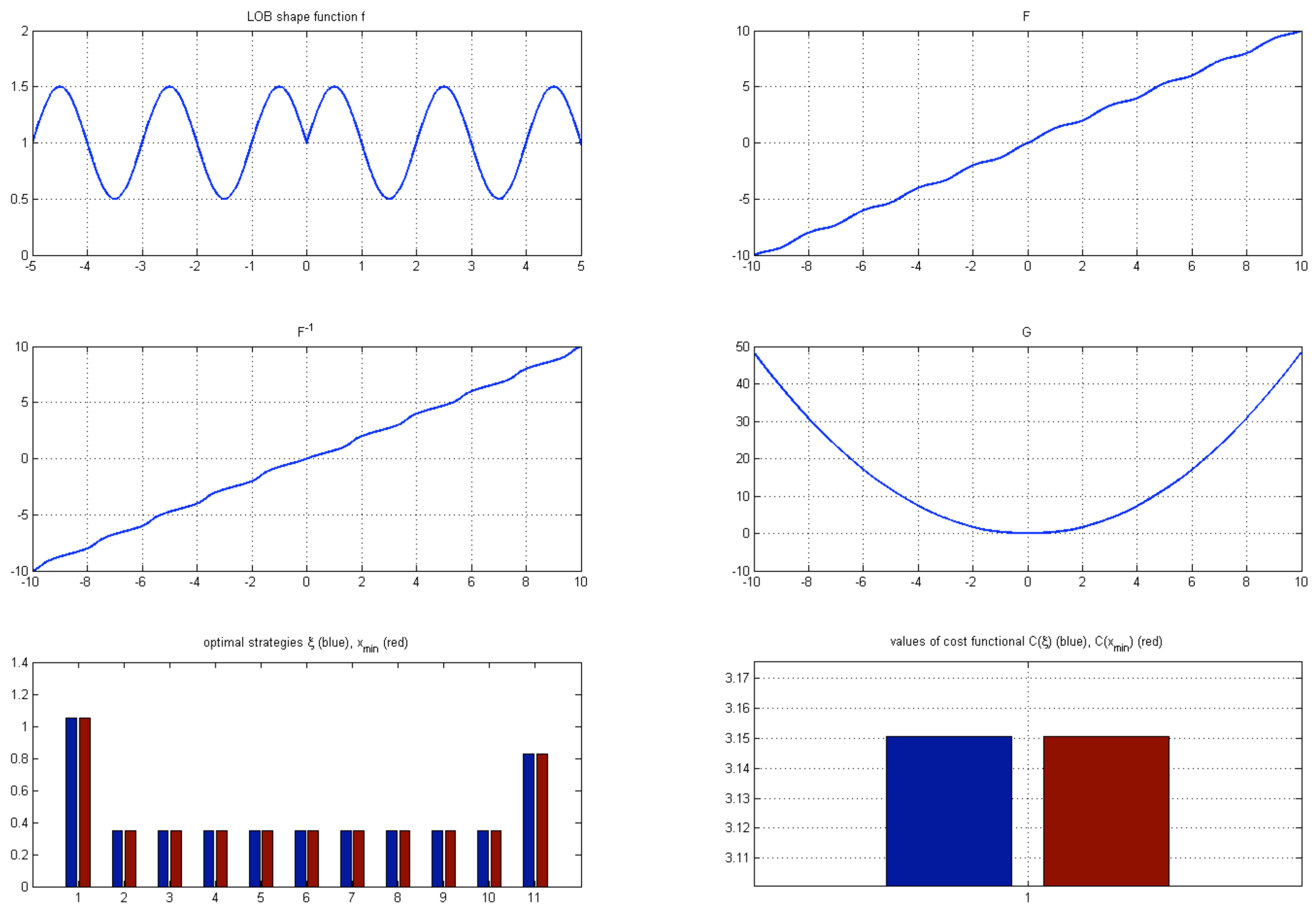


Figure 6: $f(x) = \frac{1}{2}\sin(\pi|x|) + 1$

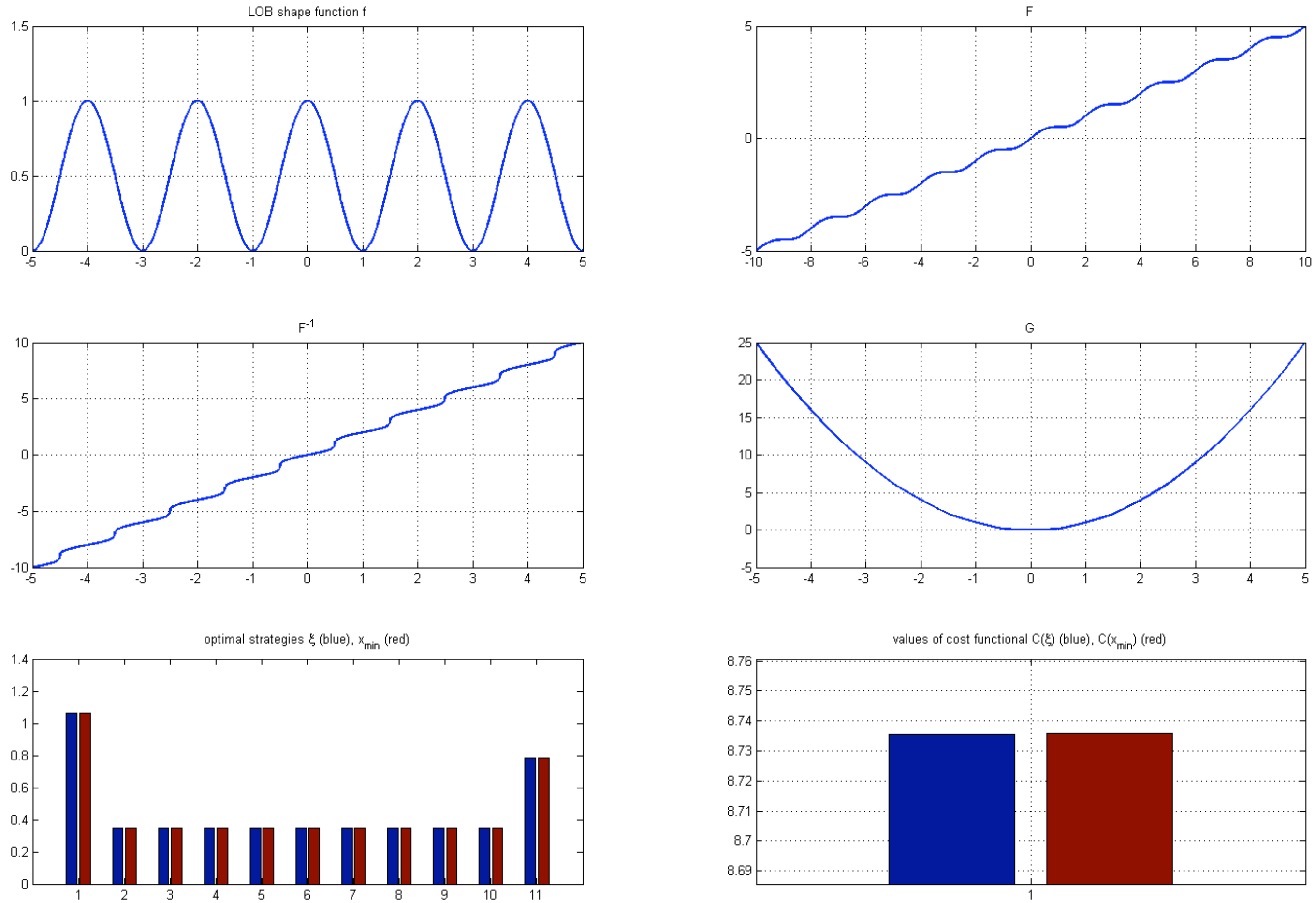


Figure 7: $f(x) = \frac{1}{2}\cos(\pi|x| + \frac{1}{2})$

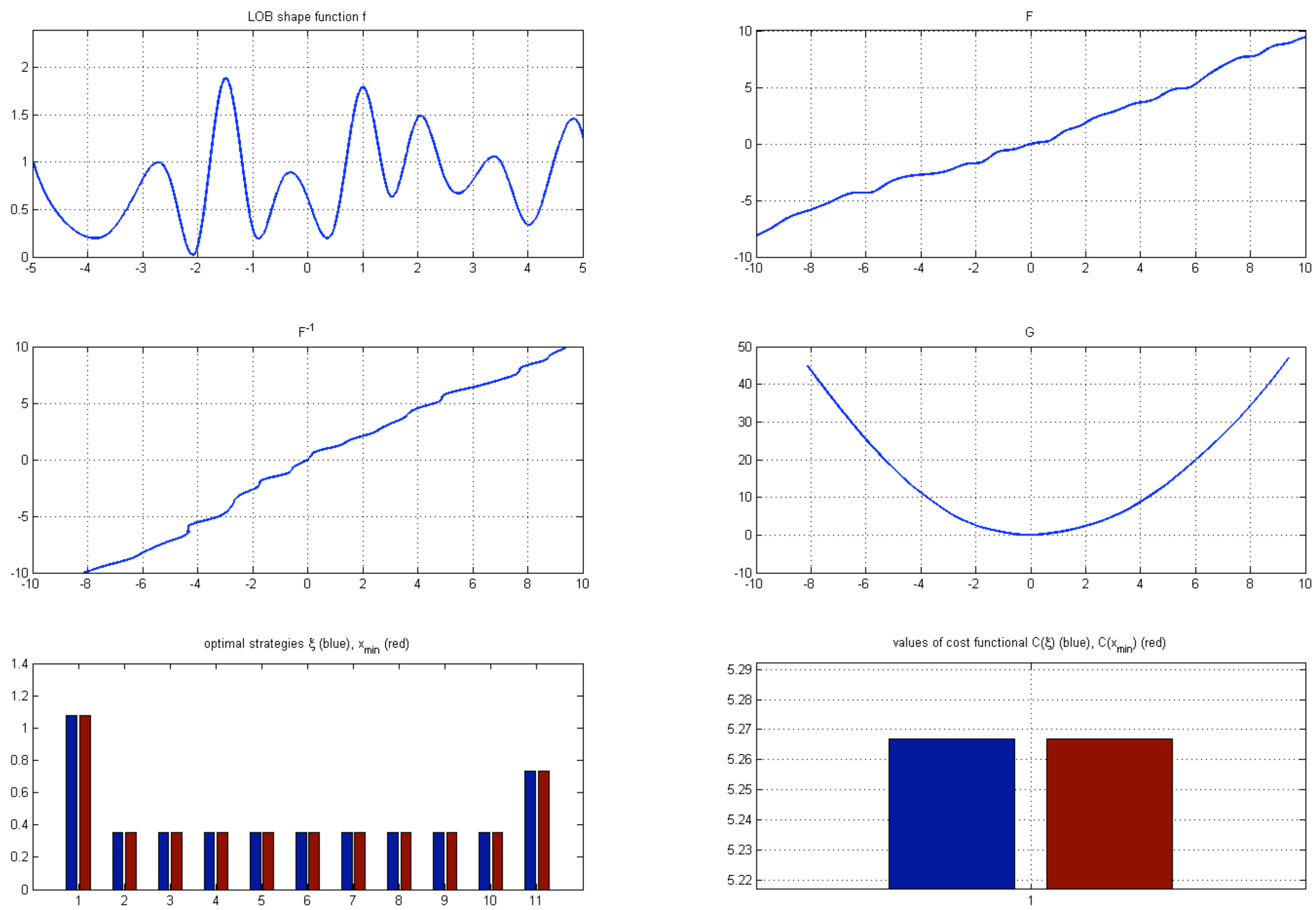


Figure 8: f random

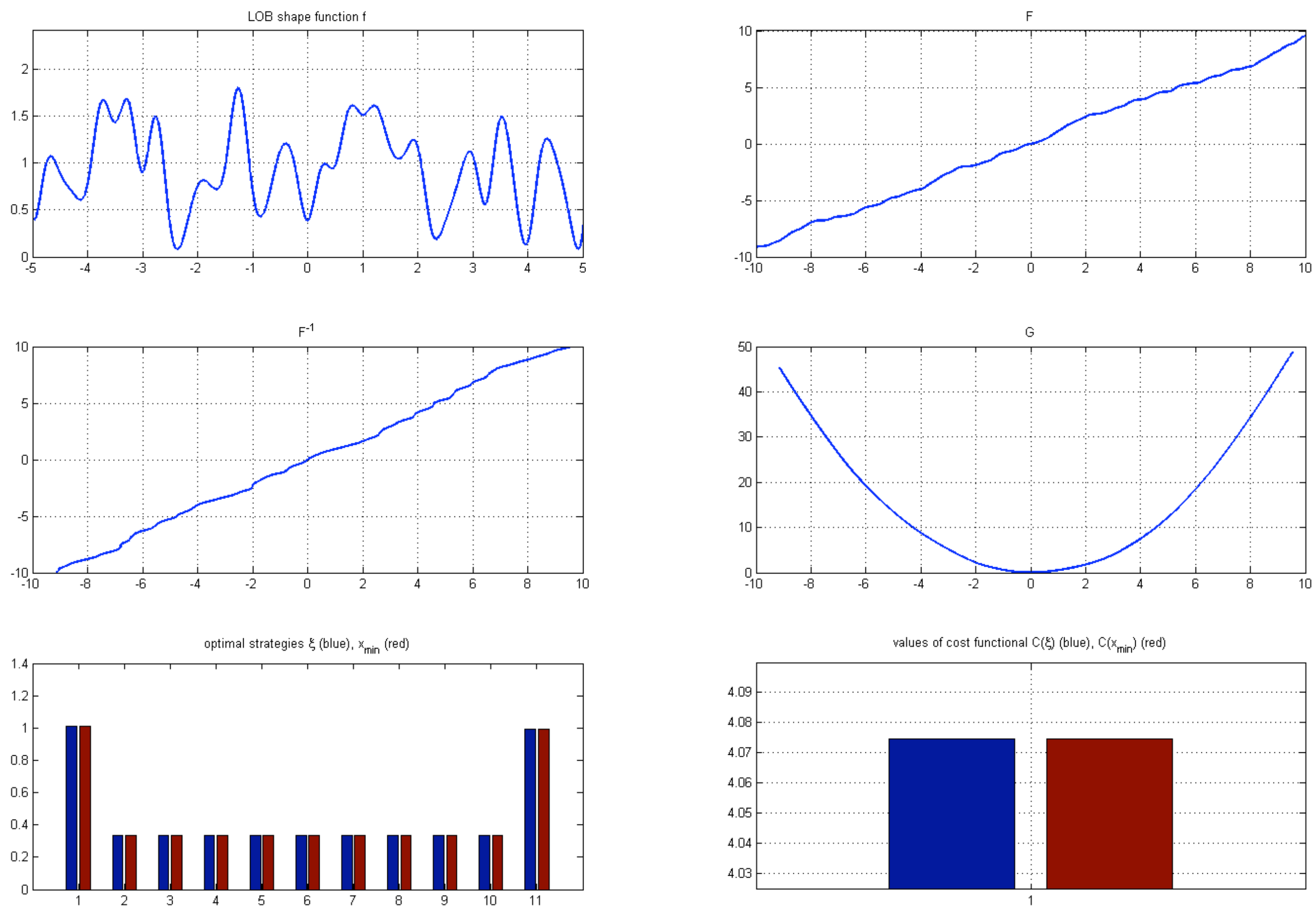


Figure 9: f random

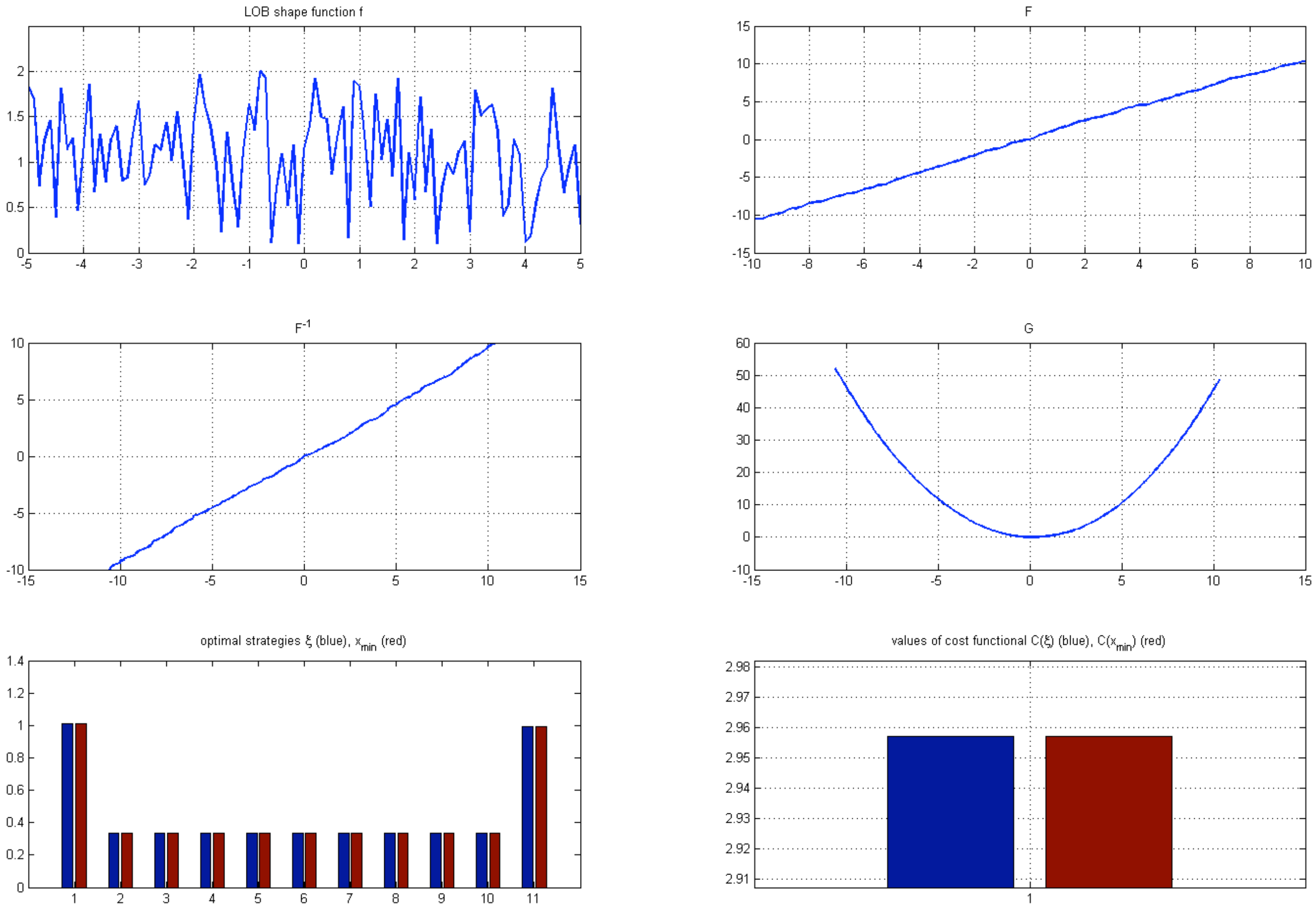


Figure 10: f random

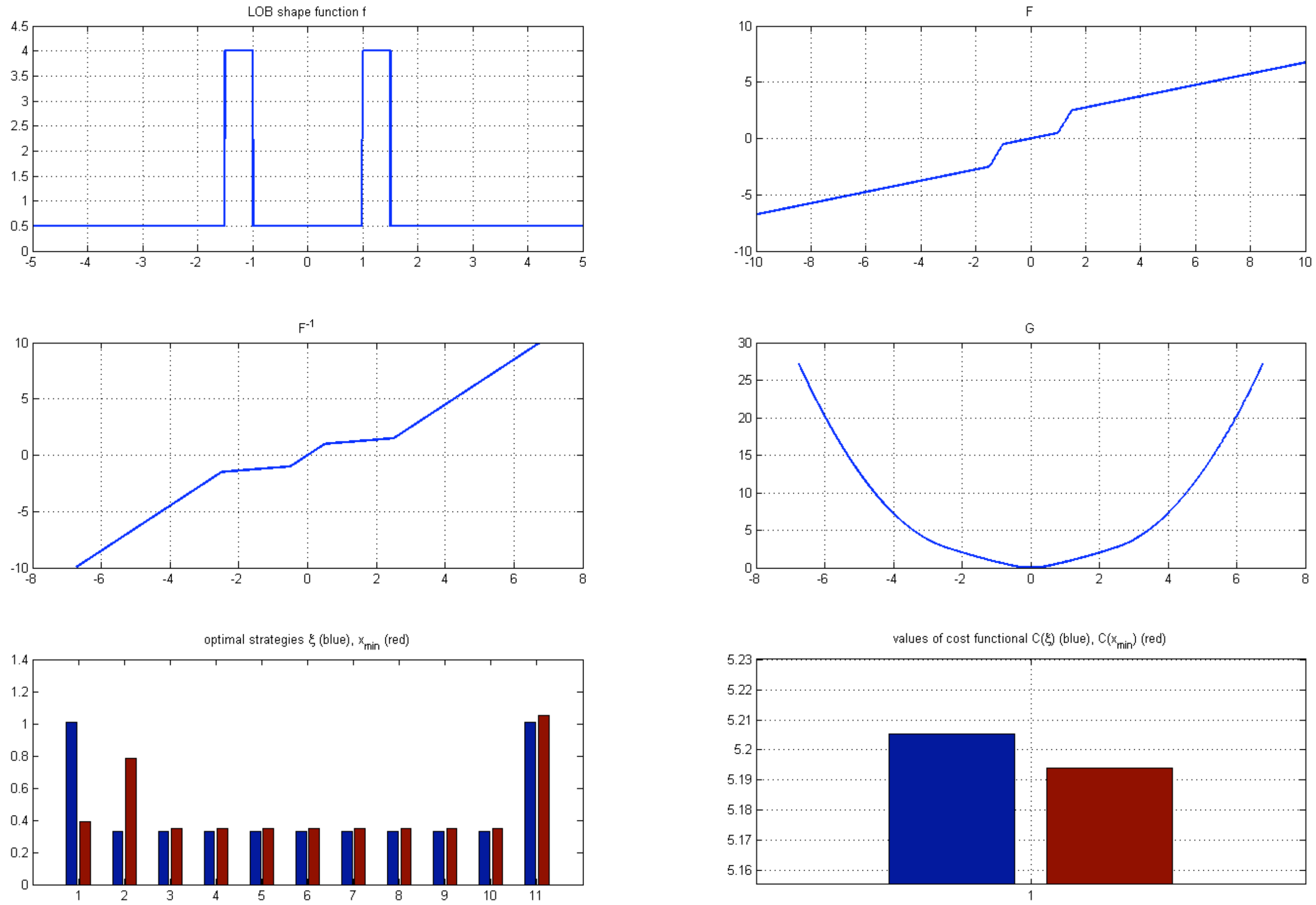


Figure 11: f piecewise constant

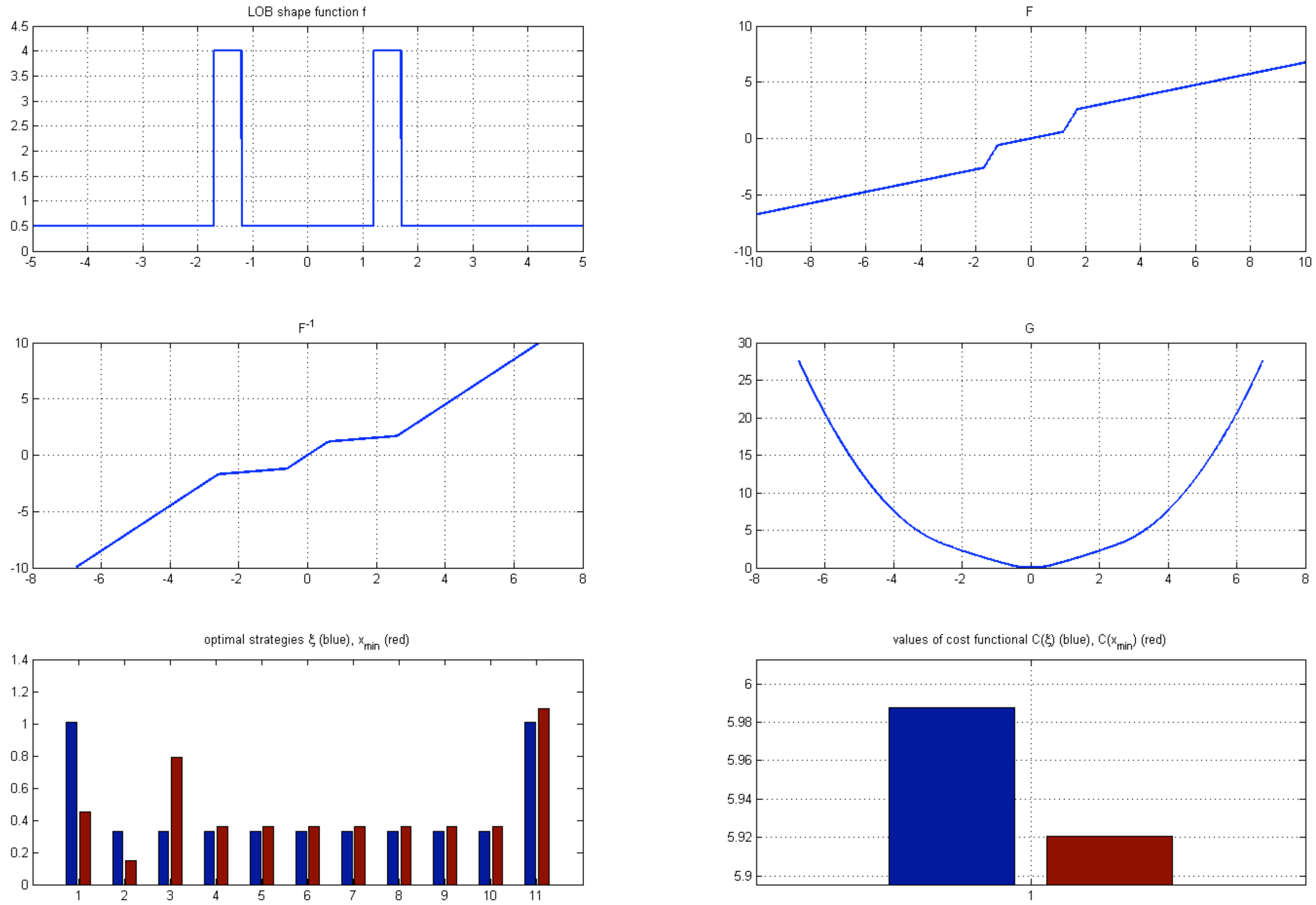


Figure 12: f piecewise constant

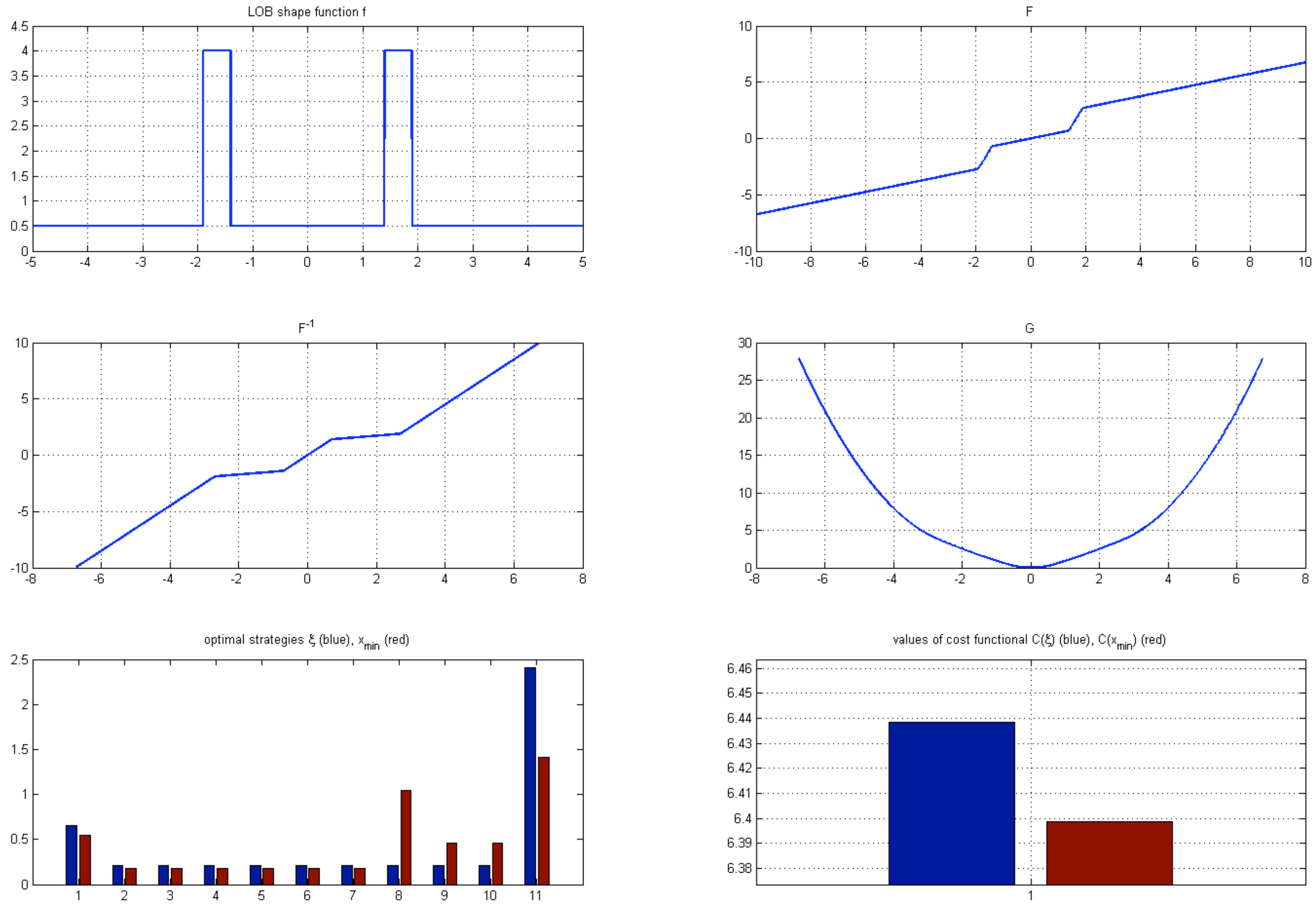


Figure 13: f piecewise constant

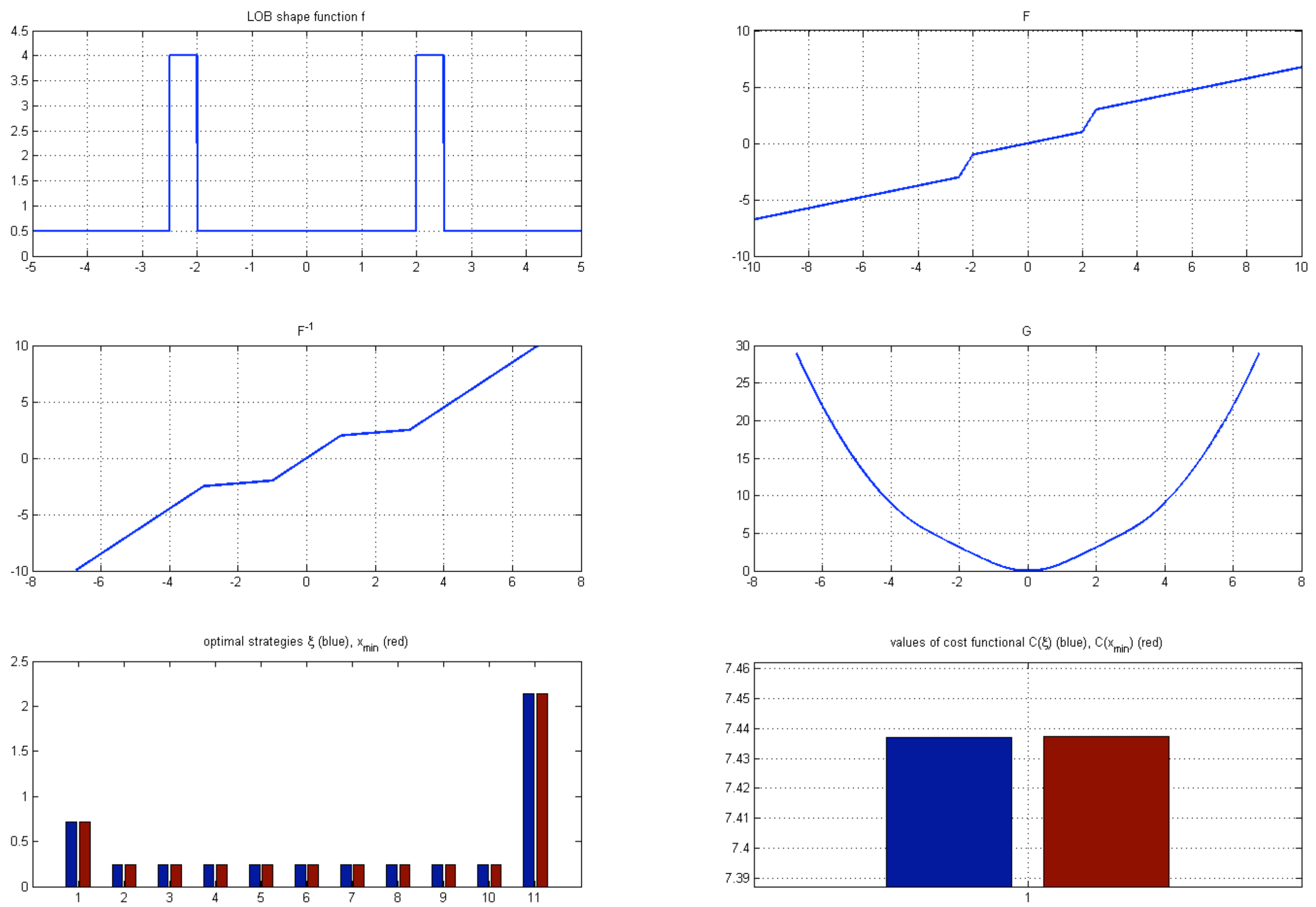


Figure 14: f piecewise constant

Continuous-time limit of the optimal strategy

- Initial block trade of size ξ_0^*, where

$$F^{-1}\Big(X_0 - \xi_0^* \int_0^T \rho_s \, ds\Big) = F^{-1}(\xi_0^*) + \frac{\xi_0^*}{f(F^{-1}(\xi_0^*))}$$

- Continuous trading in $]0, T[$ at rate

$$\xi_t^* = \rho_t \xi_0^*$$

- Terminal block trade of size

$$\xi_T^* = X_0 - \xi_0^* - \xi_0^* \int_0^T \rho_t \, dt$$

I. Order book models

1. Linear impact, general resilience

2. Nonlinear impact, exponential resilience

3. Gatheral's model

Liquidation time: $T \geq 0$.

Strategy: X adapted with $X_0 > 0$ fixed and $X_T = 0$.
Admissible: X_t bounded, absolutely continuous in t.

Liquidation time: $T \geq 0$.

Strategy: X adapted with $X_0 > 0$ fixed and $X_T = 0$.
Admissible: X_t bounded, absolutely continuous in t.

Market impact model: S^0 unaffected price, = martingale

$$S_t = S_t^0 + \int_0^t h(-\dot{X}_t)G(t-s)\,ds$$

- For $h(x) = \lambda x$ continuous-time version of simplified model in I.1.
- For nonlinear h close to continuous-time version of simplified model in I.2.
- $G \equiv const$ corresponds to purely permanent impact
- $G(t-s) = \delta(t-s)$ corresponds to purely temporary impact
- **Almgren-Chriss model:** (studied in next lectures)

$$G(t-s) = \lambda\delta(t-s) + \gamma$$

