



HAL
open science

Dealing with the Inventory Risk

Olivier Guéant, Charles-Albert Lehalle, Joaquin Fernandez Tapia

► **To cite this version:**

Olivier Guéant, Charles-Albert Lehalle, Joaquin Fernandez Tapia. Dealing with the Inventory Risk. 2011. hal-00628528

HAL Id: hal-00628528

<https://hal.science/hal-00628528>

Preprint submitted on 3 Oct 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Dealing with the Inventory Risk

Olivier Guéant*, Charles-Albert Lehalle**, Joaquin Fernandez Tapia***

First draft - December 2010. This version: September 2011.

Abstract

Market makers have to continuously set bid and ask quotes for the stocks they have under consideration. Hence they face a complex optimization problem in which their return, based on the bid-ask spread they quote and the frequency they indeed provide liquidity, is challenged by the price risk they bear due to their inventory. In this paper, we provide optimal bid and ask quotes and closed-form approximations are derived using spectral arguments.

Introduction

The optimization of the intra-day trading process on electronic markets was born with the need to split large trades to make the balance between trading too fast (and possibly degrade the obtained price via “market impact”) and trading too slow (and suffer from a too long exposure to “market risk”). This “trade scheduling” viewpoint has been mainly formalized in the late nineties by Bertsimas and Lo [8] and Almgren and Chriss [2]. More sophisticated approaches involving the use of stochastic and impulse control have been proposed since then (see for instance [9]). Another branch of proposals goes in the direction of modeling the effect of the “aggressive” (*i.e.* liquidity consuming) orders at the finest level, for instance via a martingale model of the behavior market depth and of its resilience (see [1]).

From a quantitative viewpoint, market microstructure is a sequence of auction games between market participants. It implements the balance between supply and demand, forming an equilibrium traded price to be used as reference for valuation of the listed assets. The rule of each auction game (fixing auction, continuous auction, etc), are fixed by the firm operating each trading venue. Nevertheless, most of all trading mechanisms on electronic markets rely on market participants sending orders to a “queuing system” where their open interests are consolidated as “liquidity provision” or form transactions [3]. The efficiency of such a process relies on an adequate timing between buyers and sellers, to avoid too many non-informative oscillations of the transaction price (for more details and modeling, see for example [19]).

To take profit of these oscillations, it is possible to provide liquidity to an impatient buyer (respectively seller) and maintain an inventory until the arrival of the next impatient seller

The authors wish to acknowledge the helpful conversations with Pierre-Louis Lions (Collège de France), Jean-Michel Lasry (Université Paris-Dauphine), Yves Achdou (Université Paris-Diderot), Thierry Foucault (HEC), Albert Menkveld (VU University Amsterdam), Vincent Millot (Université Paris-Diderot), Antoine Lemenant (Université Paris-Diderot) and Vincent Fardeau (London School of Economics).

*UFR de Mathématiques, Laboratoire Jacques-Louis Lions, Université Paris-Diderot. 175, rue du Chevaleret, 75013 Paris, France. olivier.gueant@ann.jussieu.fr

**Head of Quantitative Research. Crédit Agricole Cheuvreux. 9, Quai du Président Paul Doumer, 92400 Courbevoie, France. clehalle@cheuvreux.com

***PhD student, Université Paris 6 - Pierre et Marie Curie. 4 place Jussieu, 75005 Paris, France.

(respectively buyer). Market participants focused on this kind of liquidity-providing activity are called “market makers”. On one hand they are buying at the bid price and selling at the ask price they chose, taking profit of this “bid-ask spread”. On the other hand, their inventory is exposed to price fluctuations mainly driven by the volatility of the market (see [4, 7, 11, 13, 16, 23]).

The usual “market making problem” comes from the optimality of the quotes (*i.e.* the bid and ask prices) that such agents provide to other market participants with respect to the constraints on their inventory and their utility function as a proxy to their risk (see [10, 15, 17, 20, 22, 25]).

The recent evolution of market microstructure and the financial crisis reshaped the nature of the interactions of the market participants during electronic auctions, one consequence being the emergence of “high-frequency market makers” who are said to be part of 70% of the electronic trades and have a massively passive (*i.e.* liquidity providing) behavior. A typical balance between passive and aggressive orders for such market participants is around 80% of passive interactions (see [21]).

Avellaneda and Stoikov proposed in [5] an innovative framework for “market making in an order book” and studied it using different approximations. In such an approach, the “fair price” S_t is modeled via a Brownian motion with volatility σ , and the arrival of a buy or sell liquidity-consuming order at a distance δ of S_t follows a Poisson process with intensity $A \exp(-k \delta)$. Our paper extends their proposal and provides results in two main directions:

- An explicit solution to the Hamilton-Jacobi-Bellman equation coming from the optimal market making problem thanks to a non trivial change of variables and the resulting expressions for the optimal quotes:

Main Result 1 (Proposition 2 - Theorem 4). *The optimal bid and ask quotes can be expressed as:*

$$s^{b*}(t, q, s) = s - \left(\frac{1}{k} \ln \left(\frac{v_q(t)}{v_{q+1}(t)} \right) + \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right) \right)$$

$$s^{a*}(t, q, s) = s + \left(\frac{1}{k} \ln \left(\frac{v_q(t)}{v_{q-1}(t)} \right) + \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right) \right)$$

where γ is the risk aversion of the agent and where v is a family of strictly positive functions $(v_q)_{|q| \leq Q}$ solution of the linear system of ODEs that follows:

$$\forall q \in \{-Q + 1, \dots, Q - 1\}, \quad \dot{v}_q(t) = \alpha q^2 v_q(t) - \eta (v_{q-1}(t) + v_{q+1}(t))$$

$$\dot{v}_Q(t) = \alpha Q^2 v_Q(t) - \eta v_{Q-1}(t)$$

$$\dot{v}_{-Q}(t) = \alpha Q^2 v_{-Q}(t) - \eta v_{-Q+1}(t)$$

with $\forall q \in \{-Q, \dots, Q\}, v_q(T) = 1$, $\alpha = \frac{k}{2} \gamma \sigma^2$ and $\eta = A \left(1 + \frac{\gamma}{k} \right)^{-(1 + \frac{k}{\gamma})}$, T being the time horizon and Q a bound we impose on the inventory of the market maker, be it short or long.

It means that to find an exact solution to the generic high-frequency market making problem, it is enough to solve on the fly the companion ODEs in $v_q(t)$ provided by our change of variables, and to plug the result in the upper equalities to obtain the optimal quotes with respect to a given inventory and market state.

- Asymptotics of the solution that are numerically attained fast enough in most realistic cases:

Main Result 2 (Theorem 5 (asymptotics) and the associated approximation equations).

$$\begin{aligned} \lim_{T \rightarrow \infty} s - s^{b^*}(0, q, s) = \delta_{\infty}^{b^*}(q) &\simeq \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right) + \frac{2q+1}{2} \sqrt{\frac{\sigma^2 \gamma}{2kA} \left(1 + \frac{\gamma}{k} \right)^{1+\frac{k}{\gamma}}} \\ \lim_{T \rightarrow \infty} s^{a^*}(0, q, s) - s = \delta_{\infty}^{a^*}(q) &\simeq \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right) - \frac{2q-1}{2} \sqrt{\frac{\sigma^2 \gamma}{2kA} \left(1 + \frac{\gamma}{k} \right)^{1+\frac{k}{\gamma}}} \end{aligned}$$

These results open doors to new directions of research involving the modeling and control of passive interactions with electronic order books. If some attempts have been made that did not rely on stochastic control but on forward optimization (see for instance [24] for a stochastic algorithmic approach for optimal split of passive orders across competing electronic order books), they should be complemented by backward ones.

This paper starts in section 1 with the description of the model choices that had to be made. Section 2 is dedicated to the presentation of the ideas underlying the results of this paper. These ideas are developed in the framework of the initial Avellaneda and Stoikov setting. Section 3 then goes back to our model with bounds on the inventory and, after having found the optimal quotes, exposes the asymptotics of the obtained dynamics along with applications of the approximations developed in section 2. Section 4 carries out the comparative statics. Section 5 extends the framework to trends in prices and costly terminal liquidation. We finally discuss some issues with the hypotheses of the model in section 6 and end with an application to real data in section 7. Adaptations of our results are already in use at Crédit Agricole Cheuvreux to optimize the brokerage trading flow.

In our framework, we follow Avellaneda and Stoikov in using a Poisson process model pegged on a “fair price” diffusion (see section 1). As it is discussed in section 6, it is an arguable choice since it does not capture “resistances” that can be built by huge passive (*i.e.* liquidity-providing) orders preventing the market price to cross their prices. Our results cannot be used as such for large orders, but are perfectly suited for high-frequency market making as it is currently implemented in the market, using orders of small size (close to the average trade size, see [21]).

Moreover, to our knowledge, no quantitative model of “implicit market impact” of such small passive orders has never been proposed in the literature, despite very promising studies linking updates of quantities in the order books to price changes (see [12]). Its combination with recent applications of more general point processes to capture the process of arrival of orders (like Hawkes models, see [6]) should give birth to such implicit market impact models, specifying dependencies between the trend, the volatility and possible jumps in the “fair price” semi-martingale process with the parameters of the multi-dimensional point process of the market maker fill rate. At this stage, the explicit injection of such path-dependent approach (once they will be proposed in the literature) into our equations are too complex to be handled, but numerical explorations around our explicit formulas will be feasible. The outcomes of applications of our results to real data (section 7) show that they are realistic enough so that no more than small perturbations should be needed.

1 Setup of the model

1.1 Prices, orders and objective function

We consider a small high-frequency market maker operating on a single stock¹. For the sake of simplicity and since we will basically only consider short horizon problems we suppose

¹We suppose that this high-frequency market maker does not “make” the price in the sense that he has no market power. In other words, we assume that his size is small enough to consider price dynamics exogenous. Small high-frequency proprietary trader operating on both sides would be another way to describe our agent.

that a fair price² of the stock moves as an arithmetic brownian motion:

$$dS_t = \sigma dW_t$$

The market maker under consideration will continuously propose bid and ask quotes denoted respectively S_t^b and S_t^a and will hence buy and sell shares according to the rate of arrival of aggressive orders at the quoted prices.

His inventory q , that is the (signed) quantity of shares he holds, is given by $q_t = N_t^b - N_t^a$ where N^b and N^a are the jump processes giving the number of shares the market maker respectively bought and sold. These jump processes are supposed to be Poisson processes and to simplify the exposition (although this may be important, see the discussion part) we consider that jumps are of constant size, scaled to 1. Arrival rates obviously depend on the prices $S_t^b = S_t - \delta_t^b$ and $S_t^a = S_t + \delta_t^a$ quoted by the market maker and we assume, in accordance with most datasets³, that intensities λ^b and λ^a associated to N^b and N^a are of the following form⁴:

$$\lambda^b(\delta^b) = Ae^{-k\delta^b} = A \exp(-k(s - s^b)) \quad \lambda^a(\delta^a) = Ae^{-k\delta^a} = A \exp(-k(s^a - s))$$

This means for positive δ^b and δ^a that the closer to the fair price an order is quoted, the faster it will be executed.

As a consequence of his trades, the market maker has an amount of cash whose dynamics is given by:

$$dX_t = (S_t + \delta_t^a)dN_t^a - (S_t - \delta_t^b)dN_t^b$$

To this original setting introduced by Avellaneda and Stoikov (itself following Ho and Stoll [18]) we add a bound Q on the inventory that a market maker is authorized to have. In other words, we assume that a market maker with inventory Q (Q depending in practice on risk limits) will never set a bid quote and symmetrically that a market maker with inventory $-Q$, that is a short position of Q shares in the stock under consideration, will never set an ask quote.

Now, coming to the objective function and as we said above, the market maker has a time horizon T and his goal is to optimize the expected utility of his P&L at time T . In line with [5], we will focus on CARA utility functions and we suppose that the market maker optimizes:

$$\sup_{(\delta_t^a)_t, (\delta_t^b)_t \in \mathcal{A}} \mathbb{E}[-\exp(-\gamma(X_T + q_T S_T))]$$

where \mathcal{A} is the set of predictable processes bounded from below, where γ is the absolute risk aversion characterizing the market maker, where X_T is the amount of cash at time T and where $q_T S_T$ is the evaluation of the (signed) remaining quantity of shares in the inventory at time T (liquidation at the fair price S_T ⁵).

1.2 The (HJB) equation and the original Avellaneda-Stoikov (AS-HJB) equation

To the optimization problem set up in the preceding section, one can associate a Hamilton-Jacobi-Bellman (HJB) equation, or more exactly a system of HJB PDEs which consists in the following equations

²This fair price for the stock can be thought of as a smoothed mid-price for instance.

³At least inside the limit order books.

⁴Some authors also used a linear form for the intensity functions – see [18] for instance.

⁵We will discuss other hypotheses below.

$$\begin{aligned}
& \forall q \in \{-Q+1, \dots, Q-1\}, \quad \partial_t u(t, x, q, s) + \frac{1}{2} \sigma^2 \partial_{ss}^2 u(t, x, q, s) \\
& \quad + \sup_{\delta^b} \lambda^b(\delta^b) \left[u(t, x - s + \delta^b, q + 1, s) - u(t, x, q, s) \right] \\
& \quad + \sup_{\delta^a} \lambda^a(\delta^a) \left[u(t, x + s + \delta^a, q - 1, s) - u(t, x, q, s) \right] = 0 \\
& \quad \partial_t u(t, x, Q, s) + \frac{1}{2} \sigma^2 \partial_{ss}^2 u(t, x, Q, s) \\
& \quad + \sup_{\delta^a} \lambda^a(\delta^a) \left[u(t, x + s + \delta^a, Q - 1, s) - u(t, x, Q, s) \right] = 0 \\
& \quad \partial_t u(t, x, -Q, s) + \frac{1}{2} \sigma^2 \partial_{ss}^2 u(t, x, -Q, s) \\
& \quad + \sup_{\delta^b} \lambda^b(\delta^b) \left[u(t, x - s + \delta^b, -Q + 1, s) - u(t, x, -Q, s) \right] = 0
\end{aligned}$$

with the final condition:

$$\forall q \in \{-Q, \dots, Q\}, \quad u(T, x, q, s) = -\exp(-\gamma(x + qs))$$

This system differs from the initial equation introduced by Avellaneda and Stoikov because of the bounds we imposed on the inventory. The original equation was indeed

$$\begin{aligned}
(\text{AS-HJB}) \quad 0 &= \partial_t u(t, x, q, s) + \frac{1}{2} \sigma^2 \partial_{ss}^2 u(t, x, q, s) \\
& \quad + \sup_{\delta^b} \lambda^b(\delta^b) \left[u(t, x - s + \delta^b, q + 1, s) - u(t, x, q, s) \right] \\
& \quad + \sup_{\delta^a} \lambda^a(\delta^a) \left[u(t, x + s + \delta^a, q - 1, s) - u(t, x, q, s) \right]
\end{aligned}$$

with the final condition:

$$u(T, x, q, s) = -\exp(-\gamma(x + qs))$$

and we are going to solve this equation before solving our problem because the arguments underlying our closed-form results are rooted to this equation on an unbounded domain and can then be *mutatis mutandis* applied to our problem.

2 Solving the (AS-HJB) equation

The equation (AS-HJB) we introduced in the preceding section is not a *simple* 4-variable PDE. Rather, because the variable q is discrete, it is an infinite system of 3-variable PDEs. To solve it, we will use a rather non-intuitive change of variables that is different from the one used in [5] and transform the system of PDEs in a system of linear ODEs.

Proposition 1 (A system of linear ODEs). *Let us consider a family of strictly positive functions $(v_q)_{q \in \mathbb{Z}}$ solution of the linear system of ODEs (\mathcal{S}) that follows:*

$$\forall q \in \mathbb{Z}, \dot{v}_q(t) = \alpha q^2 v_q(t) - \eta (v_{q-1}(t) + v_{q+1}(t))$$

with $v_q(T) = 1$, where $\alpha = \frac{k}{2} \gamma \sigma^2$ and $\eta = A(1 + \frac{\gamma}{k})^{-(1 + \frac{k}{\gamma})}$.

Then $u(t, x, q, s) = -\exp(-\gamma(x + qs)) v_q(t)^{-\frac{\gamma}{k}}$ is solution of (AS-HJB).

Theorem 1 (Well-posedness of the system (\mathcal{S})). *There exists a unique solution of (\mathcal{S}) in $C^\infty([0, T], \ell^2(\mathbb{Z}))$ and this solution consists in strictly positive functions.*

Theorem 2 (Expression of the maximizers). *Let us consider the solution v of the system (\mathcal{S}) as in Theorem 1. Let $u(t, x, q, s) = -\exp(-\gamma(x + qs))v_q(t)^{-\frac{\gamma}{k}}$ as above. Then, $\lambda^b(\delta^b) [u(t, x - s + \delta^b, q + 1, s) - u(t, x, q, s)]$ and $\lambda^a(\delta^a) [u(t, x + s + \delta^a, q - 1, s) - u(t, x, q, s)]$ are maximal for δ^{b*} and δ^{a*} given by:*

$$\delta^{b*}(t, q) = \frac{1}{k} \ln \left(\frac{v_q(t)}{v_{q+1}(t)} \right) + \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right)$$

$$\delta^{a*}(t, q) = \frac{1}{k} \ln \left(\frac{v_q(t)}{v_{q-1}(t)} \right) + \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right)$$

Now, in addition to these new results on the (AS-HJB) equation, we will provide an approximation of $\delta^{b*}(t, q)$ and $\delta^{a*}(t, q)$ in two steps. The first step is to notice that the above maximizers have an asymptotic behavior that can easily be characterized in a spectral way. Then, using this spectral characterization, we provide in a second step a closed-form approximation for these maximizers that will turn out to be quite accurate approximations for the optimal quotes in our initial control problem.

The asymptotic behavior of the maximizers is the following – asymptotic meaning here a time horizon T going to infinity:

Theorem 3 (Asymptotic behavior of the maximizers).

$$\forall q \in \mathbb{Z}, \quad \lim_{T \rightarrow \infty} \delta^{b*}(0, q) = \delta_\infty^{b*}(q)$$

$$\forall q \in \mathbb{Z}, \quad \lim_{T \rightarrow \infty} \delta^{a*}(0, q) = \delta_\infty^{a*}(q)$$

with

$$\delta_\infty^{b*}(q) = \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right) + \frac{1}{k} \ln \left(\frac{f_q^0}{f_{q+1}^0} \right) \quad \delta_\infty^{a*}(q) = \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right) + \frac{1}{k} \ln \left(\frac{f_q^0}{f_{q-1}^0} \right)$$

where $f^0 \in \ell^2(\mathbb{Z})$ is characterized by:

$$f^0 \in \operatorname{argmin}_{\|f\|_{\ell^2(\mathbb{Z})}=1} \sum_{q \in \mathbb{Z}} \alpha q^2 f_q^2 + \eta \sum_{q \in \mathbb{Z}} (f_{q+1} - f_q)^2$$

The challenge is now to provide an approximation for f^0 that happens to fit the actual figures and the approximation we propose is based on the continuous counterpart of f^0 , namely $\tilde{f}^0 \in L^2(\mathbb{R})$, a function that verifies:

$$\tilde{f}^0 \in \operatorname{argmin}_{\|\tilde{f}\|_{L^2(\mathbb{R})}=1} \int_{-\infty}^{\infty} \left(\alpha x^2 \tilde{f}(x)^2 + \eta \tilde{f}'(x)^2 \right) dx$$

It can be proved⁶ that such a function \tilde{f}^0 must be proportional to the probability distribution function of a normal variable with mean 0 and variance $\sqrt{\frac{\eta}{\alpha}}$. Hence, we expect f_q^0

⁶To prove this we need to proceed as in the proof of Theorems 1 and 3. In a few words, we introduce the positive, compact and self-adjoint operator L_c defined for $f \in L^2(\mathbb{R})$ as the unique weak solution v of $\alpha x^2 v - \eta v'' = f$ with $\int_{-\infty}^{\infty} (\alpha x^2 v(x)^2 + \eta v'(x)^2) dx < +\infty$. L_c can be diagonalized and largest eigenvalue of L_c can be shown to be associated to the eigenvector $f(x) = \exp\left(-\frac{1}{2}\sqrt{\frac{\alpha}{\eta}}x^2\right)$.

to behave as $\exp\left(-\frac{1}{2}\sqrt{\frac{\alpha}{\eta}}q^2\right)$.

This heuristic viewpoint induces an approximation for the maximizers if we replace f_q^0 by $\exp\left(-\frac{1}{2}\sqrt{\frac{\alpha}{\eta}}q^2\right)$:

$$\begin{aligned}\delta_\infty^{b^*}(q) &\simeq \frac{1}{\gamma} \ln\left(1 + \frac{\gamma}{k}\right) + \frac{1}{2k} \sqrt{\frac{\alpha}{\eta}}(2q+1) \\ &\simeq \frac{1}{\gamma} \ln\left(1 + \frac{\gamma}{k}\right) + \frac{2q+1}{2} \sqrt{\frac{\sigma^2\gamma}{2kA}} \left(1 + \frac{\gamma}{k}\right)^{1+\frac{k}{\gamma}} \\ \delta_\infty^{a^*}(q) &\simeq \frac{1}{\gamma} \ln\left(1 + \frac{\gamma}{k}\right) - \frac{1}{2k} \sqrt{\frac{\alpha}{\eta}}(2q-1) \\ &\simeq \frac{1}{\gamma} \ln\left(1 + \frac{\gamma}{k}\right) - \frac{2q-1}{2} \sqrt{\frac{\sigma^2\gamma}{2kA}} \left(1 + \frac{\gamma}{k}\right)^{1+\frac{k}{\gamma}}\end{aligned}$$

3 A solution to our control problem

As for the resolution of the (AS-HJB) equation, we can reduce the Hamilton-Jacobi-Bellman (HJB) associated to our control problem to a linear system of ODEs. However the linear system associated to this model will be simpler since it involves $2Q+1$ equations only and can be solved using linear algebra. Then, using a verification argument, we can prove that our solution to (HJB) is indeed the value function of our control problem and find the optimal quotes.

Proposition 2 (Resolution of (HJB)). *Let us consider a family $(v_q)_{|q|\leq Q}$ of strictly positive functions solution of:*

$$\begin{aligned}\forall q \in \{-Q+1, \dots, Q-1\}, \quad \dot{v}_q(t) &= \alpha q^2 v_q(t) - \eta(v_{q-1}(t) + v_{q+1}(t)) \\ \dot{v}_Q(t) &= \alpha Q^2 v_Q(t) - \eta v_{Q-1}(t) \\ \dot{v}_{-Q}(t) &= \alpha Q^2 v_{-Q}(t) - \eta v_{-Q+1}(t)\end{aligned}$$

with $\forall q \in \{-Q, \dots, Q\}, v_q(T) = 1$, where $\alpha = \frac{k}{2}\gamma\sigma^2$ and $\eta = A(1 + \frac{\gamma}{k})^{-(1+\frac{k}{\gamma})}$. Then, $u(t, x, q, s) = -\exp(-\gamma(x+qs))v_q(t)^{-\frac{2}{k}}$ is solution of (HJB).

Theorem 4 (Resolution of the control problem). *Let us introduce the matrix M defined by:*

$$M = \begin{pmatrix} \alpha Q^2 & -\eta & 0 & \cdots & \cdots & \cdots & 0 \\ -\eta & \alpha(Q-1)^2 & -\eta & 0 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 & -\eta & \alpha(Q-1)^2 & -\eta \\ 0 & \cdots & \cdots & \cdots & 0 & -\eta & \alpha Q^2 \end{pmatrix}$$

where $\alpha = \frac{k}{2}\gamma\sigma^2$ and $\eta = A(1 + \frac{\gamma}{k})^{-(1+\frac{k}{\gamma})}$.

Then, if

$$v(t) = (v_{-Q}(t), v_{-Q+1}(t), \dots, v_0(t), \dots, v_{Q-1}(t), v_Q(t))' = \exp(-M(T-t)) \times (1, \dots, 1)'$$

$u(t, x, q, s) = -\exp(-\gamma(x + qs))v_q(t)^{-\frac{\gamma}{k}}$ is the value function of the control problem and the optimal quotes are given by:

$$\forall q \in \{-Q, \dots, 0, \dots, Q-1\}, \quad s - s^{b*}(t, q, s) = \delta^{b*}(t, q) = \frac{1}{k} \ln \left(\frac{v_q(t)}{v_{q+1}(t)} \right) + \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right)$$

$$\forall q \in \{-(Q-1), \dots, 0, \dots, Q\}, \quad s^{a*}(t, q, s) - s = \delta^{a*}(t, q) = \frac{1}{k} \ln \left(\frac{v_q(t)}{v_{q-1}(t)} \right) + \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right)$$

and the bid-ask spread quoted by the market maker is given by:

$$\forall q \in \{-(Q-1), \dots, 0, \dots, Q-1\}, \quad \psi^*(t, q) = -\frac{1}{k} \ln \left(\frac{v_{q+1}(t)v_{q-1}(t)}{v_q(t)^2} \right) + \frac{2}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right)$$

To exemplify our findings and in order to justify the asymptotic approximation we shall provide, let us present some graphs to understand the behavior in time and inventory of the optimal quotes.