Costs:

$\dot{X}_t\,dt$ shares are sold at price S_t $\Rightarrow$ infinitesimal costs $= -\dot{X}_t S_t\,dt$

$$\begin{aligned}\text{Total costs} &= -\int_0^T \dot{X}_t S_t\,dt \\ &= -\int_0^T \dot{X}_t S_t^0\,dt + \int_0^T\int_0^t (-\dot{X}_t)h(-\dot{X}_s)G(t-s)\,ds\,dt\end{aligned}$$

Letting $\xi_t := -\dot{X}_t$, we get

$$\text{Expected costs} = -X_0 S_0^0 + \mathbb{E}\Big[\int_0^T\int_0^t \xi_t h(\xi_s)G(t-s)\,ds\,dt\Big]$$

Remark: Model formulation is not complete since optimal strategies typically will not be absolutely continous (see continous-time limit in preceding section)

Are there price manipulation strategies?

Find $\xi \in L^2[0,T]$ such that

$$\int_0^T \int_0^t \xi_t h(\xi_s) G(t-s)\, ds\, dt < 0.$$

Theorem 8. [Gatheral (2008)]
Suppose that

$$G(t) = e^{-\rho t}$$

and market impact is not linear. Then the model admits price manipulation strategies in the strong sense.

Theorem 8. [Gatheral (2008)]
Suppose that

$$G(t) = e^{-\rho t}$$

and market impact is not linear. Then the model admits price manipulation strategies in the strong sense.

Very puzzling result in view of Corollary 1!

Theorem 8. [Gatheral (2008)]
Suppose that

$$G(t) = e^{-\rho t}$$

and market impact is not linear. Then the model admits price manipulation strategies in the strong sense.

Very puzzling result in view of Corollary 1!

Resolution of this paradox:

$$Costs_{\text{Gatheral}} = \int_0^T \int_0^t \xi_t h(\xi_s) G(t-s)\, ds\, dt$$

$$Costs_{\text{AFS}} = \int_0^T \xi_t F^{-1}\Big(\int_0^t \xi_s G(t-s)\, ds \Big)\, dt$$

Theorem 8. [Gatheral (2008)]
Suppose that

$$G(t) = e^{-\rho t}$$

and market impact is not linear. Then the model admits price manipulation strategies in the strong sense.

Taking $\rho \downarrow 0$ yields:

Corollary 2. [Huberman & Stanzl (2004)]
Suppose that market impact is permanent and nonlinear. Then the model admits price manipulation strategies in the strong sense.

Sketch of proof of Theorem 8: For simplicity assume

$$h(-x) = -h(x)$$

Consider a strategy of the form

$$\xi_t = v_1 \text{ for } 0 \le t \le T_0 \text{ and } \xi_t = -v_2 \text{ for } T_0 < t \le T.$$

'Round trip' requires that

$$v_1 T_0 = v_2(T - T_0)$$

A calculation yields that for this specific strategy

$$\int_0^T \int_0^t \xi_t h(\xi_s) G(t-s)\, ds\, dt = \cdots$$

$$\begin{aligned}\cdots \quad = \quad & v_1 h(v_1)\Big(e^{-\frac{v_2\rho T}{v_1+v_2}} - 1 + \frac{v_2\rho T}{v_1+v_2}\Big) + v_2 h(v_2)\Big(e^{-\frac{v_1\rho T}{v_1+v_2}} - 1 + \frac{v_1\rho T}{v_1+v_2}\Big) \\ & -v_2 h(v_1)\Big(1 + e^{-\rho T} - e^{-\frac{v_2\rho T}{v_1+v_2}} - e^{-\frac{v_1\rho T}{v_1+v_2}}\Big)\end{aligned}$$

$$\begin{aligned}\cdots &= v_1 h(v_1)\Big(e^{-\frac{v_2\rho T}{v_1+v_2}} - 1 + \frac{v_2\rho T}{v_1+v_2}\Big) + v_2 h(v_2)\Big(e^{-\frac{v_1\rho T}{v_1+v_2}} - 1 + \frac{v_1\rho T}{v_1+v_2}\Big)\\ &\qquad -v_2 h(v_1)\Big(1 + e^{-\rho T} - e^{-\frac{v_2\rho T}{v_1+v_2}} - e^{-\frac{v_1\rho T}{v_1+v_2}}\Big)\\ &\approx \frac{v_1 v_2\big[v_1 h(v_2) - v_2 h(v_1)\big](\rho T)^2}{2(v_1+v_2)^2} + O((\rho T)^3) \qquad \text{for } \rho T \to 0\end{aligned}$$

$$\begin{aligned}\cdots &= v_1h(v_1)\Big(e^{-\frac{v_2\rho T}{v_1+v_2}}-1+\frac{v_2\rho T}{v_1+v_2}\Big)+v_2h(v_2)\Big(e^{-\frac{v_1\rho T}{v_1+v_2}}-1+\frac{v_1\rho T}{v_1+v_2}\Big)\\ &\qquad -v_2h(v_1)\Big(1+e^{-\rho T}-e^{-\frac{v_2\rho T}{v_1+v_2}}-e^{-\frac{v_1\rho T}{v_1+v_2}}\Big)\\ &\approx \frac{v_1v_2\big[v_1h(v_2)-v_2h(v_1)\big](\rho T)^2}{2(v_1+v_2)^2}+O((\rho T)^3) \qquad \text{for } \rho T\to 0\end{aligned}$$

Can always choose v_1, v_2 such that $[\dots]<0$, then take T such that ρT small enough. $\square$

More econo-physics:

$G(t) = t^{-\gamma}$, $h(v) = v^{\delta}$

Gatheral finds that

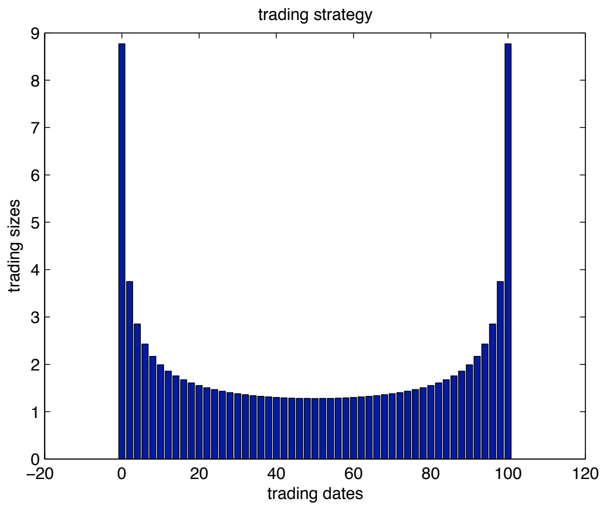
$$\gamma \text{ must be such that } \gamma \geq \gamma^* := 2 - \frac{\log 3}{\log 2} \approx 0.415$$

$$\delta + \gamma \approx 1$$

Consistent with (some) empirical studies.

Conclusion for Part I:

- Market impact should decay as a convex function of time
- Exponential or power law resilience leads to "bathtub solutions"



which are extremely robust

- Many open problems
- Minimizing *expected* costs does not take into account volatility risk. Must introduce risk aversion — see next part.

II. The qualitative effects of risk aversion

1. **Exponential utility and mean-variance**

2. **General utility functions**

3. **Mean-variance optimization for model from model from Section I.1**

References

Almgren, R. *Optimal execution with nonlinear impact functions and trading-enhanced risk.* Applied Mathematical Finance 10 , 1-18 (2003).

Almgren, R., Chriss, N. *Value under liquidation.* Risk, Dec. 1999.

Almgren, R., Chriss, N. *Optimal execution of portfolio transactions.* J. Risk 3, 5-39 (2000).

Almgren, R., Lorenz, J. *Adaptive arrival price.* In: Algorithmic Trading III: Precision, Control, Execution, Brian R. Bruce, editor, Institutional Investor Journals (2007).

A.S., Schöneborn, T. *Optimal basket liquidation with finite time horizon for CARA investors.* Preprint (2008).

A.S., Schöneborn, T. *Risk aversion and the dynamics of optimal liquidation strategies in illiquid markets.* Finance Stoch. **13**, (2009).

A.S., Schöneborn, T., Tehranci, M. *Optimal basket liquidation for CARA investors is deterministic.* To appear in Applied Math. Finance.

Schöneborn, T. *Trade execution in illiquid markets. Optimal stochastic control and multi-agent equilibria.* Ph.D. thesis, TU Berlin (2008).

II. The qualitative effects of risk aversion

1. Exponential utility and mean-variance

Liquidation time: $T \in [0, \infty]$.

Strategy: X adapted with $X_0 > 0$ fixed and $X_T = 0$.

Admissible: X_t bounded, absolutely continuous in t. Take

$$\xi_t := -\dot{X}_t$$

as controll. Then

$$X_t^\xi := X_0 - \int_0^t \xi_s \, ds$$

Market impact model: Following Almgren (2003),

$S_t^\xi =$	S_0	$+$	σB_t	$+$	$\gamma(X_t^\xi - X_0)$	$+$	$h(\xi_t)$
	initial price		Brownian motion		permanent impact		temporary impact

Most common model in practice; *drift, multiple assets, general Lévy process, Gatheral-type impact* possible.

Assumption:

$$f(x) := xh(x)$$

is convex, C^1, and satisfies $f(x) = f(-x)$ and $f(x)/x \to \infty$ for $|x| \to \infty$.
E.g., $h(x) = \alpha \operatorname{sign}(x)\sqrt{|x|} + \beta x$.

Sales revenues:

$$\begin{aligned}\mathcal{R}_T(\xi) &= \int_0^T (-\dot{X}_t) S_t^\xi \, dt = \dots \\ &= S_0 X_0 - \frac{\gamma}{2} X_0^2 + \sigma \int_0^T X_t^\xi \, dB_t - \int_0^T f(\xi_t)\, dt.\end{aligned}$$

Goal: maximize expected utility

$$\mathbb{E}[\, u(\mathcal{R}_T(\xi))\,]$$

over admissible strategies for $u(x) = -e^{-\alpha x}$

Setup as control problem

- controlled diffusion:

$$R_t^\xi = R_0 + \sigma \int_0^t X_s^\xi \, dB_s - \int_0^t f(\xi_s) \, ds$$

- value function

$$v(T, X_0, R_0) = \sup_{\xi \in \mathcal{X}(T, X_0)} \mathbb{E}\big[\, u(R_T^\xi) \,\big],$$

where

$$\mathcal{X}(T, X_0) = \Big\{ \xi \,|\, X^\xi \text{ is bounded and } \int_0^T \xi_t \, dt = X_0 \Big\}$$

Heuristic derivation of HJB equation

$$dv(T-t, X_t^\xi, R_t^\xi) = \sigma v_R X_t^\xi \, dB_t + \Big(- v_t - \xi_t v_X - v_R f(\xi_t) + \frac{\sigma^2}{2}(X_t^\xi)^2 v_{RR} \Big)\, dt$$

Hence

$$v_t = \frac{\sigma^2}{2} X^2 v_{RR} - \inf_\xi \big(\xi v_X + v_R f(\xi) \big)$$

Heuristic derivation of HJB equation

$$dv(T-t, X_t^\xi, R_t^\xi) = \sigma v_R X_t^\xi \, dB_t + \Big(- v_t - \xi_t v_X - v_R f(\xi_t) + \frac{\sigma^2}{2}(X_t^\xi)^2 v_{RR} \Big)\, dt$$

Hence

$$v_t = \frac{\sigma^2}{2} X^2 v_{RR} - \inf_\xi \big(\xi v_X + v_R f(\xi) \big)$$

What about the constraint $\int_0^T \xi_t \, dt = X_0$?

Heuristic derivation of HJB equation

$$dv(T-t, X_t^\xi, R_t^\xi) = \sigma v_R X_t^\xi \, dB_t + \Big(- v_t - \xi_t v_X - v_R f(\xi_t) + \frac{\sigma^2}{2}(X_t^\xi)^2 v_{RR}\Big)\, dt$$

Hence

$$v_t = \frac{\sigma^2}{2} X^2 v_{RR} - \inf_\xi \big(\xi v_X + v_R f(\xi)\big)$$

What about the constraint $\int_0^T \xi_t \, dt = X_0$? It is in the initial condition:

$$v(0, X, R) = \lim_{T \downarrow 0} v(T, X, R) = \begin{cases} u(R) & \text{if } X = 0, \\ -\infty & \text{otherwise.} \end{cases}$$

Heuristic derivation of HJB equation

$$dv(T-t, X_t^\xi, R_t^\xi) = \sigma v_R X_t^\xi \, dB_t$$
$$+ \Big(-v_t - \xi_t v_X - v_R f(\xi_t) + \frac{\sigma^2}{2}(X_t^\xi)^2 v_{RR} \Big)\, dt$$

Hence

$$v_t = \frac{\sigma^2}{2} X^2 v_{RR} - \inf_\xi \big(\xi v_X + v_R f(\xi)\big)$$

What about the constraint $\int_0^T \xi_t \, dt = X_0$? It is in the initial condition:

$$v(0, X, R) = \lim_{T \downarrow 0} v(T, X, R) = \begin{cases} u(R) & \text{if } X = 0, \\ -\infty & \text{otherwise.} \end{cases}$$

Theorem 9. [A.S. & Schöneborn (2008), A.S., Schöneborn & Tehranchi (2009)]
If $u(x) = -e^{-\alpha x}$ *for some* $\alpha > 0$*, then the unique optimal strategy* ξ^* *is a deterministic function of* t*. Moreover,* v *is a classical solution of the singular HJB equation.*

The fact that optimal strategies for CARA investors are deterministic is very robust. Is also true

- if Brownian motion is replaced by a Lévy process;
- for Gatheral-type impact
- other models with functionally dependent impact

Sketch of proof: For simplicity: $\sigma = 1$. We have

$$\begin{aligned}\mathbb{E}\big[\,u(R_T^\xi)\,\big] &= -e^{-\alpha R_0}\mathbb{E}\Big[\,e^{-\alpha\int_0^T X_t^\xi\,dB_t+\alpha\int_0^T f(\xi_t)\,dt}\,\Big]\\ &= -e^{-\alpha R_0}\mathbb{E}^\xi\Big[\,e^{\frac{\alpha^2}{2}\int_0^T (X_t^\xi)^2\,dt+\alpha\int_0^T f(\xi_t)\,dt}\,\Big]\end{aligned}$$

where

$$\frac{d\mathbb{P}^\xi}{d\mathbb{P}} = e^{-\alpha\int_0^T X_t^\xi\,dB_t-\frac{\alpha^2}{2}\int_0^T (X_t^\xi)^2\,dt}$$

Sketch of proof: For simplicity: $\sigma = 1$. We have

$$\begin{aligned}\mathbb{E}\big[\,u(R_T^\xi)\,\big] &= -e^{-\alpha R_0}\mathbb{E}\Big[e^{-\alpha\int_0^T X_t^\xi\,dB_t+\alpha\int_0^T f(\xi_t)\,dt}\Big]\\ &= -e^{-\alpha R_0}\mathbb{E}^\xi\Big[e^{\frac{\alpha^2}{2}\int_0^T (X_t^\xi)^2\,dt+\alpha\int_0^T f(\xi_t)\,dt}\Big]\end{aligned}$$

where

$$\frac{d\mathbb{P}^\xi}{d\mathbb{P}} = e^{-\alpha\int_0^T X_t^\xi\,dB_t-\frac{\alpha^2}{2}\int_0^T (X_t^\xi)^2\,dt}$$

Now we can minimize inside the expectation w.r.t. $\mathbb{P}^\xi$:

$$\begin{aligned}\mathbb{E}^\xi\Big[e^{\frac{\alpha^2}{2}\int_0^T (X_t^\xi)^2\,dt+\alpha\int_0^T f(\xi_t)\,dt}\Big] &\geq \mathbb{E}^\xi\Big[e^{\frac{\alpha^2}{2}\int_0^T (X_t^{\xi^*})^2\,dt+\alpha\int_0^T f(\xi_t^*)\,dt}\Big]\\ &= e^{\frac{\alpha^2}{2}\int_0^T (X_t^{\xi^*})^2\,dt+\alpha\int_0^T f(\xi_t^*)\,dt}\end{aligned}$$

where ξ^* is the deterministic minimizer of

$$\xi \longmapsto \frac{\alpha}{2}\int_0^T (X_t^\xi)^2\,dt + \int_0^T f(\xi_t)\,dt. \qquad \square$$

Hence, the value function is

$$
\begin{aligned}
v(T, X_0, R_0) &= \sup_{\xi \in \mathcal{X}(T,X_0)} \mathbb{E}\big[\, u(R_T^\xi)\,\big] = \sup_{\xi \in \mathcal{X}_{\det}(T,X_0)} \mathbb{E}\big[\, u(R_T^\xi)\,\big] \\
&= -\exp\Big(-\alpha R_0 + \alpha \inf_{\xi \in \mathcal{X}_{\det}(T,X_0)} \int_0^T L(X_t^\xi, \xi_t)\, dt \Big)
\end{aligned}
$$

where $\mathcal{X}_{\det}(T, X_0)$ are the deterministic strategies in $\mathcal{X}(T, X_0)$ and L is the Lagrangian

$$
L(q,p) = \frac{\alpha}{2} q^2 + f(-p) = \frac{\alpha}{2} q^2 + f(p)
$$

Classical mechanics: the action function

$$S(T,X) := \inf_{\xi\in\mathcal{X}_{\det}(T,X)} \int_0^T L(X_t^\xi,\xi_t)\,dt = \inf_{\xi\in\mathcal{X}_{\det}(T,X)} \int_0^T L(X_t^\xi,\dot X_t^\xi)\,dt$$

is a classical solution of the Hamilton-Jacobi equation

$$S_T(T,X) + H(X,S_X(T,X)) = 0 \qquad T>0,\ X\in\mathbb{R}$$

where H is the Hamiltonian

$$H(q,p) = -\frac{\alpha}{2}q^2 + f^*(p)$$

Boundary conditions:

$$S(0,0)=0 \qquad \text{and} \qquad S(0,X)=\infty \text{ for } X\neq 0.$$

[Side remark: this fact is classical when $f\in C^2$ but more subtle when $f\in C^1$ as for $h(x)=\sqrt{|x|}$]

Plugging the Hamilton-Jacobi equation into

$$\begin{aligned} v(T, X_0, R_0) &= -\exp\Big(-\alpha R_0 + \alpha \inf_{\xi \in \mathcal{X}_{\det}(T,X_0)} \int_0^T L(X_t^\xi, \xi_t)\,dt\Big) \\ &= -\exp\Big(-\alpha R_0 + \alpha S(T, X_0)\Big) \end{aligned}$$

yields the singular HJB-equation for v. $\square$

Alternative proof: Define the function

$$w(T, X_0, R_0) := -\exp\Big(-\alpha R_0 + \alpha S(T, X_0)\Big)$$

so that it's a classical solution of the singular HJB-equation. Then use a verification argument to show that $w = v$ (subtle).

Then there is $\xi^* \in \mathcal{X}_{\text{det}}(T, X_0)$ such that

$$S(T, X_0) = \int_0^T L(X_t^{\xi^*}, \xi_t^*)\, dt$$

and this ξ^* must hence be optimal. □

The relation with mean-variance optimization

For $\xi \in \mathcal{X}_{\text{det}}(T, X_0)$,

$$R_t^\xi = R_0 + \sigma \int_0^t X_s^\xi \, dB_s - \int_0^t f(\xi_s)\, ds$$

is Gaussian, and so

$$\mathbb{E}[\, u(R_T^\xi)\,] = -\exp\Big(-\alpha \mathbb{E}[\, R_T^\xi\,] + \frac{\alpha^2}{2} \text{var}(R_T^\xi)\Big)$$

Hence, exponential-utility maximization is equivalent to the maximization of the mean-variance functional

$$\mathbb{E}[\, R_T^\xi\,] - \frac{\alpha}{2} \text{var}(R_T^\xi)$$

for deterministic strategies [Markowitz,. . . , Almgren & Chriss (2000)]. Different for adaptive strategies [Almgren & Lorenz (2008)].

Computation of the optimal strategy

Classical mechanics: X^{ξ^*} is solution of the Euler-Lagrange equation

$$\alpha X = f''(\dot{X}_t)\ddot{X}_t \qquad \text{with } X_0 = \textit{ initial portfolio } \text{ and } X_T = 0$$

Computation of the optimal strategy

Classical mechanics: X^{ξ^*} is solution of the Euler-Lagrange equation

$$\alpha X = f''(\dot{X}_t)\ddot{X}_t \qquad \text{with } X_0 = \textit{ initial portfolio } \text{ and } X_T = 0$$

Not clear when $f \notin C^2$ as for $h(x) = \sqrt{|x|}$

Theorem 10. **[A.S. & Schöneborn (2008)]**
The optimal X^{ξ^*} is C^1 and uniquely solves the Hamilton equations

$$\begin{aligned} \dot{X}_t &= H_p(X_t, p(t)) = -(f^*)'(-p(t)) \\ \dot{p}(t) &= -H_q(X_t, p(t)) = \alpha X_t \end{aligned}$$

with initial conditions $X_0^{\xi^*} = X_0$ and $p(0) = -(f^*)'(\xi_0^*)$.

Example: For linear temporary impact, $f(x) = \lambda x^2$, the optimal strategy is

$$\xi_t^* = X_0 \sqrt{\frac{\alpha\sigma^2}{2\lambda}} \cdot \frac{\cosh\left((T-t)\sqrt{\frac{\alpha\sigma^2}{2\lambda}}\right)}{\sinh\left(T\sqrt{\frac{\alpha\sigma^2}{2\lambda}}\right)}$$

$$X_t^{\xi^*} = X_0 \cdot \frac{\cosh\left(t\sqrt{\frac{\alpha\sigma^2}{2\lambda}}\right)\sinh\left(T\sqrt{\frac{\alpha\sigma^2}{2\lambda}}\right) - \cosh\left(T\sqrt{\frac{\alpha\sigma^2}{2\lambda}}\right)\sinh\left(t\sqrt{\frac{\alpha\sigma^2}{2\lambda}}\right)}{\sinh\left(T\sqrt{\frac{\alpha\sigma^2}{2\lambda}}\right)}$$

The value function is

$$v(T, R_0, X_0) = -\exp\left[-\alpha(R_0 + S_0 X_0 - \frac{\gamma}{2}X_0^2) + X_0^2\sqrt{\frac{\lambda\alpha^3\sigma^2}{2}}\coth\left(T\sqrt{\frac{\alpha\sigma^2}{2\lambda}}\right)\right]$$

II. The qualitative effects of risk aversion

1. Exponential utility and mean-variance

2. General utility functions

Problem with $T < \infty$ difficult because of singular initial condition of HJB equation.

$\Longrightarrow$ Consider infinite time horizon instead

- Assume also linear temporary impact (for simplicity only)

$$f(x) = \lambda x^2$$

- Utility function $u \in C^6(\mathbb{R})$ such that the absolute risk aversion,

$$A(R) := -\frac{u''(R)}{u'(R)} \qquad (= \text{constant for exponential utility}),$$

satisfies

$$0 < A_{min} \leq A(R) \leq A_{max} < \infty.$$

Entire section based on A.S. & Schöneborn (2009)

Recall

$$R_t^\xi = R_0 + \sigma \int_0^t X_s^\xi \, dB_s - \lambda \int_0^t \xi_s^2 \, ds.$$

- Optimal liquidation:

$$\text{maximize} \quad \mathbb{E}[\, u(R_\infty^\xi)\,]$$

- Maximization of asymptotic portfolio value:

$$\text{maximize} \quad \lim_{t\uparrow\infty} \mathbb{E}[\, u(R_t^\xi)\,]$$

Note: Liquidation enforced by the fact that a risk-averse investor does not want to hold a stock whose price process is a martingale.