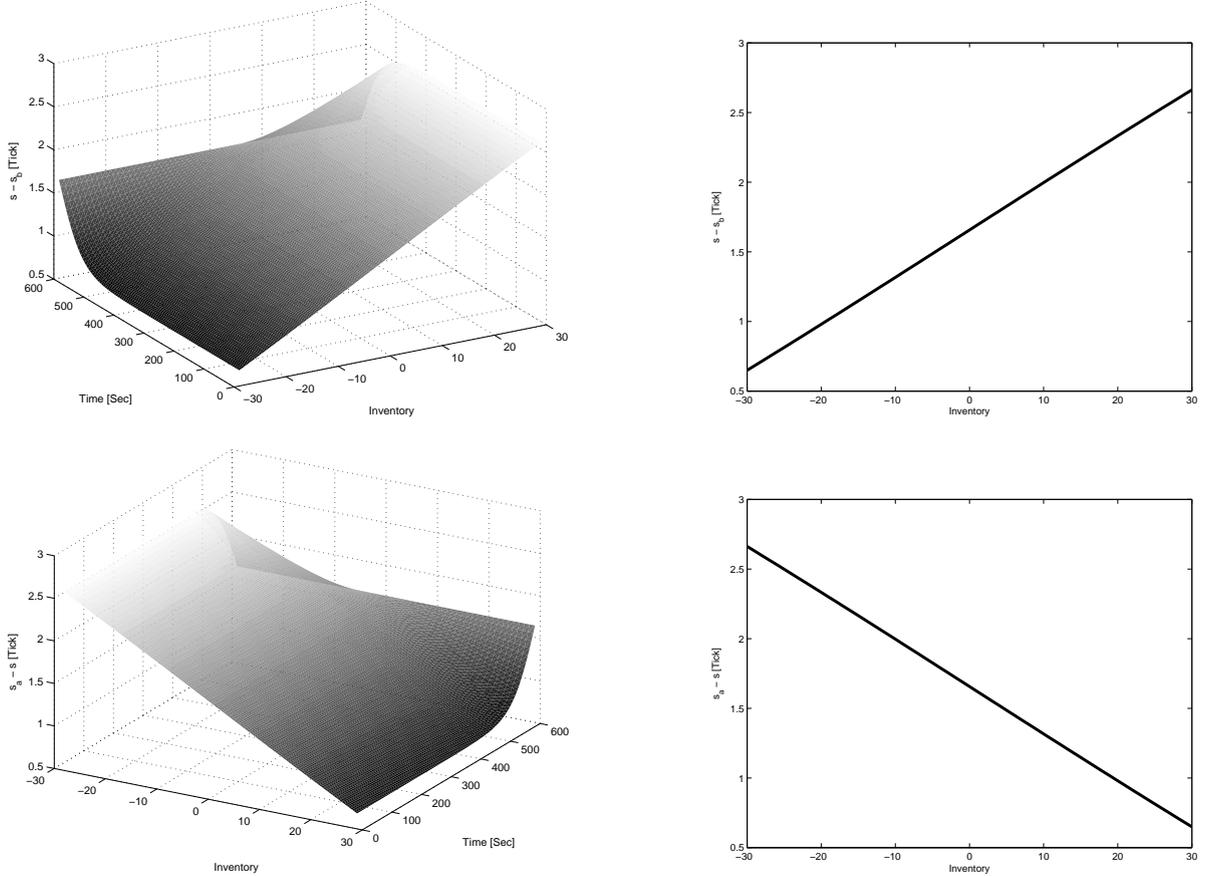


Figure 1: Left: Behavior of the optimal quotes with time and inventory. Right: Behavior of the optimal quotes with inventory, at time $t = 0$. $\sigma = 0.3 \text{ Tick} \cdot \text{s}^{-1/2}$, $A = 0.9 \text{ s}^{-1}$, $k = 0.3 \text{ Tick}^{-1}$, $\gamma = 0.01 \text{ Tick}^{-1}$, $T = 600 \text{ s}$.

We clearly see that the optimal quotes are, after a few minutes, very close to their asymptotic values and this justifies the approximation of the optimal quotes by their asymptotic limits obtained in the following theorem:

Theorem 5 (Asymptotics for the optimal quotes). *The optimal quotes have asymptotic limits*

$$\lim_{T \rightarrow \infty} \delta^{b*}(0, q) = \delta_{\infty}^{b*}(q)$$

$$\lim_{T \rightarrow \infty} \delta^{a*}(0, q) = \delta_{\infty}^{a*}(q)$$

that can be expressed as:

$$\delta_{\infty}^{b*}(q) = \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right) + \frac{1}{k} \ln \left(\frac{f_q^0}{f_{q+1}^0} \right) \quad \delta_{\infty}^{a*}(q) = \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right) + \frac{1}{k} \ln \left(\frac{f_q^0}{f_{q-1}^0} \right)$$

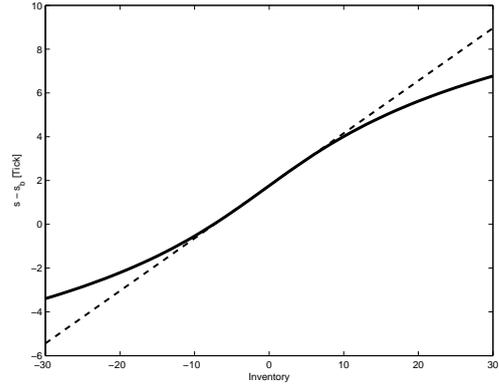
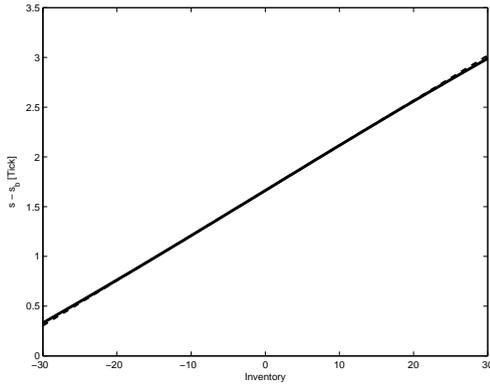
where $f^0 \in \mathbb{R}^{2Q+1}$ is an eigenvector corresponding to the smallest eigenvalue of the matrix M introduced in the preceding theorem.

The resulting bid-ask spread quoted by the market maker is asymptotically:

$$\psi_{\infty}^*(q) = -\frac{1}{k} \ln \left(\frac{f_{q+1}^0 f_{q-1}^0}{f_q^0{}^2} \right) + \frac{2}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right)$$

Now, these expressions for the limits depend on the spectral decomposition of the matrix M introduced earlier. In fact, and this is the very reason why we first solved the (AS-HJB) equation instead of heading directly to the Hamilton-Jacobi-Bellman equation of our problem, we are going to use the “gaussian” approximations introduced in the preceding section and this approximation will turn out to be quite accurate except when q (or Q) is very large⁷.

We exhibit below the values of the optimal quotes, along with their associated approximations⁸. Empirically, these approximations for the quotes are satisfactory in most cases and are always very good for small values of the inventory q . The apparent difficulty to approximate the bid-ask spread comes from the chosen scale (the bid-ask spread being almost uniform across values of the inventory).



⁷In fact, even though f^0 is certainly well approximated by the gaussian approximation, we cannot expect a very good fit for the quotes when q is large because we are approximating an expression that depends on ratios of the form $\frac{f_q^0}{f_{q+1}^0}$ or $\frac{f_q^0}{f_{q-1}^0}$.

⁸For the bid-ask spread the approximation is

$$\psi_{\infty}^*(q) \simeq \frac{2}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right) + \sqrt{\frac{\sigma^2 \gamma}{2kA} \left(1 + \frac{\gamma}{k} \right)^{1 + \frac{k}{\gamma}}}$$

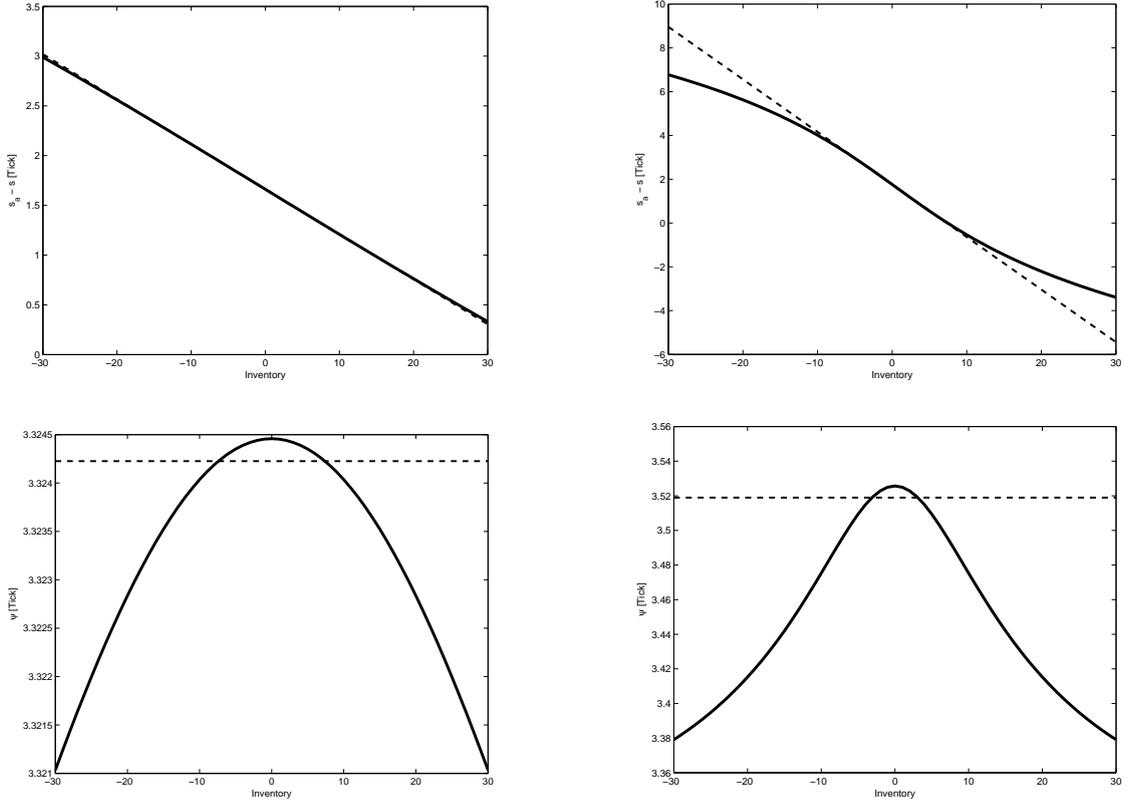


Figure 2: Asymptotic behavior of optimal quotes and the bid-ask spread (bold line). Approximation (dotted line). Left: $\sigma = 0.4 \text{ Tick} \cdot s^{-1/2}$, $A = 0.9 \text{ s}^{-1}$, $k = 0.3 \text{ Tick}^{-1}$, $\gamma = 0.01 \text{ Tick}^{-1}$, $T = 600 \text{ s}$. Right: $\sigma = 1.0 \text{ Tick} \cdot s^{-1/2}$, $A = 0.2 \text{ s}^{-1}$, $k = 0.3 \text{ Tick}^{-1}$, $\gamma = 0.01 \text{ Tick}^{-1}$, $T = 600 \text{ s}$.

4 Comparative statics

Before starting with the comparative statics, we rewrite the approximations used in the previous section to be able to have some intuition about the behavior of the optimal quotes and bid-ask spread with respect to the parameters:

$$\delta_{\infty}^{b*}(q) \simeq \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right) + \frac{2q+1}{2} \sqrt{\frac{\sigma^2 \gamma}{2kA} \left(1 + \frac{\gamma}{k} \right)^{1+\frac{k}{\gamma}}}$$

$$\delta_{\infty}^{a*}(q) \simeq \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right) - \frac{2q-1}{2} \sqrt{\frac{\sigma^2 \gamma}{2kA} \left(1 + \frac{\gamma}{k} \right)^{1+\frac{k}{\gamma}}}$$

$$\psi_{\infty}^*(q) \simeq \frac{2}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right) + \sqrt{\frac{\sigma^2 \gamma}{2kA} \left(1 + \frac{\gamma}{k} \right)^{1+\frac{k}{\gamma}}}$$

Now, from these approximations, we can “deduce” the behavior of the optimal quotes and the bid-ask spread with respect to price volatility, trading intensity and risk aversion.

4.1 Dependence on σ^2

From the above approximations we expect the dependence of optimal quotes on σ^2 to be a function of the inventory. More precisely, we expect:

$$\begin{cases} \frac{\partial \delta_{\infty}^{b*}}{\partial \sigma^2} < 0, & \frac{\partial \delta_{\infty}^{a*}}{\partial \sigma^2} > 0, & \text{if } q < 0 \\ \frac{\partial \delta_{\infty}^{b*}}{\partial \sigma^2} > 0, & \frac{\partial \delta_{\infty}^{a*}}{\partial \sigma^2} > 0, & \text{if } q = 0 \\ \frac{\partial \delta_{\infty}^{b*}}{\partial \sigma^2} > 0, & \frac{\partial \delta_{\infty}^{a*}}{\partial \sigma^2} < 0, & \text{if } q > 0 \end{cases}$$

For the bid-ask spread we expect it to be increasing with respect to σ^2 :

$$\frac{\partial \psi_{\infty}^*}{\partial \sigma^2} > 0$$

The rationale behind this is that a rise in σ^2 increases the inventory risk. Hence, to reduce this risk, a market maker that has a long position will try to reduce his exposure and hence ask less for his stocks (to get rid of some of them) and accept to buy at a cheaper price (to avoid buying new stocks). Similarly, a market maker with a short position tries to buy stocks, and hence increases its bid quote, while avoiding short selling new stocks, and he increases its ask quote to that purpose. Overall, due to the increase in risk, the bid-ask spread widens as it is well instanced in the case of a market maker with a flat position (this one wants indeed to earn more per trade to compensate the increase in inventory risk).

These intuitions can be verified numerically on Figure 3.

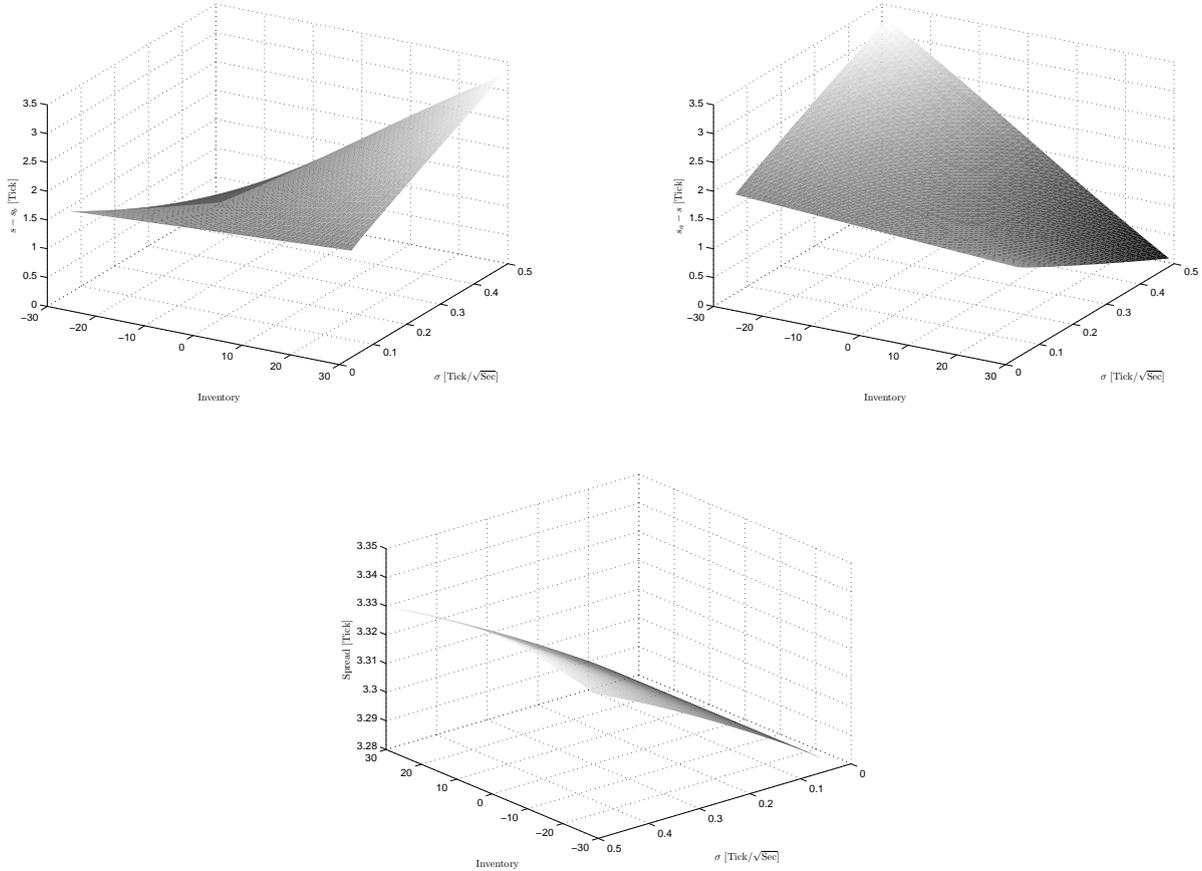


Figure 3: Asymptotic optimal quotes and bid-ask spread for different inventories and different values for the volatility σ . $A = 0.9 \text{ s}^{-1}$, $k = 0.3 \text{ Tick}^{-1}$, $\gamma = 0.01 \text{ Tick}^{-1}$, $T = 600 \text{ s}$.

4.2 Dependence on A

Because of the above approximations, and in accordance with the form of the system (\mathcal{S}), we expect the dependence on A to be the exact opposite of the dependence on σ^2 , namely

$$\begin{cases} \frac{\partial \delta_{\infty}^{b*}}{\partial A} > 0, & \frac{\partial \delta_{\infty}^{a*}}{\partial A} < 0, & \text{if } q < 0; \\ \frac{\partial \delta_{\infty}^{b*}}{\partial A} < 0, & \frac{\partial \delta_{\infty}^{a*}}{\partial A} < 0, & \text{if } q = 0 \\ \frac{\partial \delta_{\infty}^{b*}}{\partial A} < 0, & \frac{\partial \delta_{\infty}^{a*}}{\partial A} > 0, & \text{if } q > 0 \end{cases}$$

For the same reason, we expect the bid-ask spread to be decreasing with respect to A .

$$\frac{\partial \psi_{\infty}^*}{\partial A} < 0$$

The rationale behind these expectations is that an increase in A reduces the inventory risk. An increase in A indeed increases the frequency of trades and hence reduces the risk of being stuck with a large inventory (either positive or negative). For this reason, a rise in A should have the same effect as a decrease in σ^2 .

These intuitions can be verified numerically on Figure 4.

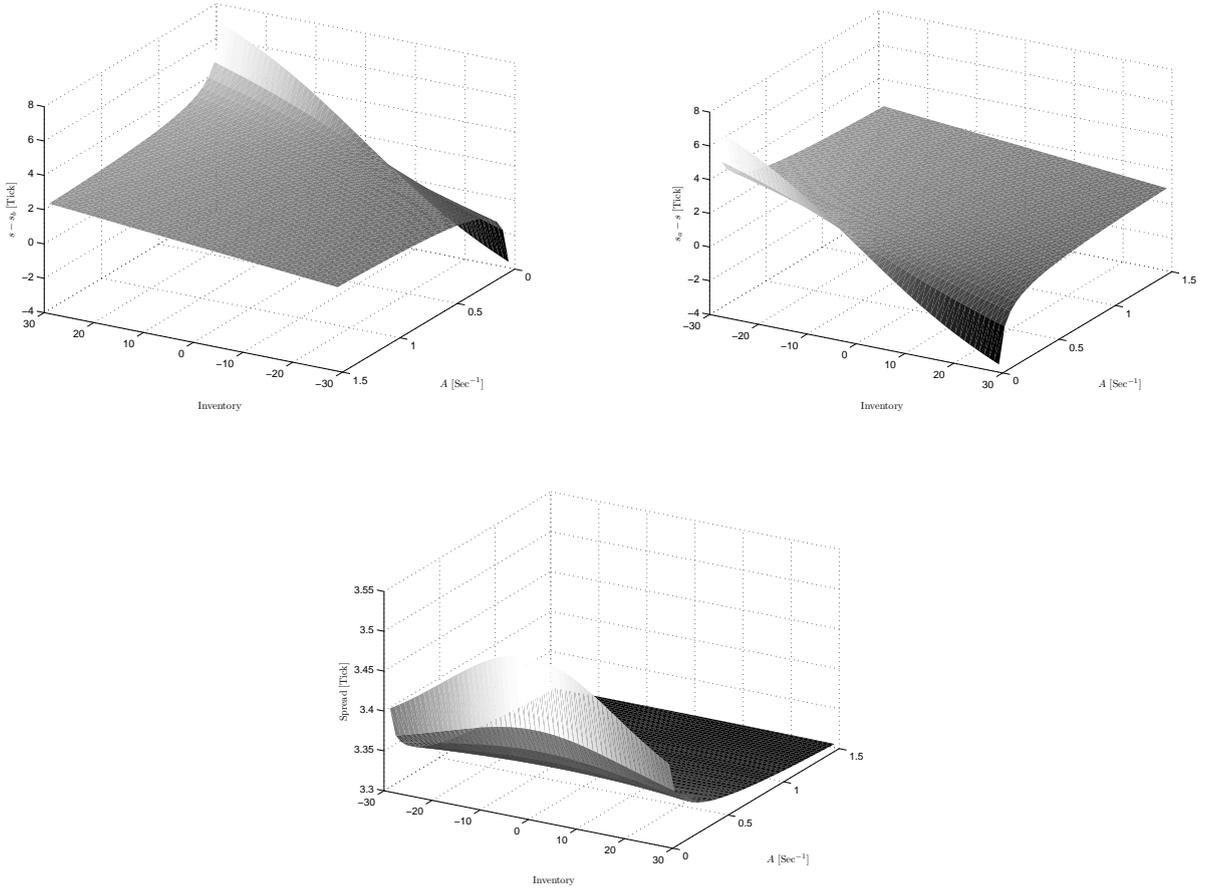


Figure 4: Asymptotic optimal quotes and bid-ask spread for different inventories and different values of A . $\sigma = 0.3 \text{ Tick} \cdot \text{s}^{-1/2}$, $k = 0.3 \text{ Tick}^{-1}$, $\gamma = 0.01 \text{ Tick}^{-1}$, $T = 600 \text{ s}$.

4.3 Dependence on k

From the above approximations we expect δ_{∞}^{b*} to be decreasing in k for q greater than some negative threshold. Below this threshold we expect it to be increasing. Similarly we expect δ_{∞}^{a*} to be decreasing in k for q smaller than some positive threshold. Above this threshold we expect it to be increasing.

Eventually, as far as the bid-ask spread is concerned, the above approximation indicates that the bid-ask spread should be a decreasing function of k .

$$\frac{\partial \psi_{\infty}^*}{\partial k} < 0$$

In fact several effects are in interaction. On one hand, there is a “no-volatility” effect that is completely orthogonal to any reasoning on the inventory risk: when k increases, in a situation where δ^b and δ^a are positive, trades occur closer to the price S_t . For this reason, and in absence of inventory risk, the optimal bid-ask spread has to shrink. However, an increase in k also affects the inventory risk since it decreases the probability to be executed (for $\delta^b, \delta^a > 0$). Hence, an increase in k is also, in some aspects, similar to a decrease in A . These two effects explain the expected behavior.

Numerically, one of two effects dominates for the values of the inventory under consideration:

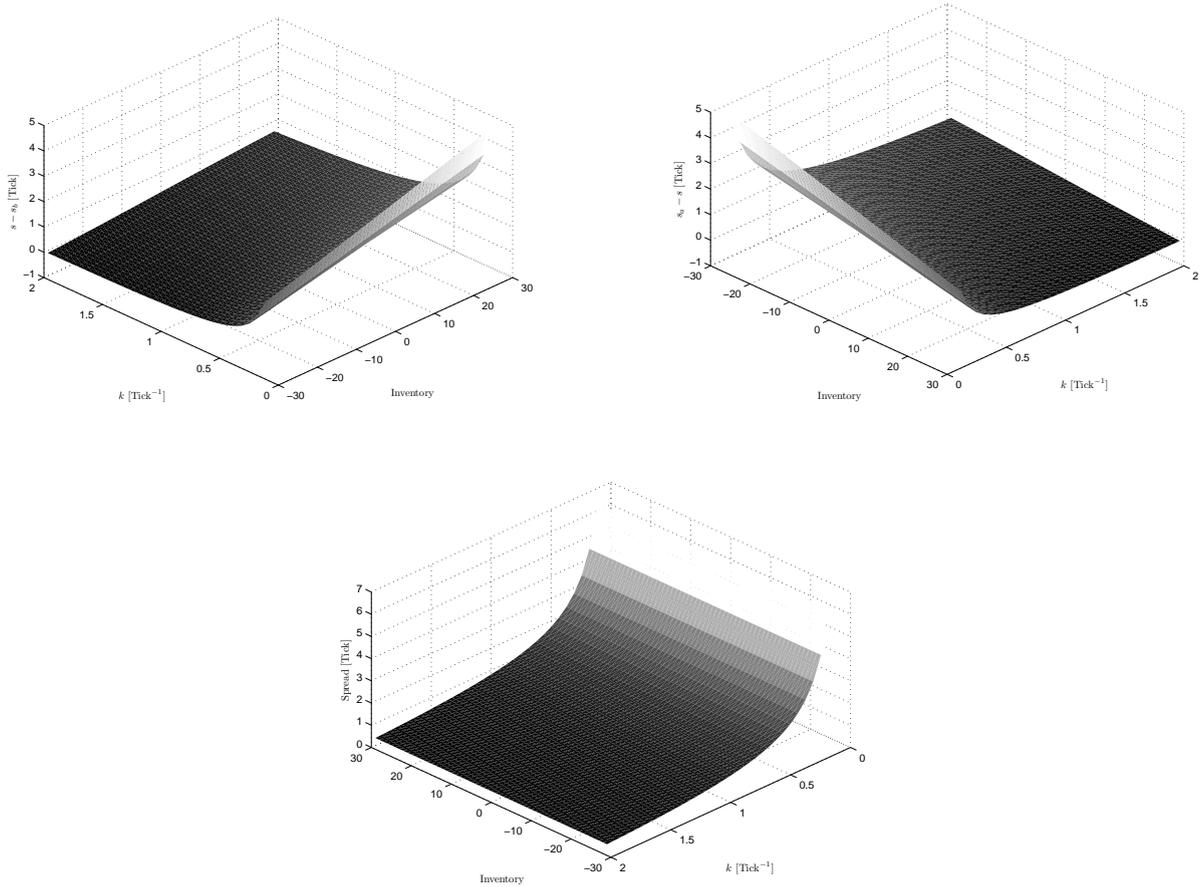


Figure 5: Asymptotic optimal quotes and bid-ask spread for different inventories and different values of k . $\sigma = 0.3 \text{ Tick} \cdot \text{s}^{-1/2}$, $A = 0.9 \text{ s}^{-1}$, $\gamma = 0.01 \text{ Tick}^{-1}$, $T = 600 \text{ s}$.