HJB equation for finite time horizon:

$$v_t = \frac{\sigma^2}{2} X^2 v_{RR} - \inf_c \left(c v_X + \lambda v_R c^2 \right)$$

Guess for infinite time horizon:

$$0 = \frac{\sigma^2}{2} X^2 v_{RR} - \inf_c \left(c v_X + \lambda v_R c^2 \right)$$

Initial condition:

$$v(0, R) = u(R).$$

HJB equation for finite time horizon:

$$v_t = \frac{\sigma^2}{2} X^2 v_{RR} - \inf_c \left(c v_X + \lambda v_R c^2 \right)$$

Guess for infinite time horizon:

$$0 = \frac{\sigma^2}{2} X^2 v_{RR} - \inf_c \left(c v_X + \lambda v_R c^2 \right)$$

Initial condition:

$$v(0, R) = u(R).$$

Corresponding reduced-form equation:

$$v_X^2 = -2\lambda \sigma^2 X^2 v_R \cdot v_{RR}$$

Not a straightforward PDE either......

Way out: consider optimal Markov control in HJB equation

$$\widehat{c}(X,R) = -\frac{v_X(X,R)}{2\lambda v_R(X,R)}$$

and let

$$\widetilde{c}(Y,R) = \frac{\widehat{c}(\sqrt{Y},R)}{\sqrt{Y}}.$$

If v solves the HJB equation, then $\widetilde{c}$ solves

$$(*) \qquad \begin{cases} \widetilde{c}_Y = \dfrac{\sigma^2}{4\widetilde{c}}\widetilde{c}_{RR} - \dfrac{3}{2}\lambda\widetilde{c}\widetilde{c}_R \\[2em] \widetilde{c}(0,R) = \sqrt{\frac{\sigma^2 A(R)}{2\lambda}} \end{cases}$$

Way out: consider optimal Markov control in HJB equation

$$\widehat{c}(X,R) = -\frac{v_X(X,R)}{2\lambda v_R(X,R)}$$

and let

$$\widetilde{c}(Y,R) = \frac{\widehat{c}(\sqrt{Y},R)}{\sqrt{Y}}.$$

If v solves the HJB equation, then $\widetilde{c}$ solves

$$(*) \qquad \begin{cases} \widetilde{c}_Y = \dfrac{\sigma^2}{4\widetilde{c}}\widetilde{c}_{RR} - \dfrac{3}{2}\lambda\widetilde{c}\widetilde{c}_R \\[2ex] \widetilde{c}(0,R) = \sqrt{\frac{\sigma^2 A(R)}{2\lambda}} \end{cases}$$

Theorem 11. $(*)$ *admits a unique classical solution* $\widetilde{c} \in C^{2,4}$ *s.th.*

$$\sqrt{\frac{\sigma^2 A_{min}}{2\lambda}} \leq \widetilde{c}(Y,R) \leq \sqrt{\frac{\sigma^2 A_{max}}{2\lambda}}$$

Follows from:

Theorem 12. [Ladyzhenskaya, Solonnikov & Uraltseva (1968)] *There is a classical $C^{2,4}$-solution for the parabolic partial differential equation*

$$f_t - \frac{\partial}{\partial x}\big[a(x,t,f,f_x)\big] + b(x,t,f,f_x) = 0$$

with initial value condition $f(0,x) = \psi_0(x)$ if all of the following conditions are satisfied:

- *$\psi_0(x)$ is smooth (C^4) and bounded*
- *a and b are smooth (C^3 respectively C^2)*
- *There are constants b_1 and $b_2 \geq 0$ such that for all x and u:*

$$\bigg(b(x,t,u,0) - \frac{\partial a}{\partial x}(x,t,u,0)\bigg)\, u \geq -b_1 u^2 - b_2.$$

- *For all $M > 0$, there are constants $\mu_M \geq \nu_M > 0$ such that for all x, t, u and p that are bounded in modulus by M:*

$$\nu_M \leq \frac{\partial a}{\partial p}(x,t,u,p) \leq \mu_M \tag{12}$$

and

$$\left(|a| + \left|\frac{\partial a}{\partial u}\right|\right)(1+|p|) + \left|\frac{\partial a}{\partial x}\right| + |b| \leq \mu_M(1+|p|)^2. \tag{13}$$

Proof: Obtained from original existence theorem by cutting off the coefficients of the PDE. □

Next, consider the transport equation

$$\begin{cases} \widetilde{w}_Y = -\lambda \widetilde{c}\widetilde{w}_R \\ \widetilde{w}(0,R) = u(R). \end{cases}$$

Proposition 5. *The transport equation admits a $C^{2,4}$-solution $\widetilde{w}$. Moreover, $w(X,R) := \widetilde{w}(X^2,R)$ is a classical solution of the HJB equation*

$$0 = \frac{\sigma^2}{2} X^2 w_{RR} - \inf_c \big(c w_X + w_R c^2 \big), \qquad w(0,R) = u(R)$$

The unique minimum above is attained at

$$c(X,R) := \widetilde{c}(X^2,R)X.$$

Sketch of proof: Existence and uniqueness of solutions follws by method of characteristics. Assume for the moment that

$$\widetilde{c}^2 = -\frac{\sigma^2 \widetilde{w}_{RR}}{2\lambda \widetilde{w}_R}.$$

Then with $Y = X^2$:

$$\begin{aligned}
0 &= -\lambda X^2 \widetilde{w}_R \left(\frac{\sigma^2 \widetilde{w}_{RR}}{2\lambda \widetilde{w}_R} + \widetilde{c}^2 \right) \\
&= -\lambda X^2 \widetilde{w}_R \left(\frac{\sigma^2 \widetilde{w}_{RR}}{2\lambda \widetilde{w}_R} + \frac{\widetilde{w}_Y^2}{\lambda^2 \widetilde{w}_R^2} \right) \\
&= -\frac{1}{2}\sigma^2 X^2 w_{RR} - \frac{w_X^2}{4\lambda w_R} \\
&= \inf_c \left[-\frac{1}{2}\sigma^2 X^2 w_{RR} + \lambda w_R c^2 + w_X c \right]
\end{aligned}$$

We now show that

$$\widetilde{c}^2 = -\frac{\sigma^2 \widetilde{w}_{RR}}{2\lambda \widetilde{w}_R}.$$

First, observe that it holds for $Y = 0$. For general Y, consider

$$\frac{d}{dY}\widetilde{c}^2 = -3\lambda \widetilde{c}^2 \widetilde{c}_R + \frac{\sigma^2}{2}\widetilde{c}_{RR}$$

$$-\frac{d}{dY}\frac{\sigma^2 \widetilde{w}_{RR}}{2\lambda \widetilde{w}_R} = \sigma^2 \widetilde{c}\frac{d}{dR}\frac{\widetilde{w}_{RR}}{2\widetilde{w}_R} + \sigma^2 \widetilde{c}_R \frac{\widetilde{w}_{RR}}{2\widetilde{w}_R} + \frac{\sigma^2}{2}\widetilde{c}_{RR}$$

The first holds by PDE for $\widetilde{c}$, the second by transport eqn. for $\widetilde{w}$.

Next,

$$\frac{d}{dY}\left(\widetilde{c}^2 + \frac{\sigma^2 \widetilde{w}_{RR}}{2\lambda \widetilde{w}_R}\right) = -3\lambda \widetilde{c}^2 \widetilde{c}_R + \frac{\sigma^2}{2}\widetilde{c}_{RR} - \sigma^2 \widetilde{c}\frac{d}{dR}\frac{\widetilde{w}_{RR}}{2\widetilde{w}_R} - \sigma^2 \widetilde{c}_R \frac{\widetilde{w}_{RR}}{2\widetilde{w}_R} - \frac{\sigma^2}{2}\widetilde{c}_{RR}$$

$$= -\lambda \widetilde{c}\frac{d}{dR}\left(\widetilde{c}^2 + \frac{\sigma^2 \widetilde{w}_{RR}}{2\lambda \widetilde{w}_R}\right) - \lambda \widetilde{c}_R \left(\widetilde{c}^2 + \frac{\sigma^2 \widetilde{w}_{RR}}{2\lambda \widetilde{w}_R}\right).$$

We now show that

$$\widetilde{c}^2 = -\frac{\sigma^2 \widetilde{w}_{RR}}{2\lambda \widetilde{w}_R}.$$

First, observe that it holds for $Y = 0$. For general Y, consider

$$\frac{d}{dY}\widetilde{c}^2 = -3\lambda\widetilde{c}^2\widetilde{c}_R + \frac{\sigma^2}{2}\widetilde{c}_{RR}$$

$$-\frac{d}{dY}\frac{\sigma^2 \widetilde{w}_{RR}}{2\lambda \widetilde{w}_R} = \sigma^2\widetilde{c}\frac{d}{dR}\frac{\widetilde{w}_{RR}}{2\widetilde{w}_R} + \sigma^2\widetilde{c}_R\frac{\widetilde{w}_{RR}}{2\widetilde{w}_R} + \frac{\sigma^2}{2}\widetilde{c}_{RR}$$

The first holds by PDE for $\widetilde{c}$, the second by transport eqn. for $\widetilde{w}$.

Next,

$$\frac{d}{dY}\left(\widetilde{c}^2 + \frac{\sigma^2 \widetilde{w}_{RR}}{2\lambda \widetilde{w}_R}\right) = -3\lambda\widetilde{c}^2\widetilde{c}_R + \frac{\sigma^2}{2}\widetilde{c}_{RR} - \sigma^2\widetilde{c}\frac{d}{dR}\frac{\widetilde{w}_{RR}}{2\widetilde{w}_R} - \sigma^2\widetilde{c}_R\frac{\widetilde{w}_{RR}}{2\widetilde{w}_R} - \frac{\sigma^2}{2}\widetilde{c}_{RR}$$
$$= -\lambda\widetilde{c}\frac{d}{dR}\left(\widetilde{c}^2 + \frac{\sigma^2 \widetilde{w}_{RR}}{2\lambda \widetilde{w}_R}\right) - \lambda\widetilde{c}_R\left(\widetilde{c}^2 + \frac{\sigma^2 \widetilde{w}_{RR}}{2\lambda \widetilde{w}_R}\right).$$

Therefore need $u \in C^6$!

Hence,

$$f(Y,R) := \widetilde{c}^2 + \frac{\sigma^2 \widetilde{w}_{RR}}{2\lambda \widetilde{w}_R}$$

satisfies the linear PDE

$$f_Y = -\lambda \widetilde{c} f_R - \lambda \widetilde{c}_R f$$

with initial value condition $f(0,R) = 0$. One obvious solution to this PDE is $f(Y,R) \equiv 0$. By the method of characteristics this is the unique solution to the PDE, since $\widetilde{c}$ and $\widetilde{c}_R$ are smooth and hence locally Lipschitz. $\square$

A (rather technical) verification argument yields:

Theorem 13. *The value functions for optimal liquidation and for maximization of asymptotic portfolio value are equal and are classical solutions of the HJB equation*

$$-\frac{1}{2}\sigma^2 X^2 v_{RR} + \inf_c \big[\lambda v_R c^2 + v_X c\big] = 0$$

with boundary condition $v(0,R) = u(R)$*. The a.s. unique optimal control* $\hat{\xi}_t$ *is Markovian and given in feedback form by*

$$\hat{\xi}_t = c(X_t^{\hat{\xi}}, R_t^{\hat{\xi}}) = -\frac{v_X}{2\lambda v_R}(X_t^{\hat{\xi}}, R_t^{\hat{\xi}}). \tag{14}$$

For the value functions, we have convergence:

$$v(X_0, R_0) = \lim_{t\to\infty} \mathbb{E}[u(R_t^{\hat{\xi}})] = \mathbb{E}[u(R_\infty^{\hat{\xi}})] \tag{15}$$

Corollary 3. *If* $u(R) = -e^{-AR}$, *then*

$$X_t^{\xi^*} = X_0 \exp\Big(-t\sqrt{\frac{\sigma^2 A}{2\lambda}}\Big).$$

Corollary 3. *If* $u(R) = -e^{-AR}$*, then*

$$X_t^{\xi^*} = X_0 \exp\Big(-t\sqrt{\frac{\sigma^2 A}{2\lambda}}\Big).$$

General result:

Theorem 14. *The optimal strategy* $c(X,R)$ *is increasing (decreasing) in* R *iff* $A(R)$ *is increasing (decreasing). I.e.,*

Utility function		*Optimal trading strategy*
DARA	$\Longleftrightarrow$	*Passive in the money*
CARA	$\Longleftrightarrow$	*Neutral in the money*
IARA	$\Longleftrightarrow$	*Aggresive in the money*

Theorem 15. *If* u^1 *and* u^0 *are such that* $A^1 \geq A^0$ *then* $c^1 \geq c^0$*.*

Idea of Proof: $g := \widetilde{c}^1 - \widetilde{c}^0$ solves

$$g_Y = \frac{1}{2} a g_{RR} + b g_R + V g,$$

where

$$a = \frac{\sigma^2}{2\widetilde{c}^0}, \qquad b = -\frac{3}{2}\lambda \widetilde{c}^1, \qquad \text{and} \qquad V = -\frac{\sigma^2 \widetilde{c}^1_{RR}}{4\widetilde{c}^0 \widetilde{c}^1} - \frac{3}{2}\lambda \widetilde{c}^0_R.$$

The boundary condition of g is

$$g(0, R) = \sqrt{\frac{\sigma^2 A^1(R)}{2\lambda}} - \sqrt{\frac{\sigma^2 A^0(R)}{2\lambda}} \geq 0$$

Now maximum principle or Feynman-Kac argument....
(plus localization) □

Relation to forward utilities

Theorem 16.
For every $X > 0$, the value function $v(X,R)$ is again a utility function in R. Moreover,

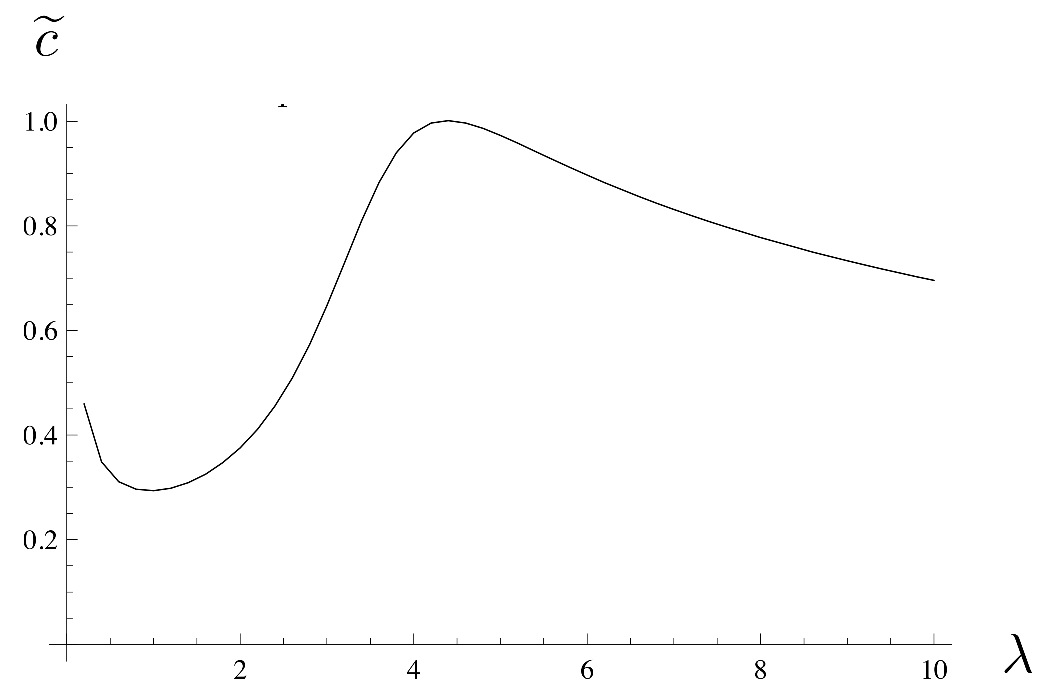
$$\widetilde{c}(Y,R) = \sqrt{\frac{\sigma^2 A(\sqrt{Y},R)}{2\lambda}}. \tag{16}$$

where

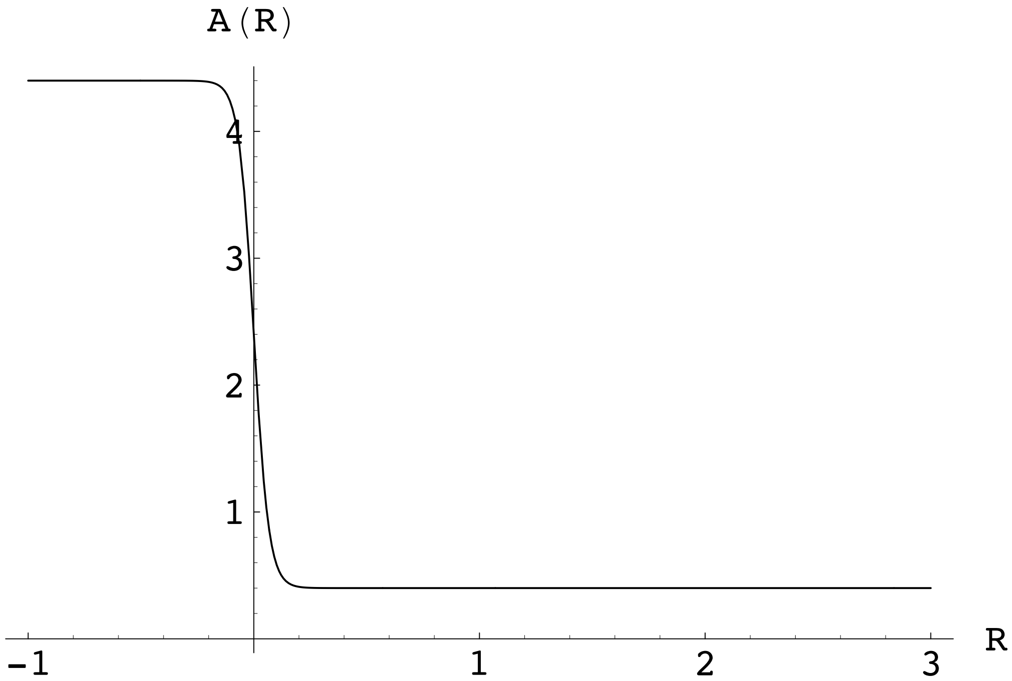
$$A(X,R) := -\frac{v_{RR}(X,R)}{v_R(X,R)}$$

What about other monotonicity relations?

• Monotonicity in λ: intuitively, an increase in liquidaton costs should lead to a decrease of liquidation speed.