4.4 Dependence on γ

Using the above approximations, we see that the dependence on γ is ambiguous. The market maker faces two different risks that contribute to the inventory risk: (i) trades occur at random times and (ii) the fair price is stochastic. But if risk aversion increases, the market maker will mitigate the two risks: (i) he may set his quotes closer to one another to reduce the randomness in execution and (ii) he may enlarge his spread to reduce price risk. The

tension between these two roles played by γ explains the different behaviors we may observe, as in the figures below:

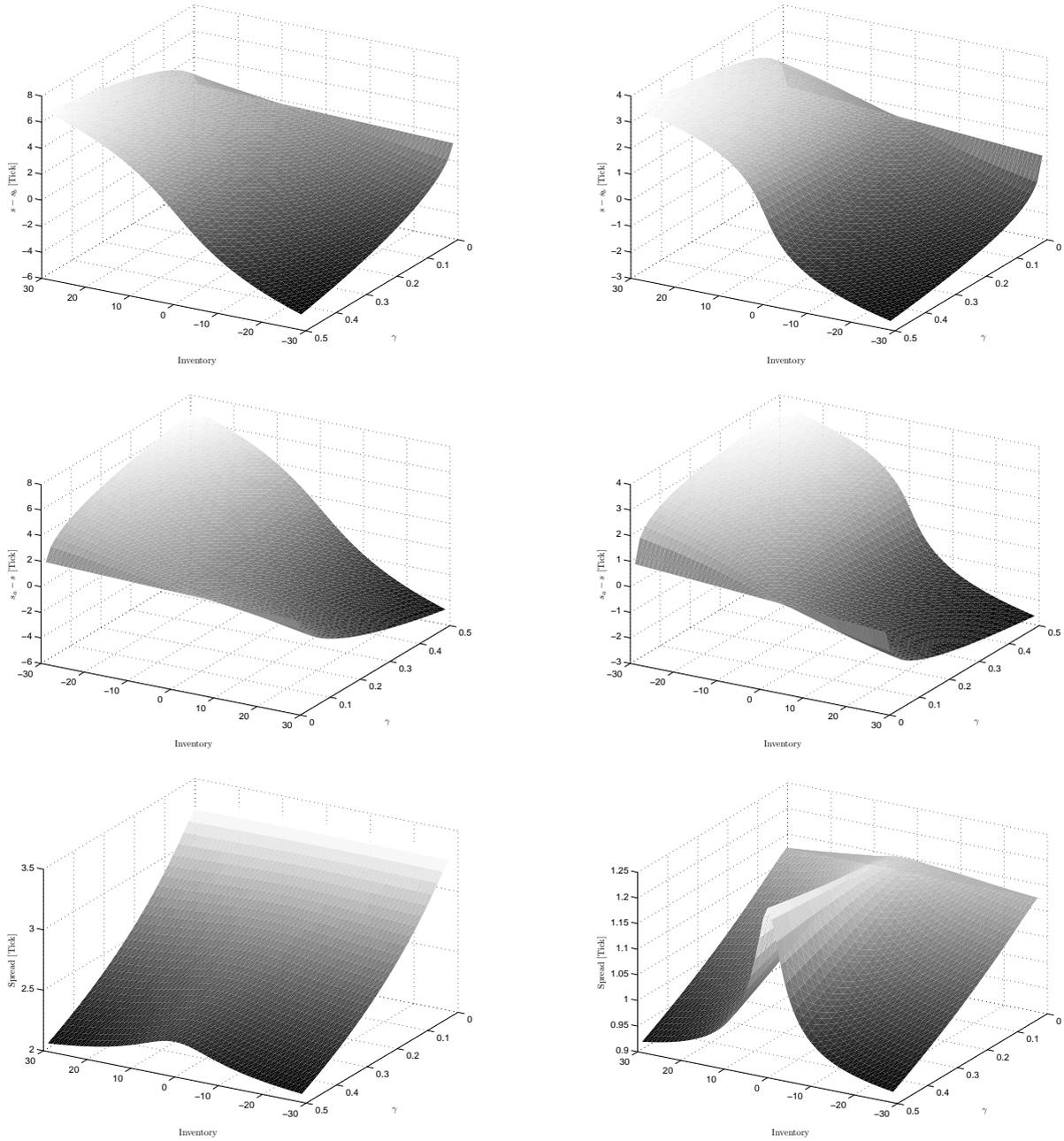


Figure 6: Asymptotic optimal quotes and bid-ask spread for different inventories and different values for the risk aversion parameter γ . Left: $\sigma = 0.3 \text{ Tick} \cdot \text{s}^{-1/2}$, $A = 0.9 \text{ s}^{-1}$, $k = 0.3 \text{ Tick}^{-1}$, $T = 600 \text{ s}$. Right: $\sigma = 0.6 \text{ Tick} \cdot \text{s}^{-1/2}$, $A = 0.9 \text{ s}^{-1}$, $k = 0.9 \text{ Tick}^{-1}$, $T = 600 \text{ s}$

5 Different settings

In what follows we provide the settings of several variants of the initial model. We will alternatively consider a model with a trend in prices and a model with a penalization term for not having cleared one's inventory.

For each model, we enounce the associated results and the general proofs are not repeated since they can easily be derived from adaptations of the proofs of the initial model.

5.1 Trend in prices

In the preceding setting, we supposed that the fair price of the stock followed a brownian motion. However, we can also build a model in presence of a trend:

$$dS_t = \mu dt + \sigma dW_t$$

In that case we have the following proposition:

Proposition 3 (Resolution with drift). *Let us consider a family of functions $(v_q)_{|q| \leq Q}$ solution of the linear system of ODEs that follows:*

$$\forall q \in \{-Q+1, \dots, Q-1\}, \dot{v}_q(t) = (\alpha q^2 - \beta q)v_q(t) - \eta(v_{q-1}(t) + v_{q+1}(t))$$

$$\dot{v}_Q(t) = (\alpha Q^2 - \beta Q)v_Q(t) - \eta v_{Q-1}(t)$$

$$\dot{v}_{-Q}(t) = (\alpha Q^2 + \beta Q)v_{-Q}(t) - \eta v_{-Q+1}(t)$$

with $\forall q \in \{-Q, \dots, Q\}, v_q(T) = 1$, where $\alpha = \frac{k}{2}\gamma\sigma^2$, $\beta = k\mu$ and $\eta = A(1 + \frac{\gamma}{k})^{-(1+\frac{k}{\gamma})}$.

Then, optimal quotes are given by:

$$s - s^{b*}(t, q, s) = \delta^{b*}(t, q) = \frac{1}{k} \ln \left(\frac{v_q(t)}{v_{q+1}(t)} \right) + \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right)$$

$$s^{a*}(t, q, s) - s = \delta^{a*}(t, q) = \frac{1}{k} \ln \left(\frac{v_q(t)}{v_{q-1}(t)} \right) + \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right)$$

and the bid-ask spread quoted by the market maker is :

$$\psi^*(t, q) = -\frac{1}{k} \ln \left(\frac{v_{q+1}(t)v_{q-1}(t)}{v_q(t)^2} \right) + \frac{2}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right)$$

Moreover,

$$\lim_{T \rightarrow \infty} \delta^{b*}(0, q) = \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right) + \frac{1}{k} \ln \left(\frac{f_q^0}{f_{q+1}^0} \right)$$

$$\lim_{T \rightarrow \infty} \delta^{a*}(0, q) = \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right) + \frac{1}{k} \ln \left(\frac{f_q^0}{f_{q-1}^0} \right)$$

$$\lim_{T \rightarrow \infty} \psi^*(0, q) = -\frac{1}{k} \ln \left(\frac{f_{q+1}^0 f_{q-1}^0}{f_q^0{}^2} \right) + \frac{2}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right)$$

where f^0 is an eigenvector corresponding to the smallest eigenvalue of:

$$\begin{pmatrix} \alpha Q^2 - \beta Q & -\eta & 0 & \cdots & \cdots & \cdots & 0 \\ -\eta & \alpha(Q-1)^2 - \beta(Q-1) & -\eta & 0 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 & -\eta & \alpha(Q-1)^2 - \beta(Q-1) & -\eta \\ 0 & \cdots & \cdots & \cdots & 0 & -\eta & \alpha Q^2 - \beta Q \end{pmatrix}$$

Using the same continuous approximation as in the initial model we find the following approximations for the optimal quotes and the bid-ask spread:

$$\begin{aligned}\delta_\infty^{b*}(q) &\simeq \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k}\right) + \left[-\frac{\mu}{\gamma\sigma^2} + \frac{2q+1}{2}\right] \sqrt{\frac{\sigma^2\gamma}{2kA} \left(1 + \frac{\gamma}{k}\right)^{1+\frac{k}{\gamma}}} \\ \delta_\infty^{a*}(q) &\simeq \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k}\right) + \left[\frac{\mu}{\gamma\sigma^2} - \frac{2q-1}{2}\right] \sqrt{\frac{\sigma^2\gamma}{2kA} \left(1 + \frac{\gamma}{k}\right)^{1+\frac{k}{\gamma}}} \\ \psi_\infty^*(q) &\simeq \frac{2}{\gamma} \ln \left(1 + \frac{\gamma}{k}\right) + \sqrt{\frac{\sigma^2\gamma}{2kA} \left(1 + \frac{\gamma}{k}\right)^{1+\frac{k}{\gamma}}}\end{aligned}$$

5.2 Costly inventory liquidation

In the initial model we imposed a terminal condition based on the assumption that the market maker liquidates his inventory at the fair price at time $t = T$. This hypothesis is questionable and we propose to introduce an additional term to model liquidation cost that can also be interpreted as a penalization term for having a non-zero inventory at time T .

From a mathematical perspective it means that the control problem is now:

$$\sup_{(\delta_t^a)_t, (\delta_t^b)_t \in \mathcal{A}} \mathbb{E}[-\exp(-\gamma(X_T + q_T S_T - c|q_T|))]$$

where $c|q|$ represents the penalization term modeling the incurred cost at the end of the period for not having cleared the inventory⁹.

The analysis can then be done in the same way as in the initial model and we get the following result:

Proposition 4 (Resolution with inventory liquidation cost). *Let us consider a family of functions $(v_q)_{|q| \leq Q}$ solution of the linear system of ODEs that follows:*

$$\begin{aligned}\forall q \in \{-Q+1, \dots, Q-1\}, \dot{v}_q(t) &= \alpha q^2 v_q(t) - \eta(v_{q-1}(t) + v_{q+1}(t)) \\ \dot{v}_Q(t) &= \alpha Q^2 v_Q(t) - \eta v_{Q-1}(t) \\ \dot{v}_{-Q}(t) &= \alpha Q^2 v_{-Q}(t) - \eta v_{-Q+1}(t)\end{aligned}$$

with $\forall q \in \{-Q, \dots, Q\}, v_q(T) = e^{-kc|q|}$, where $\alpha = \frac{k}{2}\gamma\sigma^2$ and $\eta = A(1 + \frac{\gamma}{k})^{-(1+\frac{k}{\gamma})}$.

Then, optimal quotes are given by:

$$\begin{aligned}s - s^{b*}(t, q, s) = \delta^{b*}(t, q) &= \frac{1}{k} \ln \left(\frac{v_q(t)}{v_{q+1}(t)} \right) + \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right) \\ s^{a*}(t, q, s) - s = \delta^{a*}(t, q) &= \frac{1}{k} \ln \left(\frac{v_q(t)}{v_{q-1}(t)} \right) + \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right)\end{aligned}$$

and the bid-ask spread quoted by the market maker is :

$$\psi^*(t, q) = -\frac{1}{k} \ln \left(\frac{v_{q+1}(t)v_{q-1}(t)}{v_q(t)^2} \right) + \frac{2}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right)$$

Moreover, the asymptotic behavior of the optimal quotes and the bid-ask spread does not depend on the liquidation cost coefficient c and is the same as in the initial model.

⁹ c is a positive number that represents the cost per share in the inventory at time T .

6 Discussion on the model

6.1 Exogenous nature of prices

In our model, a fair price is introduced and modeled by a brownian motion independent of the behavior of the agent. Since we are modeling a single market maker who operates through limit orders, it seems natural to consider the price process exogenous in the medium run. However, even if we neglect the impact of our market maker on the market, the very notion of fair price must be clarified. The fair price we consider may be thought of as a smoothed version of the current mid-price of the market. It is not the current mid-price because it would in that case depend on the actions of the market maker. It is neither a medium-run fair value for the stock (or fundamental value). The only important characteristic, that in fact somehow defines S , is that the execution probabilities must be well described as depending on the distance to the fair price we consider.

Another issue linked to the exogenous nature of prices is the absence of market impact or more generally the independence between the Poisson processes and the brownian motion. This issue is a deep one and, for a part of it, it points out the lack of market impact models for passive orders in the financial literature.

6.2 Constant parameters

A related issue is that the parameters σ , A and k are constant. While models can be developed to take account of deterministic or stochastic variations of the parameters, the most important point is to take account of the links between the different parameters. σ , A and k should not indeed be considered independent of one another since, for instance, an increase in A should induce an increase in the number of trades and hence an increase in price volatility.

Some attempts have been made in this direction to model the link between volatility and trades intensity. Hawkes processes (see [6]) for instance may provide good modeling perspectives to link the parameters but this has been left aside for future work.

6.3 Constant size of orders

Another apparent issue of the model is that market makers set orders of size 1 at all times. A first remark is that, if this is an issue, it is only limited to the fact that orders are of constant size since we can consider that the unitary orders stand for orders of constant size δq or equivalently, though more abstractly, orders of size 1 on a bunch of δq stocks.

If all the orders are of size δq then the stochastic process representing cash is:

$$d\tilde{X}_t = \delta q(S_t^a dN_t^a - S_t^b dN_t^b) = \delta q \times dX_t$$

where the jump processes model the event of being hit by an aggressive order (of size δq). Then, if we consider that an order of size δq is a unitary order on a bunch of δq stocks, we can write $\tilde{q}_t = \delta q \times q_t$ and the optimization criterion becomes:

$$\sup_{(\delta_t^a)_t, (\delta_t^b)_t \in \mathcal{A}} \mathbb{E} \left[-\exp \left(-\gamma(\tilde{X}_T + \tilde{q}_T S_T) \right) \right] = \sup_{(\delta_t^a)_t, (\delta_t^b)_t \in \mathcal{A}} \mathbb{E} \left[-\exp \left(-\gamma \delta q (X_T + q_T S_T) \right) \right]$$

Hence, we can solve the problem for orders of size δq using a modified risk aversion, namely solving the problem for unitary orders, with γ multiplied by δq .

However, if we can transform the problem with orders of constant size δq into the initial problem where $\delta q = 1$, the parameters must be adjusted in accordance with the fact that orders are of size δq . As such, it must be noticed that A has to be estimated to take account of the expected proportion of an order of size δq filled by a single trade and approximations must be made to take account of market making with orders of constant size. In fact, we can consider that, after each trade that partially filled the order, the market maker sends a new order so that the total size of his orders is δq , using a convex combination of the model recommendations for the price¹⁰, since in that case the inventory is not a multiple of δq .

In our view, this issue is important but it should not be considered a problem to describe qualitative market maker's behavior and appropriate approximation on A allows us to believe that the error made, as far as quantitative modeling results are concerned, is relatively small.

7 Applications

In spite of the limitations discussed above we used this model to backtest the strategy on real data. We rapidly discuss the change that had to be made to the model to use it in reality and the way backtests have been carried out. Then we present an instance of the results on the French stock AXA.

7.1 Empirical use

Before using the above model in reality, we need to discuss some features of the model that need to be adapted before any backtest is possible.

First of all, the model is continuous in both time and space while the real control problem under scrutiny is intrinsically discrete in space, because of the tick size, and in time, because orders have a certain priority and changing position too often reduces the actual chance to be reached by a market order. Hence, the model has to be reinterpreted in a discrete way. In terms of prices, quotes must not be between two ticks and we decided to ceil or floor the optimal quotes with probabilities that depend on the respective proximity to the neighboring quotes. In terms of time, an order is sent to the market and is not canceled nor modified for a given period Δt (say 20 or 60 seconds), unless a trade occurs and, though perhaps partially, fills one of the market maker's orders. Now, when a trade occurs and changes the inventory or when an order stayed in the order book for longer than Δt , then the optimal quotes on both sides are updated and, if necessary, new orders are inserted.

Now, concerning the parameters, σ , A and k can be calibrated on trade-by-trade limit order book data while γ has to be chosen. However, it is well known by practitioners that A and k have to depend at least on the actual market bid-ask spread. Since we do not explicitly take into account the underlying market, there is no market bid-ask spread in the model. Thus, we simply chose to calibrate k and A as functions of the market bid-ask spread, making then an off-model hypothesis.

Turning to the backtests, they were carried out with trade-by-trade data and we assumed that our orders were entirely filled when a trade occurred above (resp. below) the ask (resp. bid) price quoted by the market maker.

¹⁰If the optimal quote changed, the remaining part of the order is canceled before a new order is inserted. Otherwise, the market maker may just insert a new order so that the cumulated size of his orders is δq .

7.2 Results

To present the results, we chose to illustrate the case of the French stock AXA on November 2nd 2010.

We first show the evolution of the inventory and we see that this inventory mean-reverts around 0.

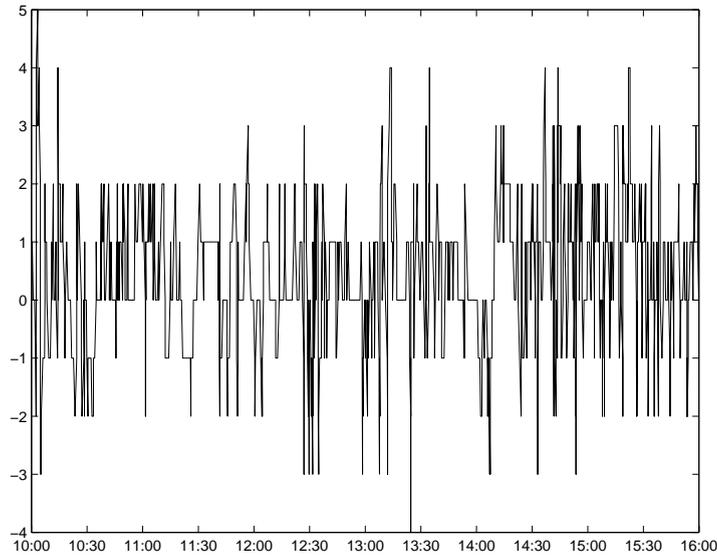


Figure 7: Inventory when the strategy is used on AXA (02/11/2010) from 10:00 to 16:00 with $\gamma = 0.05 \text{ Tick}^{-1}$

Now, to better understand the very nature of the strategy, we focused on a subperiod of 20 minutes and we plotted the state of the market both with the quotes of the market maker. Trades occurrences involving the market maker are signalled by a dot.

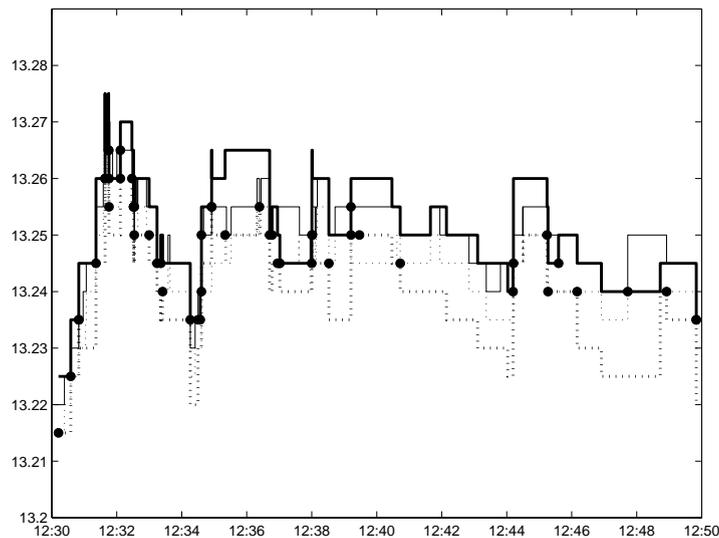


Figure 8: Details for the quotes and trades when the strategy is used on AXA (02/11/2010) with $\gamma = 0.05 \text{ Tick}^{-1}$. Thin lines represent the market while bold lines represent the quotes of the market maker. Dotted lines are associated to the bid side while plain lines are associated to the ask side. Black points represent trades in which the market maker is involved.

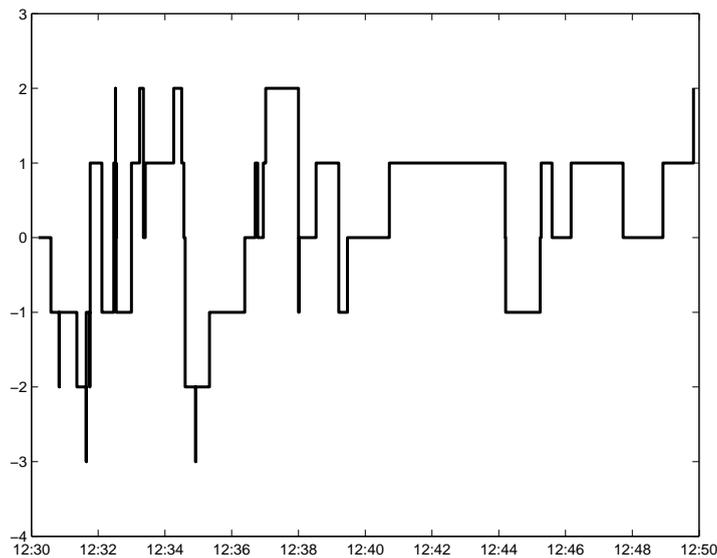


Figure 9: Details for the inventory when the strategy is used on AXA (02/11/2010) with $\gamma = 0.05 \text{ Tick}^{-1}$

Conclusion

In this paper we presented a model for the optimal quotes of a market maker. Starting from a model in line with Avellaneda and Stoikov [5] and rooted to Ho and Stoll [18], we introduced a change of variables¹¹ that allows to find exact expressions for the quotes. Then, we exhibited the asymptotic value of the optimal quotes and argued that the asymptotic values were very good approximations for the quotes even for rather small times. Closed-form approximations were then obtained using spectral arguments. The model is finally backtested on real data and the results are promising.

¹¹In a companion paper (see [14]) we used a change of variables similar to the one introduced above to solve the Hamilton-Jacobi-Bellman equation associated to an optimal execution problem with passive orders.

Appendix

Proof of Proposition 1 and Theorem 2:

Let us consider a strictly positive solution $(v_q)_q$ of (\mathcal{S}) and introduce $u(t, x, q, s) = -\exp(-\gamma(x + qs))v_q(t)^{-\frac{\gamma}{k}}$.

Then:

$$\partial_t u + \frac{1}{2}\sigma^2 \partial_{ss}^2 u = -\frac{\gamma \dot{v}_q(t)}{k v_q(t)} u + \frac{\gamma^2 \sigma^2}{2} q^2 u$$

Now, concerning the first part of the hamiltonian, we have:

$$\begin{aligned} & \sup_{\delta^b} \lambda^b(\delta^b) \left[u(t, x - s + \delta^b, q + 1, s) - u(t, x, q, s) \right] \\ &= \sup_{\delta^b} A e^{-k\delta^b} u(t, x, q, s) \left[\exp(-\gamma\delta^b) \left(\frac{v_{q+1}(t)}{v_q(t)} \right)^{-\frac{\gamma}{k}} - 1 \right] \end{aligned}$$

The first order condition of this problem corresponds to a maximum (because u is negative) and writes:

$$(k + \gamma) \exp(-\gamma\delta^{b*}) \left(\frac{v_{q+1}(t)}{v_q(t)} \right)^{-\frac{\gamma}{k}} = k$$

Hence:

$$\delta^{b*} = \frac{1}{k} \ln \left(\frac{v_q(t)}{v_{q+1}(t)} \right) + \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right)$$

and

$$\begin{aligned} & \sup_{\delta^b} \lambda^b(\delta^b) \left[u(t, x - s + \delta^b, q + 1, s) - u(t, x, q, s) \right] \\ &= -\frac{\gamma}{k + \gamma} A \exp(-k\delta^{b*}) u(t, x, q, s) \\ &= -\frac{\gamma A}{k + \gamma} \left(1 + \frac{\gamma}{k} \right)^{-\frac{k}{\gamma}} \frac{v_{q+1}(t)}{v_q(t)} u(t, x, q, s) \end{aligned}$$