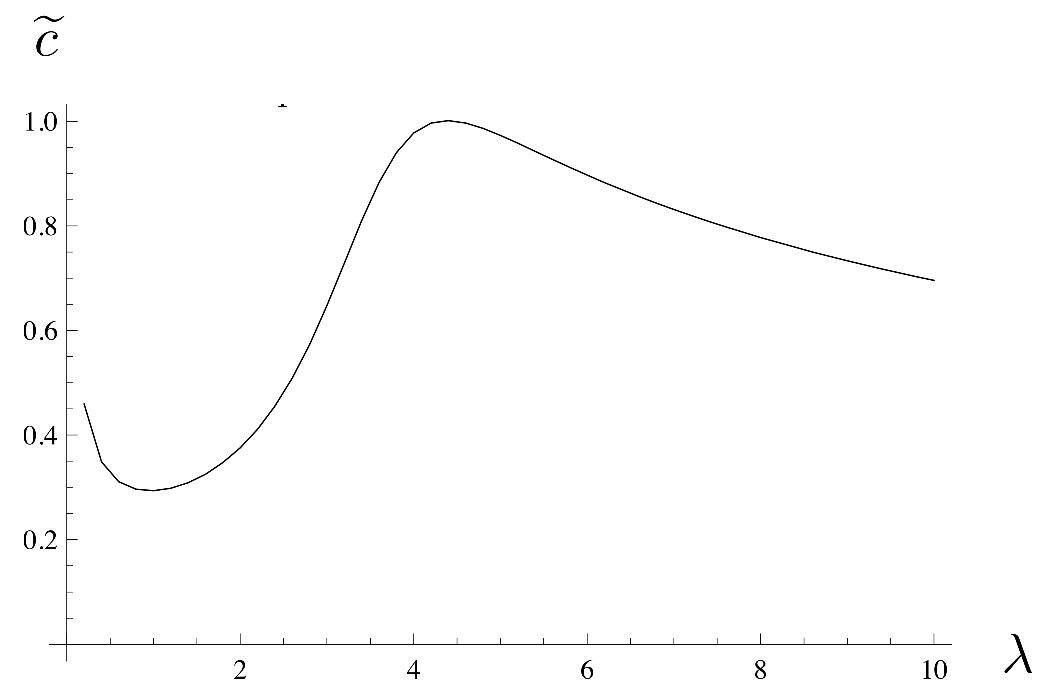


Dependence of the transformed optimal strategy $\widetilde{c}$ on λ for the DARA utility function with $A(R) = 2(1.2 - \tanh(15R))^2$.



The shape of the absolute risk aversion

$$A(R) = 2(1.2 - \tanh(15R))^2$$



Dependence of the transformed optimal strategy $\widetilde{c}$ on λ for the DARA utility function with $A(R) = 2(1.2 - \tanh(15R))^2$.

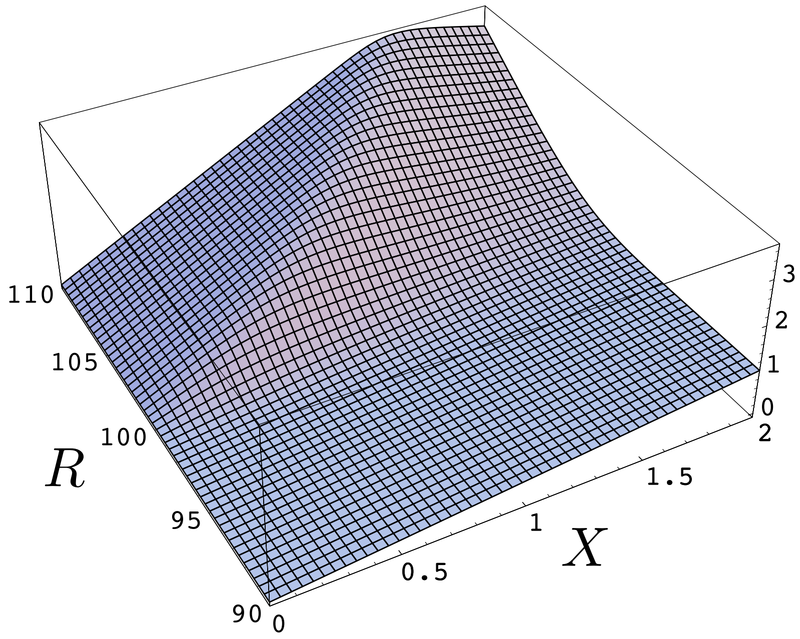
Theorem 17. *IARA* $\Longrightarrow$ *c is decreasing in λ.*

Proof similiar to Theorem 15. $\square$

What about other monotonicity relations?

• Monotonicity in λ: intuitively, an increase in liquidaton costs should lead to a decrease of liquidation speed.

• Monotonicity in X: intuitively, larger asset position should lead to an *increased* liquidation speed.



$$\hat{\xi}(X, R)$$

IARA utility function with $A(R) = 2(1.5 + \tanh(R - 100))^2$ and parameter $\lambda = \sigma = 1$.

What about other monotonicity relations?

• Monotonicity in λ: intuitively, an increase in liquidaton costs should lead to a decrease of liquidation speed.

• Monotonicity in X: intuitively, larger asset position should lead to an *increased* liquidation speed.

What about other monotonicity relations?

• Monotonicity in λ: intuitively, an increase in liquidaton costs should lead to a decrease of liquidation speed.

• Monotonicity in X: intuitively, larger asset position should lead to an *increased* liquidation speed.

• Monotonicity in σ: intuitively, an increase in volatility should lead to an increase in the liquidation speed.

What about other monotonicity relations?

- Monotonicity in λ: intuitively, an increase in liquidaton costs should lead to a decrease of liquidation speed.

- Monotonicity in X: intuitively, larger asset position should lead to an *increased* liquidation speed.

- Monotonicity in σ: intuitively, an increase in volatility should lead to an increase in the liquidation speed.

?

The multi-asset case

Initial portfolio of d assets

$$\boldsymbol{X}_0 = (X_0^1, \ldots, X_0^d)$$

Strategy

$$\boldsymbol{X}_t^{\boldsymbol{\xi}} = \boldsymbol{X}_0 - \int_0^t \boldsymbol{\xi}_s \, ds$$

Price process:

$$\boldsymbol{S}_t = \boldsymbol{S}_0^0 + \sigma \boldsymbol{B}_t + \boldsymbol{\gamma}^\top (\boldsymbol{X}_t^{\boldsymbol{\xi}} - \boldsymbol{X}_0) - \boldsymbol{h}(\boldsymbol{\xi}_t)$$

for d-dim Brownian motion $\boldsymbol{B}$ and covariance matrix $\Sigma := \sigma\sigma^\top$.

Letting

$$f(\boldsymbol{\xi}) := \xi^\top \boldsymbol{h}(\boldsymbol{\xi}),$$

The revenues are

$$R_t^{\boldsymbol{\xi}} = R_0 + \int_0^t (\boldsymbol{X}_2^{\boldsymbol{\xi}})^\top \sigma \, d\boldsymbol{B}_s - \int_0^t f(\boldsymbol{\xi}_s) \, ds.$$

Guess for HJB equation

$$0 = \frac{1}{2} \boldsymbol{X}^\top \Sigma \boldsymbol{X} v_{RR} - \inf_{\boldsymbol{c}} \big(\boldsymbol{c}^\top \nabla_X v + v_R f(\boldsymbol{c}) \big)$$

with initial condition

$$v(0, R) = u(R).$$

The revenues are

$$R_t^{\boldsymbol{\xi}} = R_0 + \int_0^t (\boldsymbol{X}_2^{\boldsymbol{\xi}})^\top \sigma \, d\boldsymbol{B}_s - \int_0^t f(\boldsymbol{\xi}_s)\, ds.$$

Guess for HJB equation

$$0 = \frac{1}{2}\boldsymbol{X}^\top \Sigma \boldsymbol{X} v_{RR} - \inf_{\boldsymbol{c}} \big(\boldsymbol{c}^\top \nabla_{\boldsymbol{X}} v + v_R f(\boldsymbol{c})\big)$$

with initial condition

$$v(0, R) = u(R).$$

Formally: Nonlinear PDE of "parabolic" type with d time parameters

The revenues are

$$R_t^{\boldsymbol{\xi}} = R_0 + \int_0^t (\boldsymbol{X}_2^{\boldsymbol{\xi}})^\top \sigma \, d\boldsymbol{B}_s - \int_0^t f(\boldsymbol{\xi}_s)\, ds.$$

Guess for HJB equation

$$0 = \frac{1}{2}\boldsymbol{X}^\top \Sigma \boldsymbol{X} v_{RR} - \inf_{\boldsymbol{c}} \big(\boldsymbol{c}^\top \nabla_{\boldsymbol{X}} v + v_R f(\boldsymbol{c})\big)$$

with initial condition

$$v(0, R) = u(R).$$

Formally: Nonlinear PDE of "parabolic" type with d time parameters

Solvability completely unclear, a priori:

$$\nabla_{\boldsymbol{X}} v = g$$

typically not solvable (Poincaré lemma)

Theorem 18. [Schöneborn (2008)]
Under analogous conditions as in the onedimensional case and f having the scaling property

$$f(a\boldsymbol{\xi}) = a^{\alpha+1} f(\boldsymbol{\xi}), \qquad a \geq 0,$$

the value function is a classical solution of the HJB equation

$$0 = \frac{1}{2}\boldsymbol{X}^\top \Sigma \boldsymbol{X} v_{RR} - \inf_{\boldsymbol{c}} \big(\boldsymbol{c}^\top \nabla_{\boldsymbol{X}} v + v_R f(\boldsymbol{c}) \big)$$

with initial condition

$$v(0, R) = u(R).$$

The minimizer $\widehat{c}$ determines the optimal strategy....

Theorem 18. [Schöneborn (2008)]
Under analogous conditions as in the onedimensional case and f having the scaling property

$$f(a\boldsymbol{\xi}) = a^{\alpha+1} f(\boldsymbol{\xi}), \qquad a \geq 0,$$

the value function is a classical solution of the HJB equation

$$0 = \frac{1}{2} \boldsymbol{X}^\top \Sigma \boldsymbol{X} v_{RR} - \inf_{\boldsymbol{c}} \big(\boldsymbol{c}^\top \nabla_{\boldsymbol{X}} v + v_R f(\boldsymbol{c}) \big)$$

with initial condition

$$v(0, R) = u(R).$$

The minimizer $\widehat{c}$ determines the optimal strategy....

How can this be proved??

Theorem 19. [Schöneborn (2008)]
The optimal control is

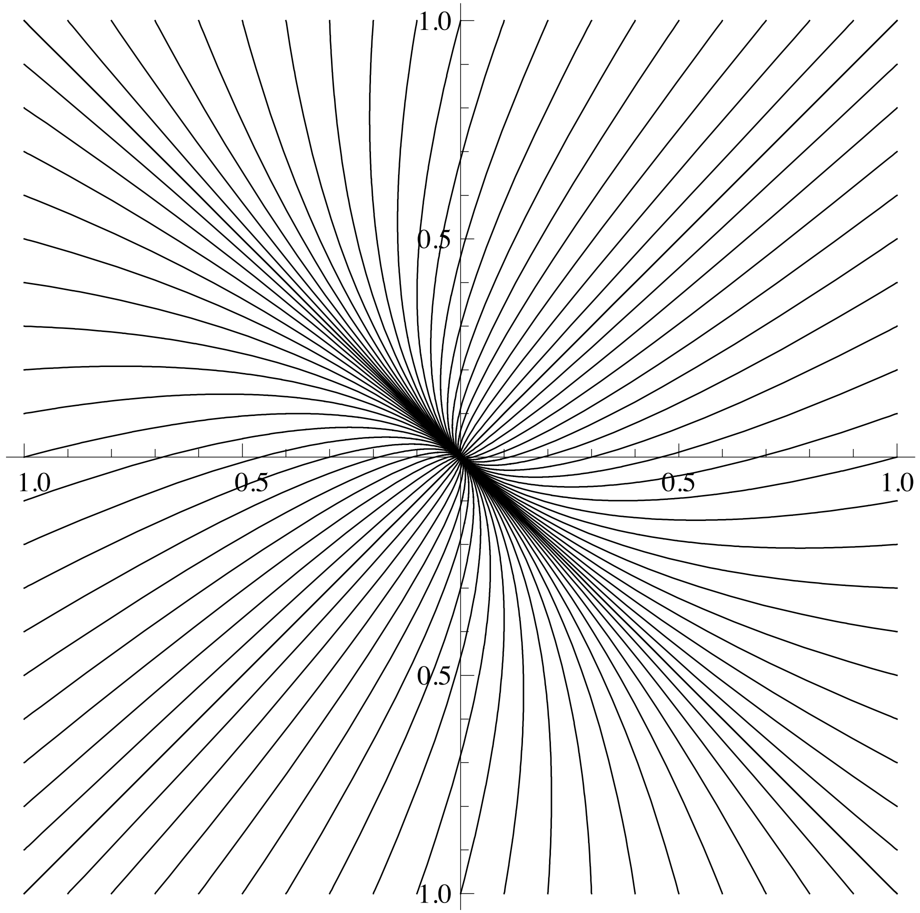
$$\widehat{c}(\boldsymbol{X},R)=\widetilde{c}\big(\overline{v}(\boldsymbol{X}),R)\overline{c}(\boldsymbol{X}),$$

where $\overline{v}(\boldsymbol{X})$ *is the cost and* $\overline{c}(\boldsymbol{X})$ *is the vector field (optimal strategy) for mean-variance optimal liquidation of* $\boldsymbol{X}$, *and* $\widetilde{c}(Y,R)$ *is the unique solution of the nonlinear PDE*

$$\widetilde{c}_Y=-\frac{2\alpha+1}{\alpha+1}\widetilde{c}^{\alpha}\widetilde{c}_R+\frac{\alpha(\alpha-1)}{\alpha+1}\Big(\frac{\widetilde{c}_R}{\widetilde{c}}\Big)^2+\frac{\alpha}{\alpha+1}\frac{\widetilde{c}_{RR}}{\widetilde{c}}$$

with initial condition

$$\widetilde{c}(0,R)=A(R)^{\frac{1}{\alpha+1}}$$



Trajectories for mean-variance optimal strategies for various initial portfolios $\boldsymbol{X}_0$ and two correlated assets.

II. The qualitative effects of risk aversion

1. **Exponential utility and mean-variance**
2. **General utility functions**
3. **Mean-variance optimization for model from model from Section I.1**

Consider return $R(X) = -\text{costs}$ instead of costs in model from Section I.1.

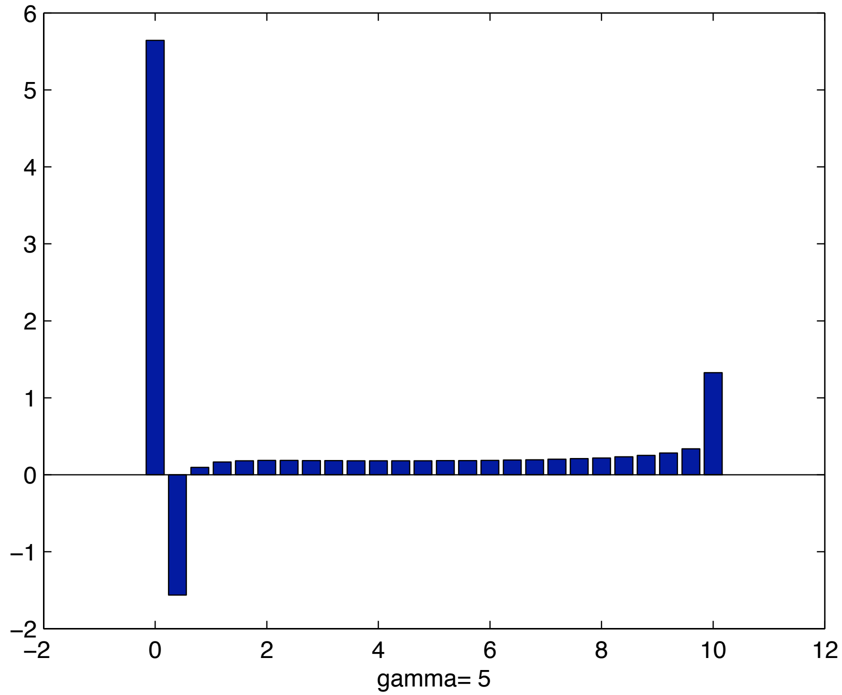
Theorem 20. *Suppose that G is strictly positive definite and that the unaffected price process S^0 satisfies $dS^0_t = \sigma_t\, dW_t$ for a Brownian motion W and a bounded and deterministic volatility function σ_s. Then the following conditions are equivalent for any strategy X^*.*

(a) *X^* maximizes the expected utility $\mathbb{E}[\,-e^{-\gamma R(X)}\,]$ in the class of all strategies X.*

(b) *X^* is deterministic and maximizes*

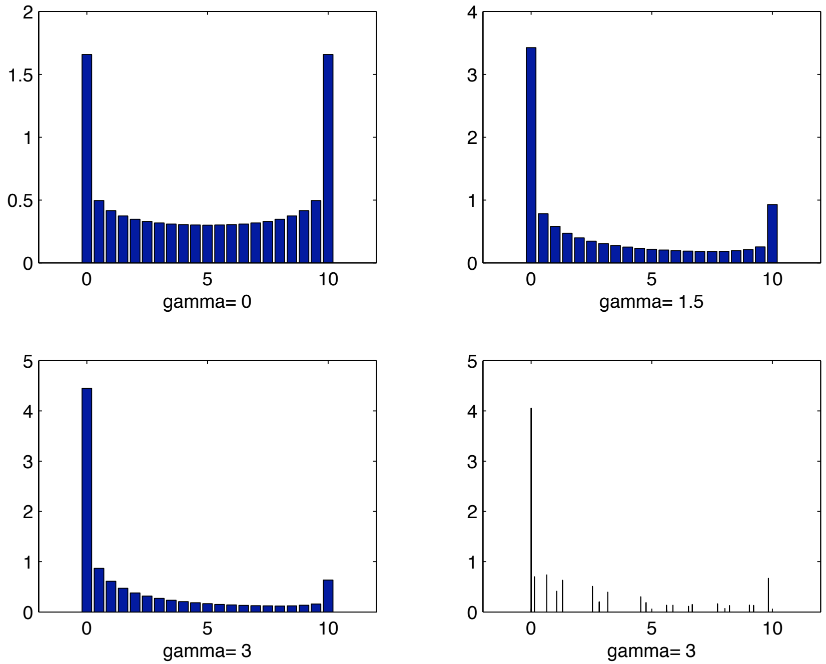
$$\mathbb{E}[\,R(X)\,] - \frac{\gamma}{2}\text{var}(R(X)),$$

in the class of deterministic strategies X.



Mean-variance optimal strategy for power-law decay $G(t) = (1+t)^{-0.4}$, covariance function $\varphi(t) = \sigma^2 t^{1/5}$ with volatility $\sigma = 0.3$, risk aversion $\gamma = 5$, and $N = 25$.

Theorem 21. *Suppose that $G(t)$ is convex, $\mathbb{T}$ is discrete, and the variance of S_t^0 increases as a convex function of t. Then any mean-variance optimal deterministic strategy X^* is monotone.*



Mean-variance optimal strategies for power-law decay $G(t) = (1+t)^{-0.4}$, linear covariance $\varphi(t) = \sigma^2 t$ with volatility $\sigma = 0.3$, and various risk aversion parameters γ.

III. Multi-agent equilibrium

References

Brunnermeier and Pedersen: *Predatory trading,* Journal of Finance 60, 1825–1863, (2005).

Carlin, Lobo, and Viswanathan: *Episodic liquidity crises: Cooperative and predatory trading, Journal of Finance* (2007).