Similarly, the maximizer for the second part of the hamiltonian is:

$$\delta^{a*} = \frac{1}{k} \ln \left(\frac{v_q(t)}{v_{q-1}(t)} \right) + \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right)$$

and

$$\begin{aligned} & \sup_{\delta^a} \lambda^a(\delta^a) \left[u(t, x + s + \delta^a, q - 1, s) - u(t, x, q, s) \right] \\ &= -\frac{\gamma}{k + \gamma} A \exp(-k\delta^{a*}) u(t, x, q, s) \\ &= -\frac{\gamma A}{k + \gamma} \left(1 + \frac{\gamma}{k} \right)^{-\frac{k}{\gamma}} \frac{v_{q-1}(t)}{v_q(t)} u(t, x, q, s) \end{aligned}$$

Hence, putting the three terms together we get:

$$\begin{aligned} & \partial_t u(t, x, q, s) + \frac{1}{2}\sigma^2 \partial_{ss}^2 u(t, x, q, s) \\ &+ \sup_{\delta^b} \lambda^b(\delta^b) \left[u(t, x - s + \delta^b, q + 1, s) - u(t, x, q, s) \right] \end{aligned}$$

$$\begin{aligned}
& + \sup_{\delta^a} \lambda^a(\delta^a) [u(t, x + s + \delta^a, q - 1, s) - u(t, x, q, s)] \\
& = -\frac{\gamma \dot{v}_q(t)}{k v_q(t)} u + \frac{\gamma^2 \sigma^2}{2} q^2 u - \frac{\gamma A}{k + \gamma} \left(1 + \frac{\gamma}{k}\right)^{\frac{k}{\gamma}} \left[\frac{v_{q+1}(t)}{v_q(t)} + \frac{v_{q-1}(t)}{v_q(t)} \right] u \\
& = -\frac{\gamma}{k} \frac{u}{v_q(t)} \left[\dot{v}_q(t) - \frac{k \gamma \sigma^2}{2} q^2 v_q(t) + A \left(1 + \frac{\gamma}{k}\right)^{-\left(1 + \frac{k}{\gamma}\right)} (v_{q+1}(t) + v_{q-1}(t)) \right] = 0
\end{aligned}$$

Now, noticing that the terminal condition for v_q is consistent with the terminal condition for u , we get that u verifies (HJB). \square

Proof of Theorem 1 and Theorem 3:

Before starting the very proof, let us introduce the necessary functional framework.

Let us introduce $H = \left\{ u \in \ell^2(\mathbb{Z}), \sum_{q \in \mathbb{Z}} \alpha q^2 u_q^2 + \eta(u_{q+1} - u_q)^2 < +\infty \right\}$. H is a Hilbert space equipped with the scalar product:

$$\langle v, w \rangle_H = \sum_{q \in \mathbb{Z}} \alpha q^2 v_q w_q + \eta(v_{q+1} - v_q)(w_{q+1} - w_q), \forall v, w \in H$$

The first preliminary lemma indicates that the $\ell^2(\mathbb{Z})$ -norm can be controlled by the H -norm.

Lemma 1.

$$\exists C > 0, \forall w \in H, \|w\|_{\ell^2(\mathbb{Z})} \leq C \|w\|_H$$

Proof:

$$\begin{aligned}
\forall w \in H, \|w\|_{\ell^2(\mathbb{Z})}^2 & = |w_0|^2 + \sum_{q \in \mathbb{Z}^*} |w_q|^2 \leq |w_0|^2 + \frac{1}{\alpha} \|w\|_H^2 \\
& \leq (|w_1| + |w_0 - w_1|)^2 + \frac{1}{\alpha} \|w\|_H^2 \leq \left[\left(\sqrt{\frac{1}{\alpha}} + \sqrt{\frac{1}{\eta}} \right)^2 + \frac{1}{\alpha} \right] \|w\|_H^2
\end{aligned}$$

\square

A second result that is central in the proof of our results is that H is compactly embedded in $\ell^2(\mathbb{Z})$.

Lemma 2. H is compactly embedded in $\ell^2(\mathbb{Z})$

Proof:

To prove the result we consider a bounded sequence $(s^k)_{k \in \mathbb{N}}$ in $H^{\mathbb{N}}$. For each $k \in \mathbb{N}$ we introduce \hat{s}^k defined by:

$$\hat{s}_0^k = s_0^k, \quad \forall q \in \mathbb{Z}^*, \hat{s}_q^k = q s_q^k$$

For $(s^k)_{k \in \mathbb{N}}$ is a bounded sequence in $H^{\mathbb{N}}$, $(\hat{s}^k)_{k \in \mathbb{N}}$ is a bounded sequence in $\ell^2(\mathbb{Z})^{\mathbb{N}}$. Hence, we can find $\hat{s}^\infty \in \ell^2(\mathbb{Z})$ so that there exists a subsequence indexed by $(k_j)_{j \in \mathbb{N}}$ such that $(\hat{s}^{k_j})_{j \in \mathbb{N}}$ weakly converges toward \hat{s}^∞ in $\ell^2(\mathbb{Z})$.

Now, we can define $s^\infty \in H$ by the inverse transformation:

$$s_0^\infty = \hat{s}_0^\infty, \quad \forall q \in \mathbb{Z}^*, s_q^\infty = \frac{1}{q} \hat{s}_q^\infty$$

and it is easy to check that $(s^{k_j})_{\mathbb{N}}$ converges in the $\ell^2(\mathbb{Z})$ sense toward s^∞ . Indeed,

$$\begin{aligned} \|s^{k_j} - s^\infty\|_{\ell^2(\mathbb{Z})}^2 &\leq |s_0^{k_j} - s_0^\infty|^2 + \sum_{q \in \mathbb{Z}^*} \frac{1}{q^2} |s_q^{k_j} - s_q^\infty|^2 \\ &= |s_0^{k_j} - s_0^\infty|^2 + \sum_{q \neq 0, |q| \leq N} \frac{1}{q^2} |s_q^{k_j} - s_q^\infty|^2 + \sum_{|q| > N} \frac{1}{q^2} |s_q^{k_j} - s_q^\infty|^2 \end{aligned}$$

The first two terms tend to 0 because $(s^{k_j})_{\mathbb{N}}$ weakly converges in $\ell^2(\mathbb{Z})$ towards s^∞ . The last one can be made smaller than any $\epsilon > 0$ as N becomes large because $\sum_{|q| > N} \frac{1}{q^2}$ tends to zero as N tends to infinity and $(\|s^{k_j} - s^\infty\|_{\ell^2(\mathbb{N})})_{j \in \mathbb{N}}$ is a bounded sequence. \square

Now, we are going to consider a linear operator L that is linked to the system (\mathcal{S}) . L is defined by

$$L : f \in \ell^2(\mathbb{Z}) \mapsto v \in H \subset \ell^2(\mathbb{Z})$$

where

$$\forall q \in \mathbb{Z}, \quad \alpha q^2 v_q - \eta(v_{q+1} - 2v_q + v_{q-1}) = f_q$$

We need to prove that L is well-defined and we use the weak formulation of the equation.

Lemma 3. *L is a well-defined linear (continuous) operator.*

Moreover $\forall f \in \ell^2(\mathbb{Z}), \forall w \in H, \langle Lf, w \rangle_H = \langle f, w \rangle_{\ell^2(\mathbb{Z})}$

Proof:

Let us consider $f \in \ell^2(\mathbb{Z})$.

Because of Lemma 1, $w \in H \mapsto \langle f, w \rangle_{\ell^2(\mathbb{Z})}$ is continuous. Hence, by Riesz representation Theorem there exists a unique $v \in H$ such that $\forall w \in H, \langle v, w \rangle_H = \langle f, w \rangle_{\ell^2(\mathbb{Z})}$.

This equation writes

$$\begin{aligned} \forall w \in H, \sum_{q \in \mathbb{Z}} f_q w_q &= \sum_{q \in \mathbb{Z}} \alpha q^2 v_q w_q + \eta(v_{q+1} - v_q)(w_{q+1} - w_q) \\ &= \alpha \sum_{q \in \mathbb{Z}} q^2 v_q w_q + \eta \sum_{q \in \mathbb{Z}} (v_{q+1} - v_q) w_{q+1} - \eta \sum_{q \in \mathbb{Z}} (v_{q+1} - v_q) w_q \\ &= \alpha \sum_{q \in \mathbb{Z}} q^2 v_q w_q + \eta \sum_{q \in \mathbb{Z}} (v_q - v_{q-1}) w_q - \eta \sum_{q \in \mathbb{Z}} (v_{q+1} - v_q) w_q \end{aligned}$$

This proves $\forall q \in \mathbb{Z}, \quad \alpha q^2 v_q - \eta(v_{q+1} - 2v_q + v_{q-1}) = f_q$.

Conversely, if $\forall q \in \mathbb{Z}, \quad \alpha q^2 v_q - \eta(v_{q+1} - 2v_q + v_{q-1}) = f_q$, then we have by the same manipulations as before that:

$$\forall w \in H, \langle v, w \rangle_H = \langle f, w \rangle_{\ell^2(\mathbb{Z})}$$

Hence L is well-defined, obviously linear and $Lf \in H$.

Now, if we take $w = Lf$ we get

$$\langle Lf, Lf \rangle_H = \langle f, Lf \rangle_{\ell^2(\mathbb{Z})} \leq \|f\|_{\ell^2(\mathbb{Z})} \|Lf\|_{\ell^2(\mathbb{Z})} \leq C \|f\|_{\ell^2(\mathbb{Z})} \|Lf\|_H$$

so that $\|Lf\|_H \leq C \|f\|_{\ell^2(\mathbb{Z})}$ and L is hence continuous. \square

Now, we are able to prove important properties about L .

Lemma 4. *L is a positive, compact, self-adjoint operator*

Proof:

As far as the positiveness of the operator is concerned we just need to notice that, by definition:

$$\forall f \in \ell^2(\mathbb{Z}), \langle Lf, f \rangle_{\ell^2(\mathbb{Z})} = \|Lf\|_H^2 \geq 0$$

For compactness, we know that $\|Lf\|_H \leq C\|f\|_{\ell^2(\mathbb{Z})}$ and Lemma 2 allows to conclude.

Now, we prove that the operator L is self-adjoint.

Consider $f^1, f^2 \in \ell^2(\mathbb{Z})$, we have:

$$\langle f^1, Lf^2 \rangle_{\ell^2(\mathbb{Z})} = \langle Lf^1, Lf^2 \rangle_H = \langle Lf^2, Lf^1 \rangle_H = \langle f^2, Lf^1 \rangle_{\ell^2(\mathbb{Z})} = \langle Lf^1, f^2 \rangle_{\ell^2(\mathbb{Z})}$$

□

Now, we can go to the very proof of Theorem 1 and Theorem 3.

Step 1: Spectral decomposition and building of a solution when the terminal condition is in $\ell^2(\mathbb{Z})$.

We know that there exists an orthogonal basis $(f^k)_{k \in \mathbb{N}}$ of $\ell^2(\mathbb{Z})$ made of eigenvectors of L (that in fact belongs to H and we can take for instance $\|f^k\|_H = 1$) and we denote $\lambda^k > 0$ the eigenvalue¹² associated to f^k (we suppose that the eigenvalues are ordered, λ^0 being the largest one). We have:

$$\alpha q^2 f_q^k - \eta(f_{q+1}^k - 2f_q^k + f_{q-1}^k) = \frac{1}{\lambda^k} f_q^k$$

Hence, if we want to solve (\mathcal{S}') that is similar to (\mathcal{S}) but with a terminal condition $v(T) \in \ell^2(\mathbb{Z})$ instead of $v(T) = 1$ (where 1 stands for the sequence equal to 1 for all indices), classical argument shows that we can search for a solution of the form $v(t) = \sum_{k \in \mathbb{N}} \mu^k(t) f^k$.

Since $\forall q \in \mathbb{Z}$

$$\begin{aligned} \dot{v}_q(t) &= \alpha q^2 v_q(t) - \eta(v_{q-1}(t) + v_{q+1}(t)) \\ &= \alpha q^2 v_q(t) - \eta(v_{q+1}(t) - 2v_q(t) + v_{q-1}(t)) - 2\eta v_q(t) \end{aligned}$$

We must have $\frac{d\mu^k}{dt}(t) = (\frac{1}{\lambda^k} - 2\eta)\mu^k(t)$ and hence, since $\lambda^k \rightarrow 0$ we can easily define a solution of (\mathcal{S}') by:

$$v(t) = \sum_{k \in \mathbb{N}} \langle v(T), f^k \rangle_{\ell^2(\mathbb{Z})} \exp\left(\left(2\eta - \frac{1}{\lambda^k}\right)(T-t)\right) f^k$$

and the solution is in fact in $C^\infty([0, T], \ell^2(\mathbb{Z}))$

Step 2: Building of a solution when $v(T) = 1$

The first thing to notice is that $H \subset \ell^1(\mathbb{Z})$ (indeed, $w \in H \Rightarrow (qw_q)_q \in \ell^2(\mathbb{Z}) \Rightarrow (w_q)_q \in \ell^1(\mathbb{Z})$ by Cauchy-Schwarz inequality). Hence, the sequence $v(T)$ that equals 1 at each index

¹²0 cannot be an eigenvalue. If indeed $\lambda^k = 0$ then $\forall w \in H, \langle f^k, w \rangle_{\ell^2(\mathbb{Z})} = \langle Lf^k, w \rangle_H = 0$. But because H is dense in $\ell^2(\mathbb{Z})$, $f^k = 0$.

is in H' , the dual of H . As a consequence, to build a solution of (\mathcal{S}) , we can consider a similar formula:

$$v(t) = \sum_{k \in \mathbb{N}} \langle 1, f^k \rangle_{H', H} \exp \left(\left(2\eta - \frac{1}{\lambda^k} \right) (T - t) \right) f^k$$

Step 3: Uniqueness

Uniqueness follows easily from the $\ell^2(\mathbb{Z})$ analysis. If indeed $v(T) = 0$ we see that we must have that

$$\forall k \in \mathbb{Z}, \langle f^k, \dot{v}(t) \rangle_{\ell^2(\mathbb{Z})} = \frac{d \langle f^k, v(t) \rangle_{\ell^2(\mathbb{Z})}}{dt} = \left(\frac{1}{\lambda^k} - 2\eta \right) \langle f^k, v(t) \rangle_{\ell^2(\mathbb{Z})}$$

Hence, $\forall k \in \mathbb{Z}, \langle f^k, v(T) \rangle_{\ell^2(\mathbb{Z})} = 0 \implies \forall k \in \mathbb{Z}, \forall t \in [0, T], \langle f^k, v(t) \rangle_{\ell^2(\mathbb{Z})} = 0$ and $v = 0$.