T. Schöneborn and A.S.: *Liquidation in the face of adversity: stealth vs. sunshine trading.* Preprint, 2007.

C.C. Moallemi, B. Park, and B. Van Roy: *The execution game.* Preprint, 2008

Entire section based on Schöneborn and A.S. (2007)

Information leakage creates multi-player situations

- One trader ('the seller') must liquidate large portfolio by T_1
- Informed traders ('the predators') can exploit the resulting drift:
  - first short the asset
  - buy back shortly before T_1 at lower price

  "predatory trading"
- Suggests 'stealth trading strategy' for seller

Information leakage creates multi-player situations

- One trader ('the seller') must liquidate large portfolio by T_1
- Informed traders ('the predators') can exploit the resulting drift:
  - first short the asset
  - buy back shortly before T_1 at lower price

  "predatory trading"
- Suggests 'stealth trading strategy' for seller

- **But why, then, do some sellers practice 'sunshine trading'?**

- $n+1$ traders with positions $X_0(t)$, $X_1(t)$, ..., $X_n(t)$
- Trades at time t are executed at the price

$$S(t) = S(0) + \sigma B(t) + \gamma \sum_{i=0}^{n} (X_i(t) - X_i(0)) + \lambda \sum_{i=0}^{n} \dot{X}_i(t)$$

- Player 0 (the seller) has $X_0(0) > 0$, $X_0(t) = 0$ for $t \geq T_1$
- Players $1, \dots, n$ have $X_i(0) = 0$, $X_i(T_1) = \text{arbitrary}$, $X_i(T_2) = 0$
- Strategies are deterministic
- Players are risk-neutral and aim to maximize expected return

Goal: Find Nash equilibrium

Situation in a one-stage framework

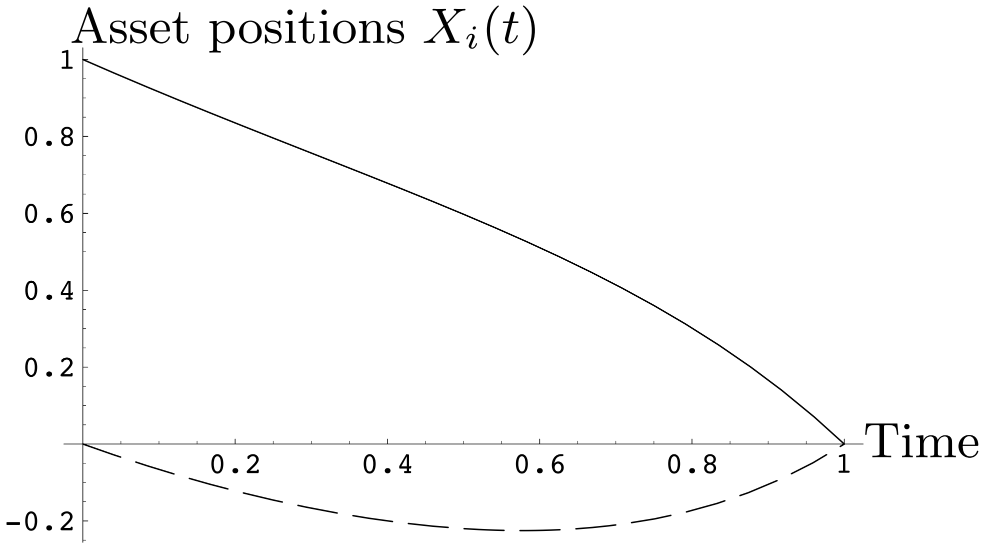
Theorem 1. [Carlin, Lobo, Viswanathan]
If $T_1 = T_2$, then the unique optimal strategies for these $n+1$ players are given by:

$$\dot{X}_i(t) = ae^{-\frac{n}{n+2}\frac{\gamma}{\lambda}t} + b_i e^{\frac{\gamma}{\lambda}t}$$

with

$$a = \frac{n}{n+2}\frac{\gamma}{\lambda}\left(1 - e^{-\frac{n}{n+2}\frac{\gamma}{\lambda}T_1}\right)^{-1}\frac{\sum_{i=0}^{n}(X_i(T_1) - X_i(0))}{n+1}$$

$$b_i = \frac{\gamma}{\lambda}\left(e^{\frac{\gamma}{\lambda}T_1} - 1\right)^{-1}\left(X_i(T_1) - X_i(0) - \frac{\sum_{j=0}^{n}(X_j(T_1) - X_j(0))}{n+1}\right).$$



Solid line $\sim$ seller, dashed line $\sim$ predator

- Predation occurs irrespective of the market parameters
- Predators always decrease the seller's return
- Predation becomes fiercer when the number of predators increases

$\Longrightarrow$ Model cannot explain sunshine trading or liquidity provision

Theorem 2.
In the two-stage framework, $T_2 > T_1$*, there is a unique Nash equilibrium, in which all predators acquire the same asset positions, and these are determined by their value at* T_1*:*

$$X_i(T_1) = \frac{A_2 n^2 + A_1 n + A_0}{B_3 n^3 + B_2 n^2 + B_1 n + B_0} X_0.$$

The coefficients A_i *and* B_i *are functions of* n *that converge in the limit* $n \uparrow \infty$*.*

Idea of Proof: Use result from Carlin et al., optimize over $X_i(T_1)$.

Coefficients in theorem can be computed exlicitly, e.g.,

$$
\begin{aligned}
A_0 \quad = \quad & 2\Bigg(- e^{\frac{\gamma(-T_1+(2+n)T_2)}{(1+n)\lambda}} - e^{\frac{\gamma(n(3+2n)T_1+(2+n)T_2)}{\left(2+3n+n^2\right)\lambda}} + \\
& e^{\frac{\gamma\left(\left(2+2n+n^2\right)T_1+n(2+n)T_2\right)}{\left(2+3n+n^2\right)\lambda}} + e^{\frac{\gamma\left(\left(-2+n^2\right)T_1+(2+n)^2T_2\right)}{\left(2+3n+n^2\right)\lambda}} + \\
& e^{\frac{\gamma(-nT_1+(1+2n)T_2)}{(1+n)\lambda}} - e^{\frac{\gamma\left(-nT_1+\left(2+5n+2n^2\right)T_2\right)}{\left(2+3n+n^2\right)\lambda}} + e^{\frac{n\gamma T_1+\gamma T_2}{\lambda+n\lambda}} - \\
& e^{\frac{\gamma T_1+n\gamma T_2}{\lambda+n\lambda}}\Bigg).
\end{aligned}
$$

Are there new effects in the two-stage model?

- **Plastic market:**

$$\text{temporary impact } \lambda \ll \text{permanent impact } \gamma$$

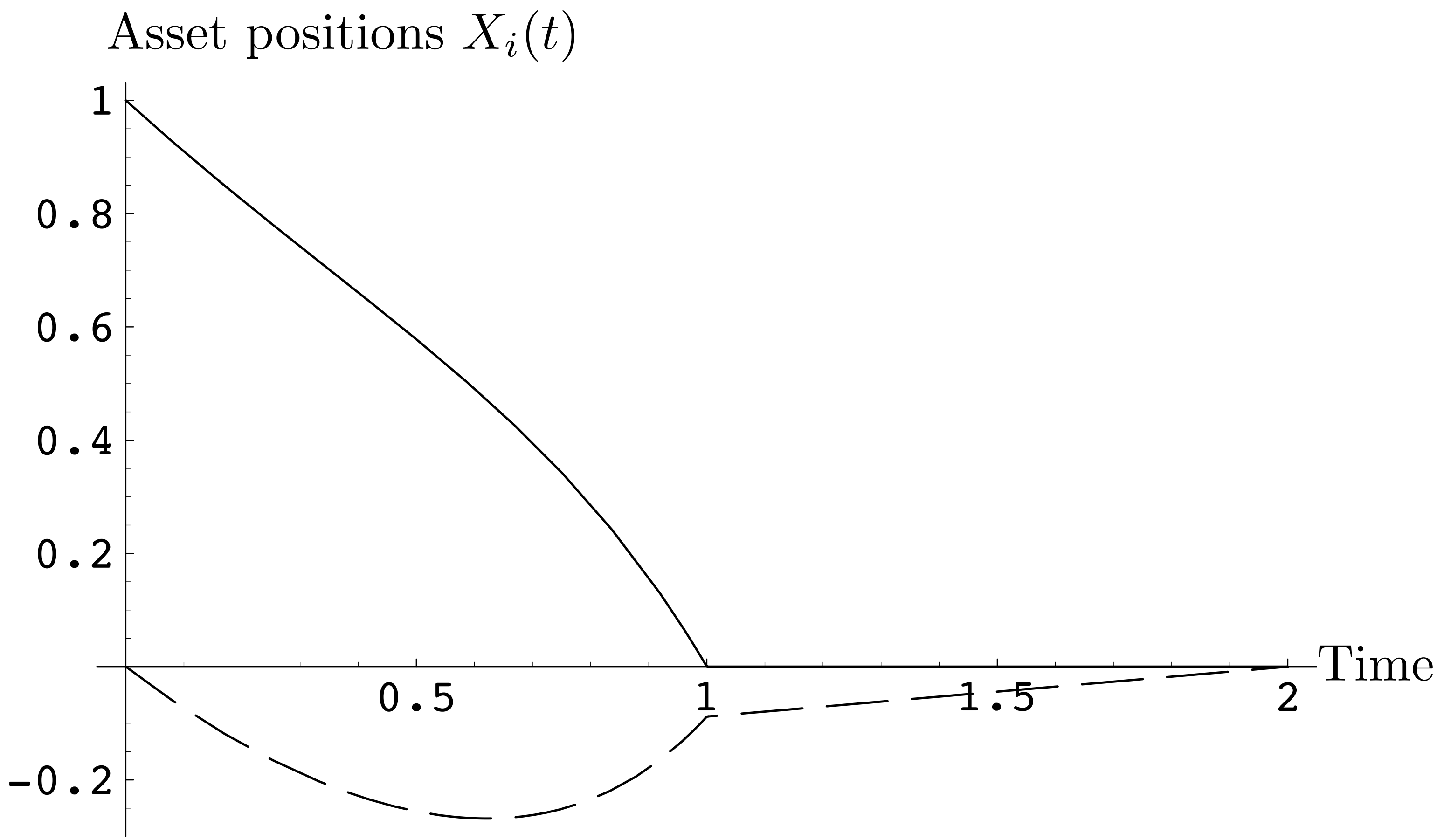
- **Elastic market:**

$$\text{temporary impact } \lambda \gg \text{permanent impact } \gamma$$

- **Intermediate market:**

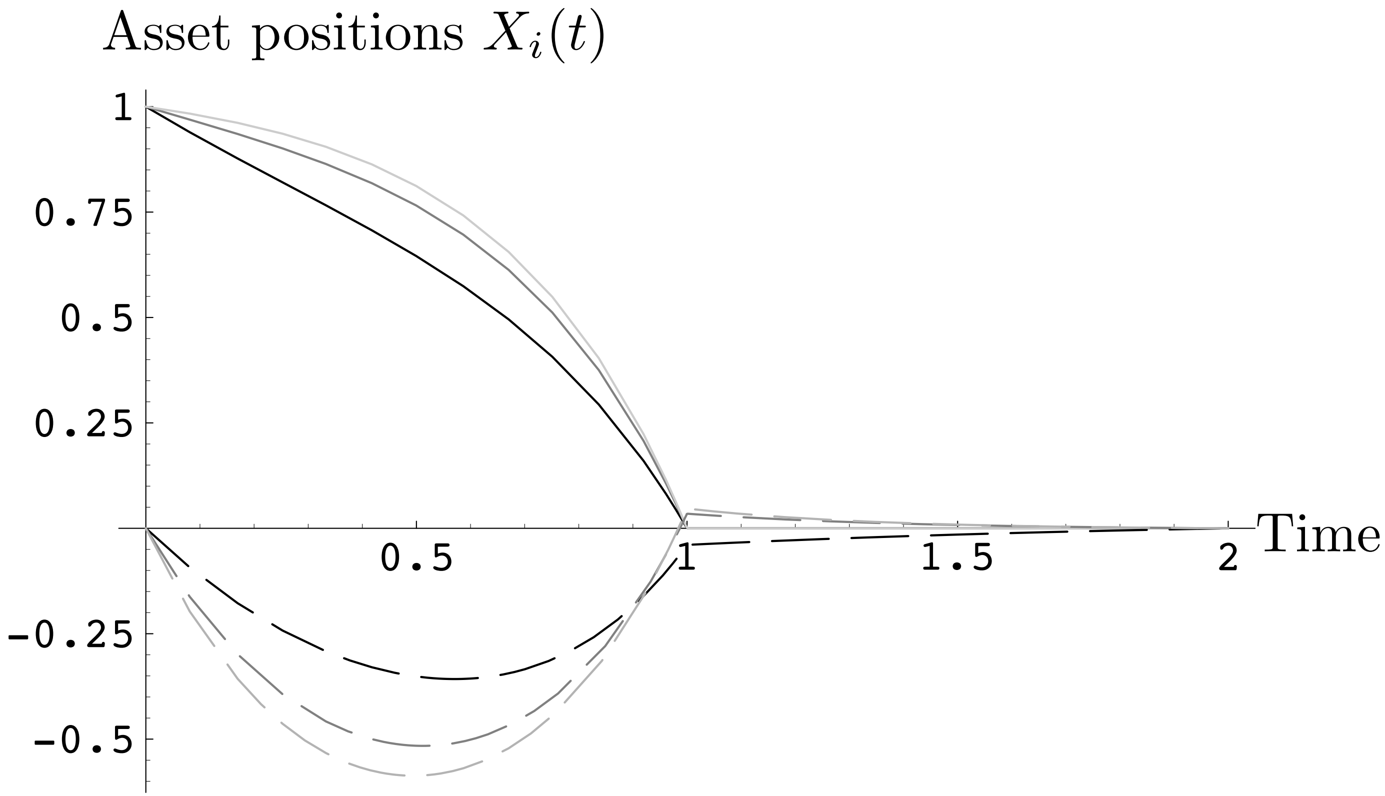
$$\text{temporary impact } \lambda \sim \text{permanent impact } \gamma$$

Plastic market (large perm. impact) one predator



Solid line ∼ seller, dashed line ∼ predator

Plastic market (large perm. impact)

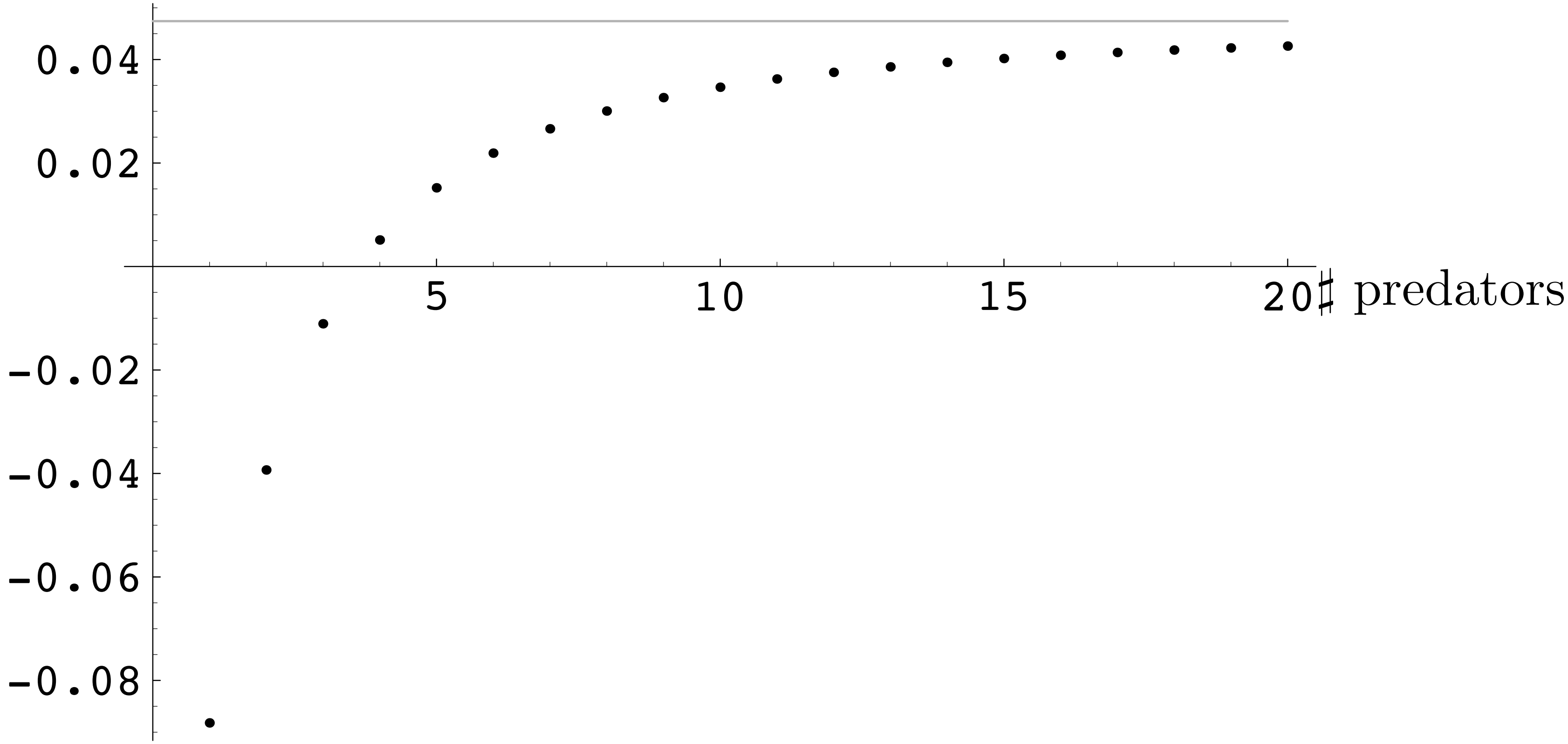


Solid lines $\sim$ seller, dashed lines $\sim$ n predators

Black $\sim n = 2$, dark grey $\sim n = 10$, light grey $\sim n = 100$

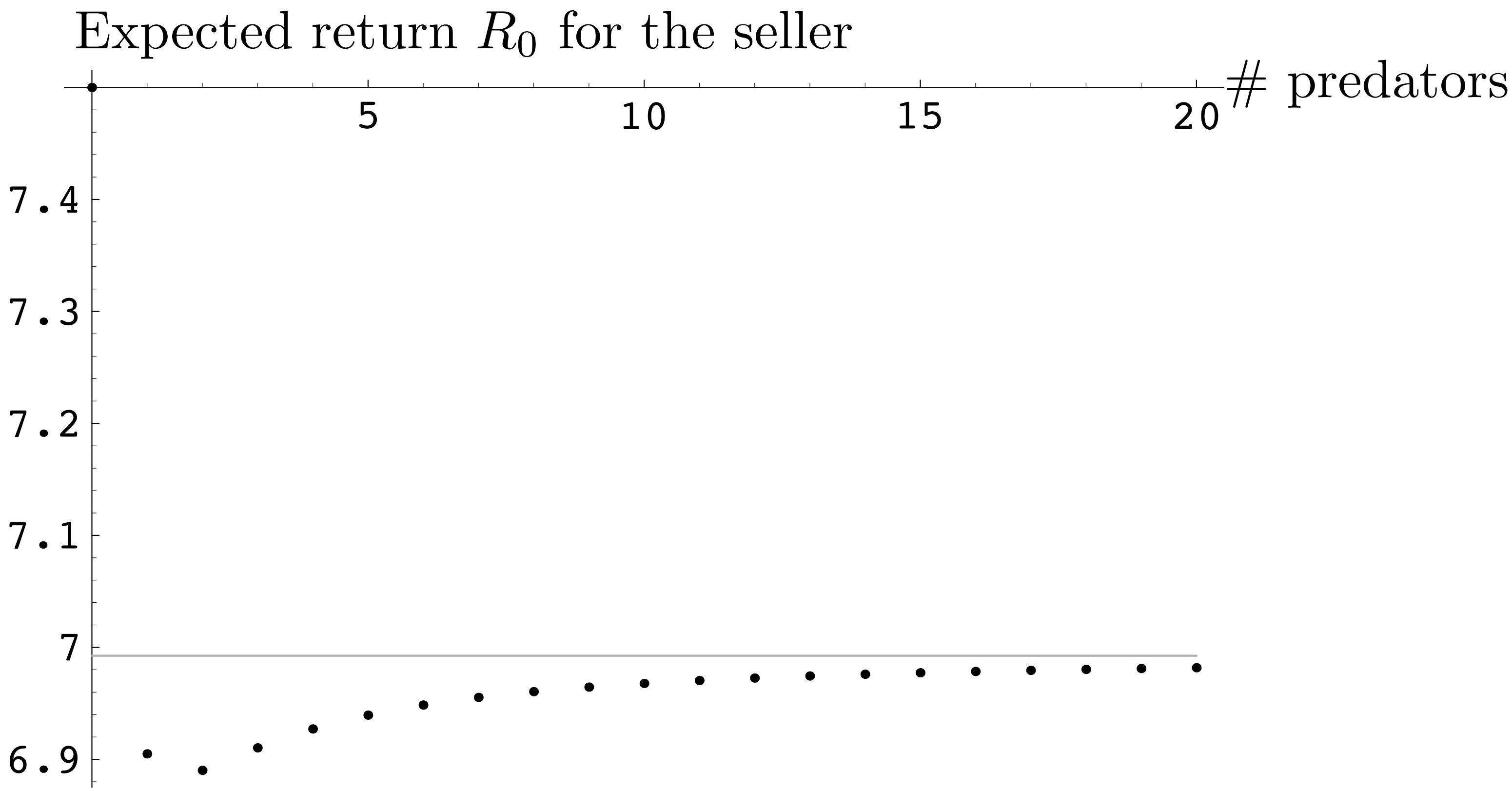
Plastic market (large perm. impact)

Joint asset position $\sum_{i=1}^{n} X_i(T_1)$ of all predators



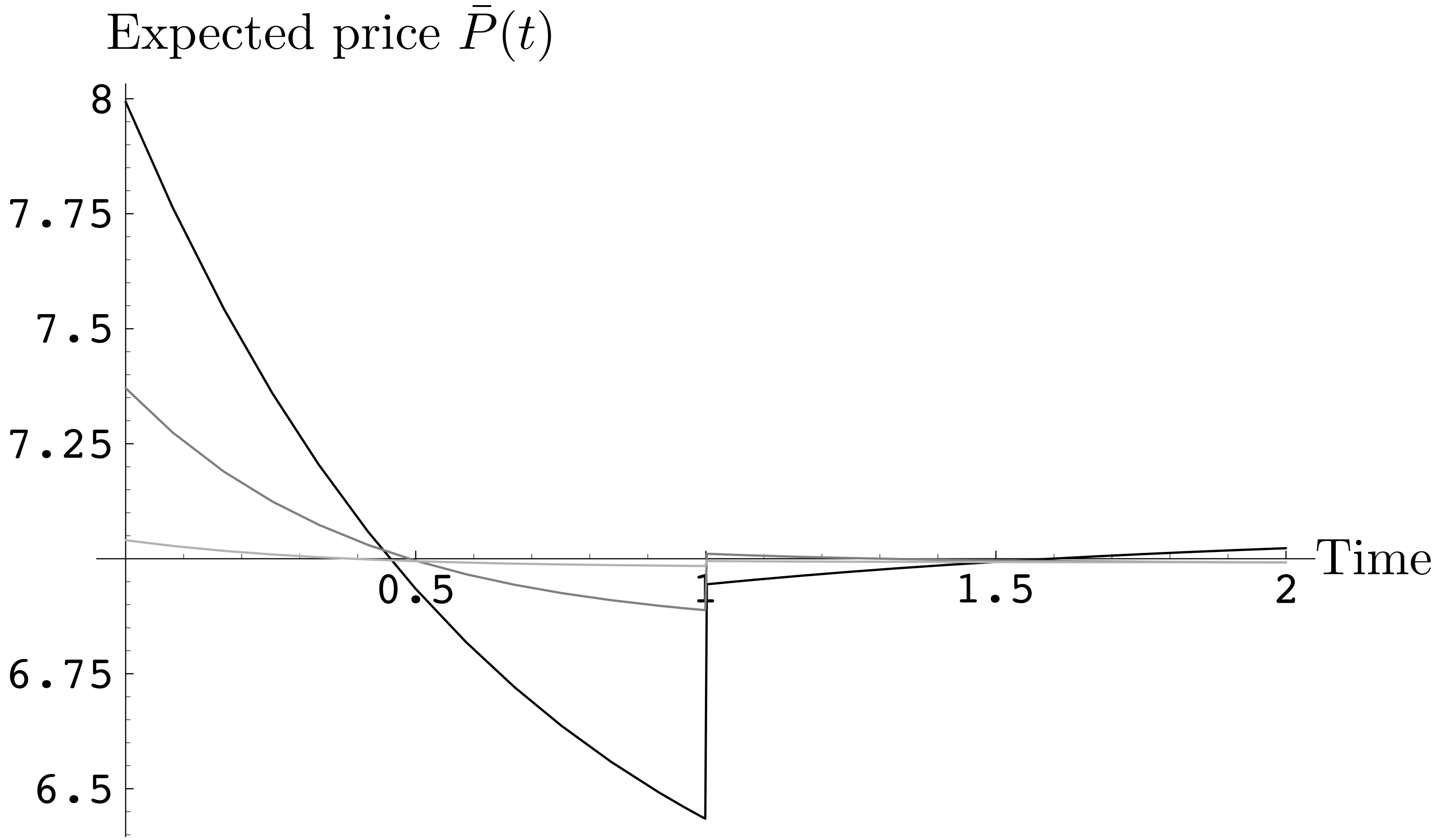
Upper grey line $= \lim_{n\to\infty} \sum_{i=1}^{n} X_i(T_1)$

Plastic market (large perm. impact)



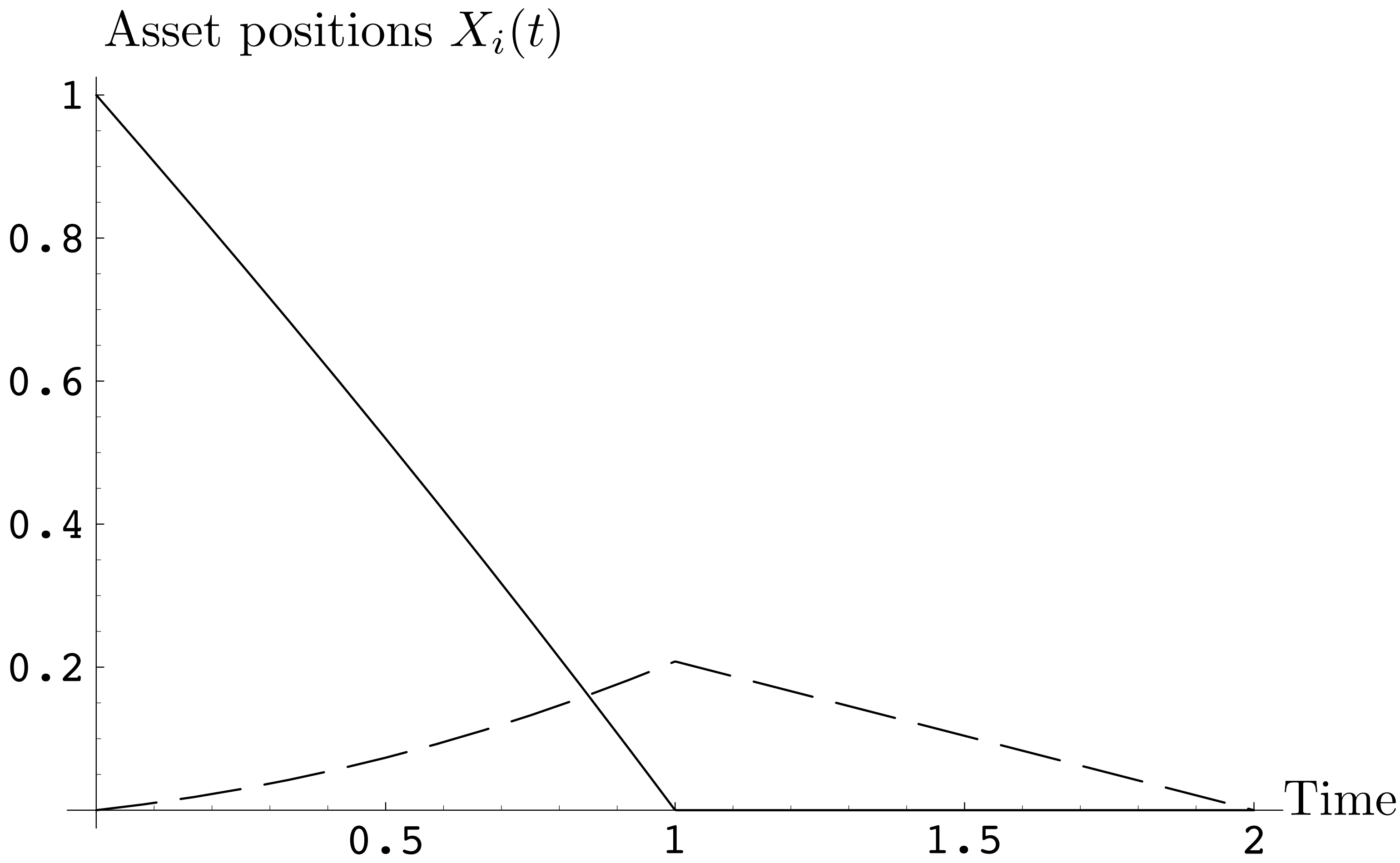
The grey line represents the limit $n \to \infty$. The return for the seller without predators is at the intersection of x- and y-axis.

Plastic market (large perm. impact)



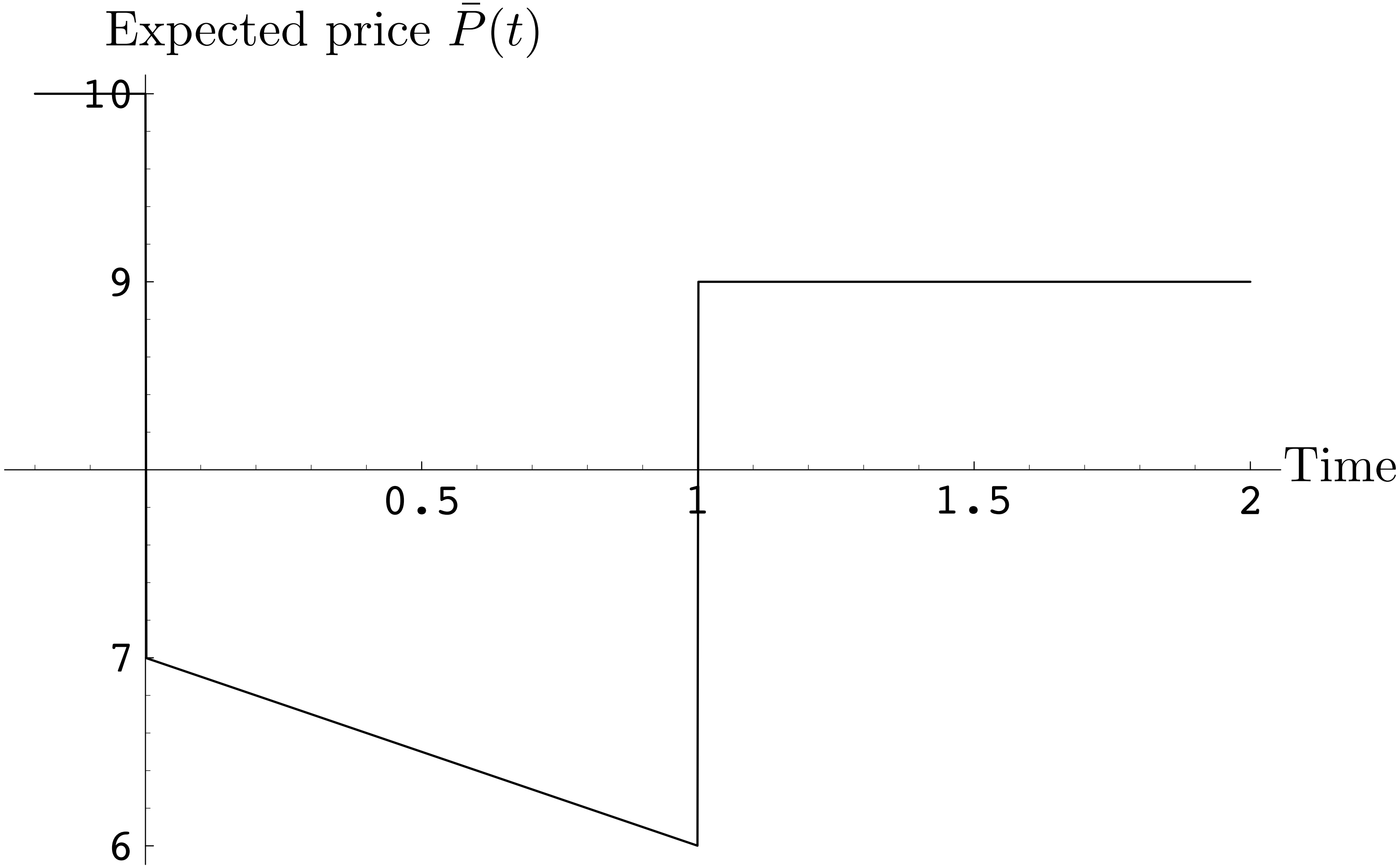
Black $\sim n = 2$, dark grey $\sim n = 10$, light grey $\sim n = 100$

Elastic market (large temp. impact) with one predator

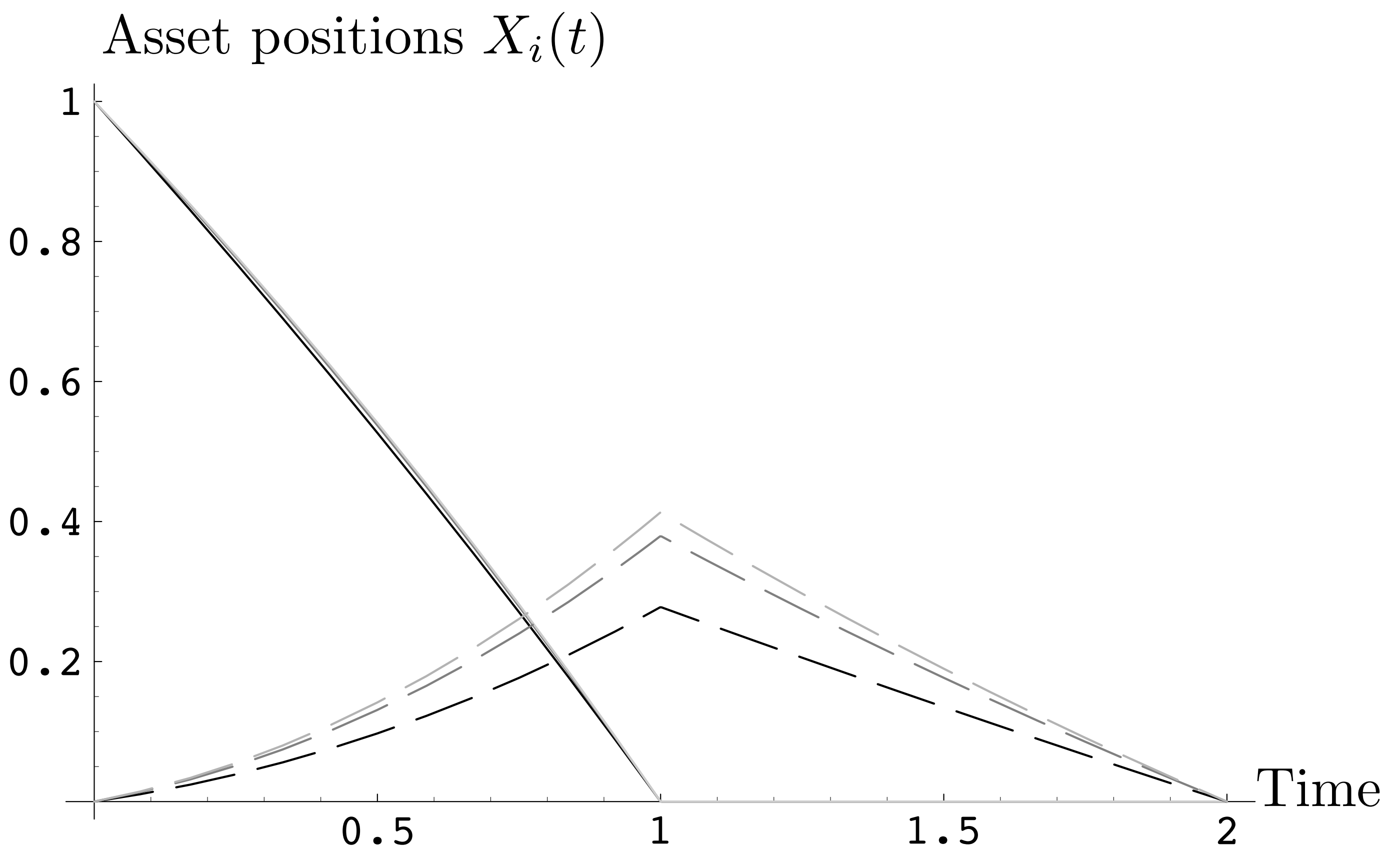


Solid line ~ seller, dashed line ~ predator

Elastic market (large temp. impact) without predators



Elastic market market (large temp. impact)

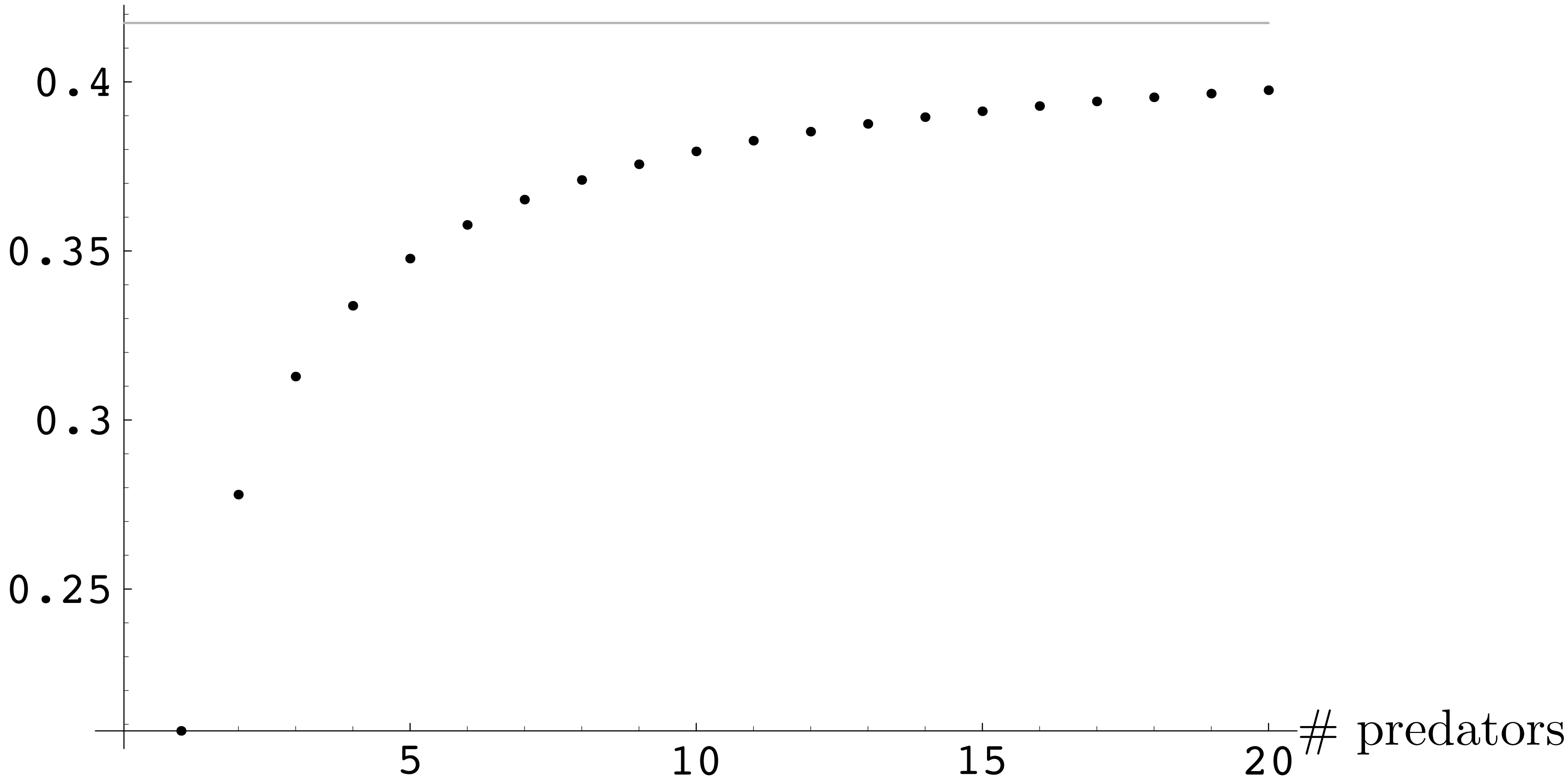


Solid lines $\sim$ seller, dashed lines $\sim$ n predators

Black $\sim n = 2$, dark grey $\sim n = 10$, light grey $\sim n = 100$

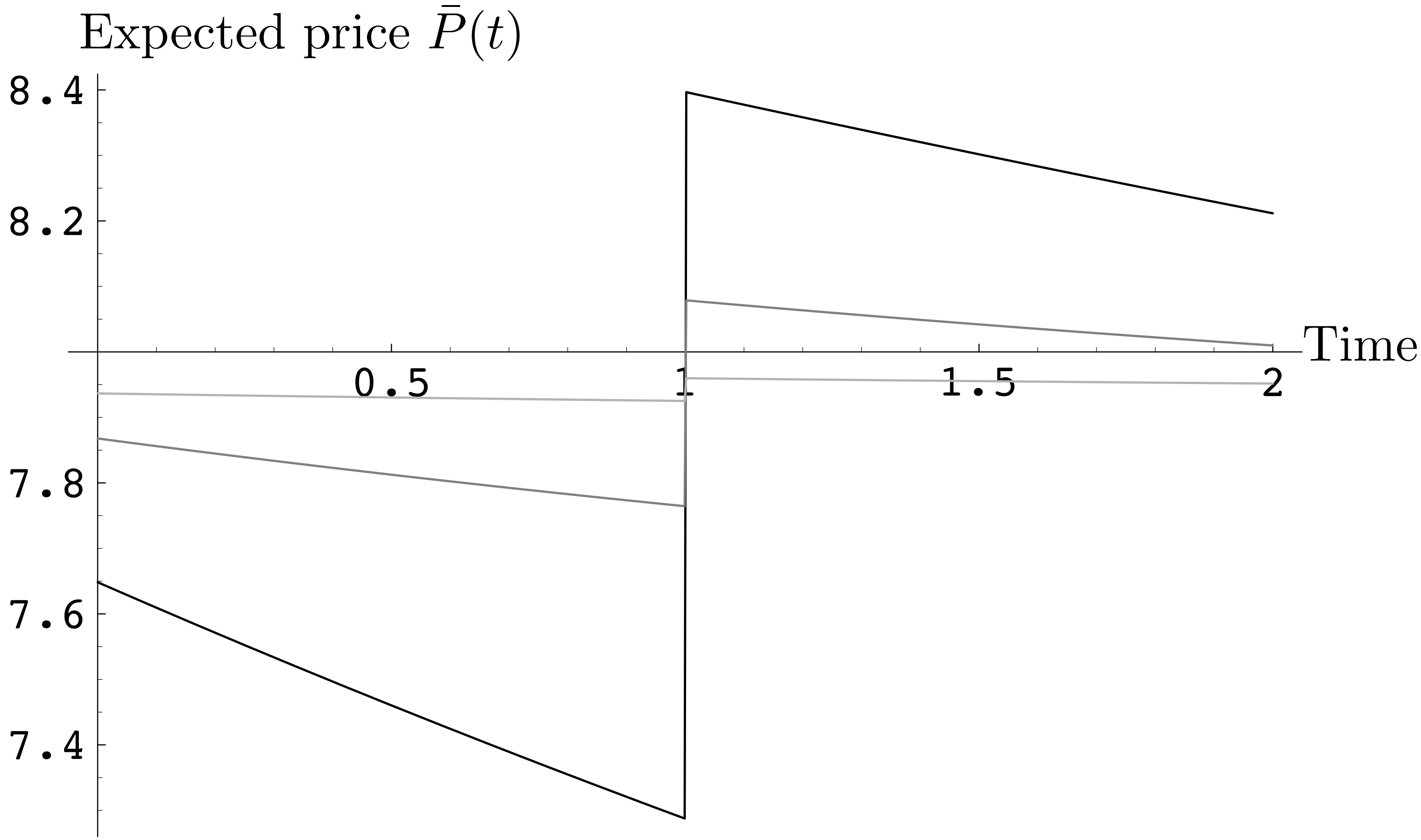
Elastic market (large temp. impact)

Joint asset position $\sum_{i=1}^{n} X_i(T_1)$ of all predators



The grey line represents the limit $\lim_{n\to\infty} \sum_{i=1}^{n} X_i(T_1)$

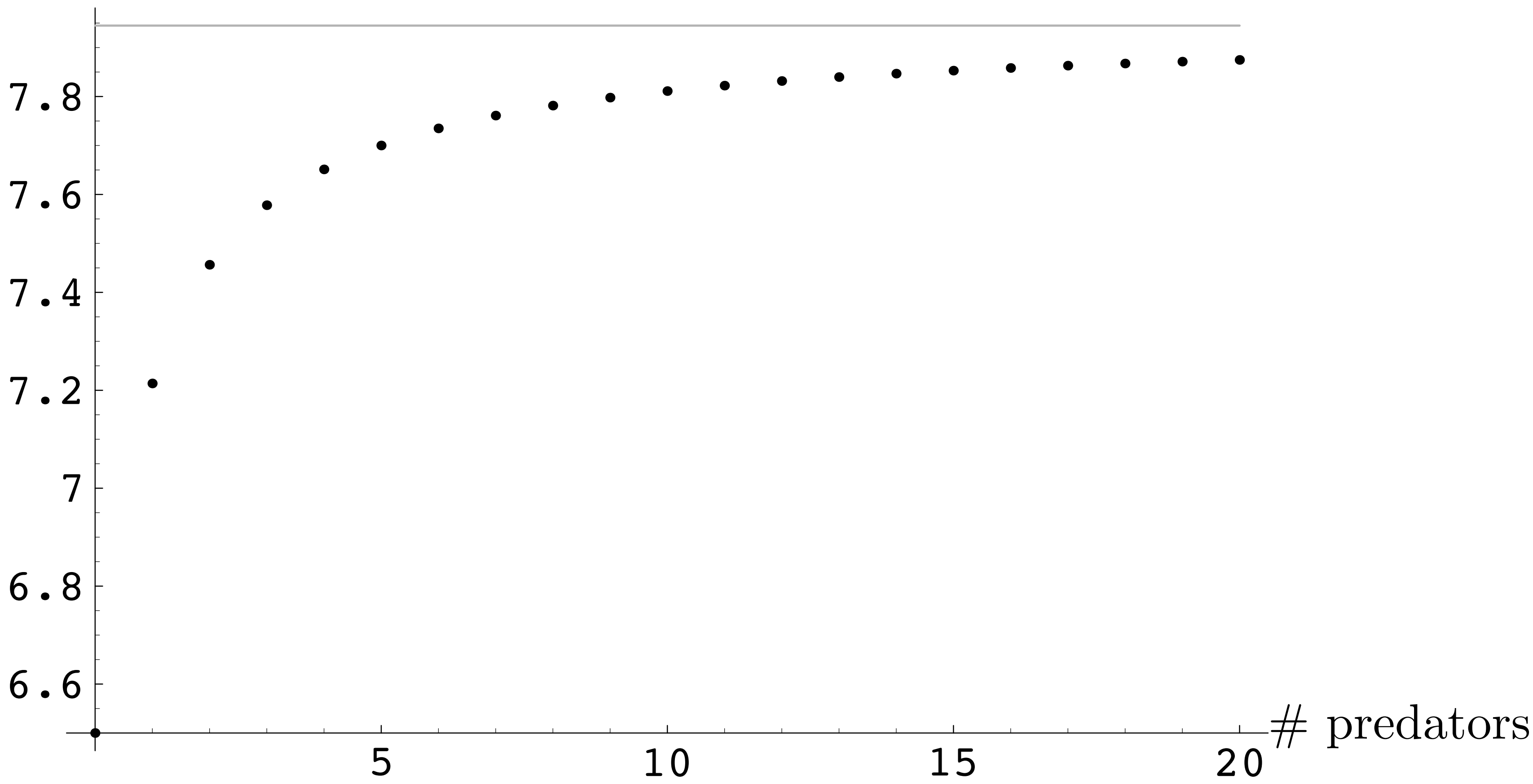
Elastic market (large temp. impact)



Black $\approx n = 2$, dark grey $\approx n = 10$, light grey $\approx n = 100$

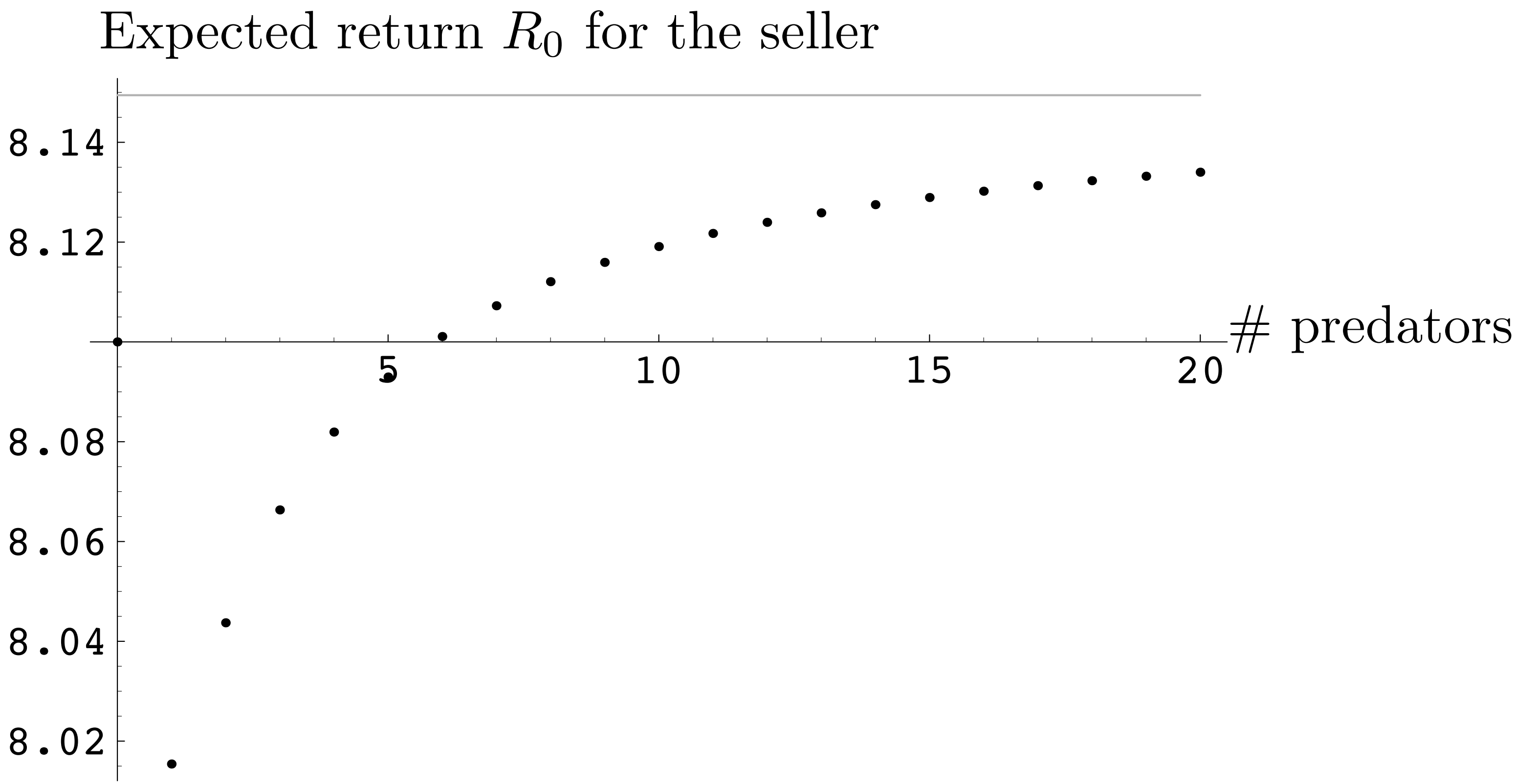
Elastic market (large temp. impact)

Expected return R_0 for the seller



The grey line represents the limit $n \to \infty$.

Moderate market ($\lambda \approx \gamma$)



The grey line represents the limit $n \to \infty$. The return for the seller without predators is at the intersection of x- and y-axis.

Theorem 3.

- *For all n, the asset position of the combined asset positions of the competitors is decreasing in $\gamma T_1/\lambda$*

- *As $n \uparrow \infty$, it converges to*

$$\lim_{n\to\infty} \sum_{i=1}^{n} X_i(T_1) = \lim_{n\to\infty} nX_1(T_1) = \frac{e^{\frac{\gamma(T_2-T_1)}{\lambda}} - 1}{e^{\frac{\gamma T_2}{\lambda}} - 1} X_0 > 0$$

- *For all n,*

$$\lim_{\gamma T_1/\lambda \downarrow 0} X_i(T_1) = \frac{T_2 - T_1}{(n+1)T_2} X_0 > 0 \qquad \lim_{\gamma T_1/\lambda \uparrow \infty} X_i(T_1) = \frac{-2X_0}{n^3 + 4n^2 + n - 2} < 0$$

- *For all n, $\dot{X}_i(t)$ is increasing in t and decreasing in $\gamma T_1/\lambda$ with*

$$\dot{X}_i(0) = \frac{T_2 - T_1}{(n+1)T_1 T_2} X_0 > 0 \qquad \textit{for } \gamma T_1/\lambda = 0$$

Corollary 4.
There are $L \leq P \in]0, \infty]$ *such that*

- *For* $0 \leq \gamma T_1/\lambda \leq L$, *the competitors are pure liquidity providers, i.e.,* $X_i(t) \geq 0$ *for* $0 \leq t \leq T$
- *For* $L \leq \gamma T_1/\lambda \leq P$, *there is first predatory trading, then liquidity provision, i.e.,* $\dot{X}_i(0) \leq 0$ *and* $X_i(T_1) \geq 0$
- *For* $P < \gamma T_1/\lambda$, *there is pure predation, i.e.,* $X_i(T_1) < 0$

Theorem 4.

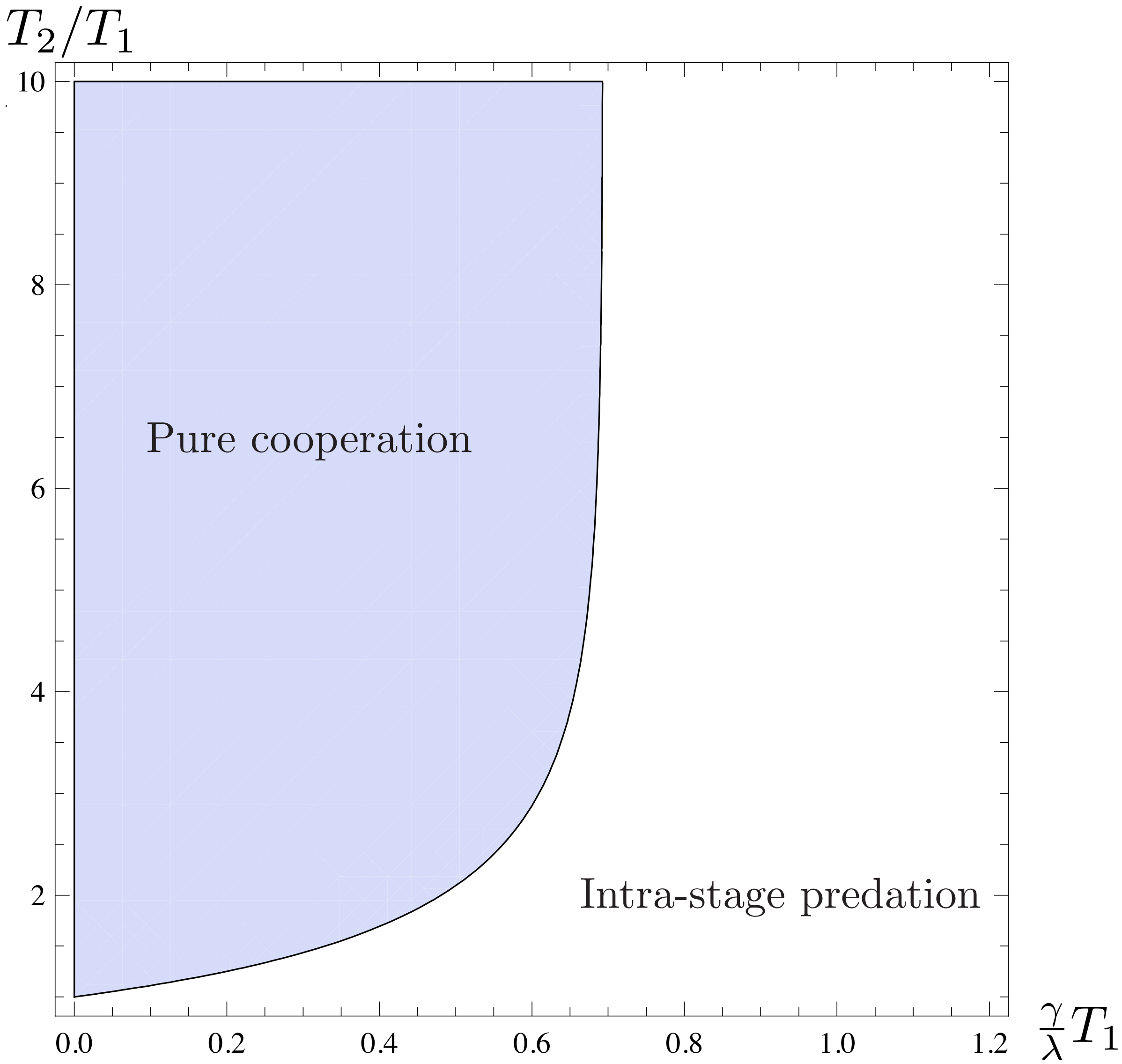
In competitive markets (i.e. in the limit $n \uparrow \infty$), the competitors are pure liquidity providers, i.e.,

$$\lim_{n\uparrow\infty} \sum_{i=1}^{n} X_i(t) > 0 \qquad \text{for } 0 < t \leq T_1$$

if and only if

$$\frac{T_2}{T_1} > -\frac{\log(2 - e^{\gamma T_1/\lambda})^+}{\frac{\gamma}{\lambda} T_1}$$

Otherwise, they engage in intra-stage predatory trading (i.e., $\sum_i \dot{X}_i(0) < 0$)


T_2/T_1
10
8
6
4
2
Pure cooperation
Intra-stage predation
0.0
0.2
0.4
0.6
0.8
1.0
1.2
$\frac{\gamma}{\lambda}T_1$

Stealth trading: no predators, expected return

$$X_0(P_0 - \gamma X_0/2 - \lambda X_0/T_1).$$

Sunshine trading: large number of predators, expected return

$$X_0\Big(P_0 - \frac{\gamma X_0}{1 - e^{-\gamma T_2/\lambda}}\Big)$$

Proposition 6. *For $n \uparrow \infty$, sunshine trading is superior to steath trading if*

$$\frac{1}{2} + \frac{\lambda}{\gamma T_1} > \frac{1}{1 - e^{-\frac{\gamma}{\lambda} T_2}}.$$

For $T_2 \uparrow \infty$, a stealth algorithm is beneficial if

$$\frac{\gamma}{\lambda} T_1 < 2$$

Predatory trading vs. liquidity provision: anecdotal evidence

Conclusion

Have studied optimal execution problems on three different levels

- **Microscopic:** Order book models
- **Mesoscopic:** Expected utility maximization in stylized model
- **Macroscopic:** Multi-agent situation; stealth vs. sunshine trading, predation vs. liquidity provision

Thank you