Step 4: Asymptotics

To prove Theorem 3, we will show that the largest eigenvalue λ^0 of L is simple and that the associated eigenvector f^0 can be chosen so as to be a strictly positive sequence.

If this is true then we have $\forall q \in \mathbb{Z}, v_q(0) \underset{T \rightarrow \infty}{\sim} \langle 1, f^0 \rangle_{H', H} f_q^0 \exp \left((2\eta - \frac{1}{\lambda^0}) T \right)$ so that the result is proved with

$$\delta_\infty^{b*}(q) = \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right) + \frac{1}{k} \ln \left(\frac{f_q^0}{f_{q+1}^0} \right) \quad \delta_\infty^{a*}(q) = \frac{1}{\gamma} \ln \left(1 + \frac{\gamma}{k} \right) + \frac{1}{k} \ln \left(\frac{f_q^0}{f_{q-1}^0} \right)$$

Hence we just need to prove the following lemma:

Lemma 5. *The eigenvalue λ^0 is simple and any associated eigenvector is of constant sign (in a strict sense).*

Proof:

Let us consider the following characterization of λ^0 and of the associated eigenvectors (this characterization follows from the spectral decomposition):

$$\frac{1}{\lambda^0} = \inf_{f \in \ell^2(\mathbb{Z})} \frac{\|f\|_H^2}{\|f\|_{\ell^2(\mathbb{Z})}^2} = \inf_{\|f\|_{\ell^2(\mathbb{Z})} = 1} \|f\|_H^2 = \inf_{\|f\|_{\ell^2(\mathbb{Z})} = 1} \sum_{q \in \mathbb{Z}} \alpha q^2 f_q^2 + \eta (f_{q+1} - f_q)^2$$

Let us consider f an eigenvector associated to λ^0 . We have that:

$$\sum_{q \in \mathbb{Z}} \alpha q^2 |f_q|^2 + \eta (|f_{q+1}| - |f_q|)^2 \leq \sum_{q \in \mathbb{Z}} \alpha q^2 f_q^2 + \eta (f_{q+1} - f_q)^2$$

Hence, since $|f|$ has the same $\ell^2(\mathbb{Z})$ -norm as f , we know that $|f|$ is an eigenvector associated to λ^0 .

Now, since $\alpha q^2 |f_q| - \eta (|f_{q+1}| - 2|f_q| + |f_{q-1}|) = \frac{1}{\lambda^0} |f_q|$, if $|f_q| = 0$ at some point q , we have $-\eta (|f_{q+1}| + |f_{q-1}|) = 0$ and this induces $|f_{q+1}| = |f_{q-1}| = 0$, and $|f| = 0$ by immediate induction. Since $f \neq 0$, we must have therefore $|f| > 0$.

This proves that there exists a strictly positive eigenvector associated to λ^0 .

Now, if the eigenvalue λ^0 were not simple, there would exist an eigenvector g associated to λ^0 with $\langle |f|, g \rangle_{\ell^2(\mathbb{Z})} = 0$. Hence, there would exist both positive and negative values in the sequence g . But, in that case, since $|g|$ must also be an eigenvector associated to λ^0 , we must have equality in the following inequality:

$$\sum_{q \in \mathbb{Z}} \alpha q^2 |g_q|^2 + \eta (|g_{q+1}| - |g_q|)^2 \leq \sum_{q \in \mathbb{Z}} \alpha q^2 g_q^2 + \eta (g_{q+1} - g_q)^2$$

In particular, we must have that $||g_{q+1}| - |g_q|| = |g_{q+1} - g_q|, \forall q$. This implies that $\forall q$, either g_q and g_{q+1} are of the same sign or at least one of the two terms is equal to 0. Thus, since g cannot be of constant sign, there must exist \hat{q} so that $g_{\hat{q}} = 0$. But then, because $|g|$ is also an eigenvector associated to λ^0 we have by immediate induction, as above, that $g = 0$.

This proves that there is no such g and that the eigenvalue is simple. \square

Step 5: Positiveness

So far, we did not prove that v , the solution of (\mathcal{S}) was strictly positive. A rapid way to prove that point is to use a Feynman-Kac-like representation of v . If $(q_s)_s$ is a continuous Markov chain on \mathbb{Z} with intensities η to jump to immediate neighbors, then we have the following representation for v :

$$v_q(t) = \mathbb{E} \left[\exp \left(- \int_t^T (\alpha q_s^2 - 2\eta) ds \right) \middle| q_t = q \right]$$

This representation guarantees that $v > 0$.

This ends the proof of Theorems 1 and 3 \square

Proof of Proposition 2:

The same computations as for Proposition 1 lead to the result. \square

Proof of Theorem 4:

Using Proposition 2, because the functions v_q s are strictly positive as above, we see that u is a solution of (HJB). Then, we are going to use a verification argument to prove that u is the value function of the optimal control problem under consideration and prove subsequently that the optimal controls are as given in the Theorem.

Let us consider processes (ν^b) and $(\nu^a) \in \mathcal{A}$.

Let $t \in [0, T)$ and let us consider the following processes for $\tau \in [t, T]$:

$$\begin{aligned} dS_\tau^{t,s} &= \sigma dW_\tau, & S_t^{t,s} &= s \\ dX_\tau^{t,x,\nu} &= (S_\tau + \nu_\tau^a) dN_\tau^a - (S_\tau - \nu_\tau^b) dN_\tau^b, & X_t^{t,x,\nu} &= x \\ dq_\tau^{t,q,\nu} &= dN_\tau^b - dN_\tau^a, & q_t^{t,q,\nu} &= q \end{aligned}$$

where the point process N^b has stochastic intensity $(\lambda_\tau^b)_\tau$ with $\lambda_\tau^b = Ae^{-k\nu_\tau^b} 1_{q_{\tau-} < Q}$ and where the point process N^a has stochastic intensity $(\lambda_\tau^a)_\tau$ with $\lambda_\tau^a = Ae^{-k\nu_\tau^a} 1_{q_{\tau-} > -Q}$ ¹³.

¹³These intensities are bounded since ν^b and ν^a are bounded from below.

Now, since u is smooth, let us write Itô's formula for u , between t and t_n where $t_n = T \wedge \inf\{\tau > t, |S_\tau - s| \geq n \text{ or } |N_\tau^a - N_t^a| \geq n \text{ or } |N_\tau^b - N_t^b| \geq n\}$ (n being a strictly positive integer):

$$\begin{aligned}
& u(t_n, X_{t_n-}^{t,x,\nu}, q_{t_n-}^{t,q,\nu}, S_{t_n}^{t,s}) = u(t, x, q, s) \\
& + \int_t^{t_n} \left(\partial_\tau u(\tau, X_{\tau-}^{t,x,\nu}, q_{\tau-}^{t,q,\nu}, S_\tau^{t,s}) + \frac{\sigma^2}{2} \partial_{ss}^2 u(\tau, X_{\tau-}^{t,x,\nu}, q_{\tau-}^{t,q,\nu}, S_\tau^{t,s}) \right) d\tau \\
& + \int_t^{t_n} \left(u(\tau, X_{\tau-}^{t,x,\nu} + S_\tau^{t,s} + \nu_\tau^a, q_{\tau-}^{t,q,\nu} - 1, S_\tau^{t,s}) - u(\tau, X_{\tau-}^{t,x,\nu}, q_{\tau-}^{t,q,\nu}, S_\tau^{t,s}) \right) \lambda_\tau^a d\tau \\
& + \int_t^{t_n} \left(u(\tau, X_{\tau-}^{t,x,\nu} - S_\tau^{t,s} + \nu_\tau^b, q_{\tau-}^{t,q,\nu} + 1, S_\tau^{t,s}) - u(\tau, X_{\tau-}^{t,x,\nu}, q_{\tau-}^{t,q,\nu}, S_\tau^{t,s}) \right) \lambda_\tau^b d\tau \\
& + \int_t^{t_n} \sigma \partial_s u(\tau, X_{\tau-}^{t,x,\nu}, q_{\tau-}^{t,q,\nu}, S_\tau^{t,s}) dW_\tau \\
& + \int_t^{t_n} \left(u(\tau, X_{\tau-}^{t,x,\nu} + S_\tau^{t,s} + \nu_\tau^a, q_{\tau-}^{t,q,\nu} - 1, S_\tau^{t,s}) - u(\tau, X_{\tau-}^{t,x,\nu}, q_{\tau-}^{t,q,\nu}, S_\tau^{t,s}) \right) dM_\tau^a \\
& + \int_t^{t_n} \left(u(\tau, X_{\tau-}^{t,x,\nu} - S_\tau^{t,s} + \nu_\tau^b, q_{\tau-}^{t,q,\nu} + 1, S_\tau^{t,s}) - u(\tau, X_{\tau-}^{t,x,\nu}, q_{\tau-}^{t,q,\nu}, S_\tau^{t,s}) \right) dM_\tau^b
\end{aligned}$$

where M^b and M^a are compensated processes associated respectively to N^b and N^a for the intensity processes $(\lambda_\tau^b)_\tau$ and $(\lambda_\tau^a)_\tau$.

Now, because v is strictly positive on a compact set, it has a strictly positive lower bound and $v_{q_\tau}(\tau)^{-\frac{1}{k}}$ is bounded along the trajectory, independently of the trajectory. Also, because ν^b and ν^a are bounded from below, and because of the definition of t_n , all the terms in the stochastic integrals are bounded and, local martingales being in fact martingales, we have:

$$\begin{aligned}
& \mathbb{E} \left[u(t_n, X_{t_n-}^{t,x,\nu}, q_{t_n-}^{t,q,\nu}, S_{t_n}^{t,s}) \right] = u(t, x, q, s) \\
& + \mathbb{E} \left[\int_t^{t_n} \left(\partial_\tau u(\tau, X_{\tau-}^{t,x,\nu}, q_{\tau-}^{t,q,\nu}, S_\tau^{t,s}) + \frac{\sigma^2}{2} \partial_{ss}^2 u(\tau, X_{\tau-}^{t,x,\nu}, q_{\tau-}^{t,q,\nu}, S_\tau^{t,s}) \right) d\tau \right. \\
& + \int_t^{t_n} \left(u(\tau, X_{\tau-}^{t,x,\nu} + S_\tau^{t,s} + \nu_\tau^a, q_{\tau-}^{t,q,\nu} - 1, S_\tau^{t,s}) - u(\tau, X_{\tau-}^{t,x,\nu}, q_{\tau-}^{t,q,\nu}, S_\tau^{t,s}) \right) \lambda_\tau^a d\tau \\
& \left. + \int_t^{t_n} \left(u(\tau, X_{\tau-}^{t,x,\nu} - S_\tau^{t,s} + \nu_\tau^b, q_{\tau-}^{t,q,\nu} + 1, S_\tau^{t,s}) - u(\tau, X_{\tau-}^{t,x,\nu}, q_{\tau-}^{t,q,\nu}, S_\tau^{t,s}) \right) \lambda_\tau^b d\tau \right]
\end{aligned}$$

Using the fact that u solves (HJB), we have that

$$\mathbb{E} \left[u(t_n, X_{t_n-}^{t,x,\nu}, q_{t_n-}^{t,q,\nu}, S_{t_n}^{t,s}) \right] \leq u(t, x, q, s)$$

with equality when the controls are taken equal the maximizers of the hamiltonians (these controls being in \mathcal{A} because v is bounded and has a strictly positive lower bound).

Now, if we prove that $\lim_{n \rightarrow \infty} \mathbb{E} \left[u(t_n, X_{t_n-}^{t,x,\nu}, q_{t_n-}^{t,q,\nu}, S_{t_n}^{t,s}) \right] = \mathbb{E} \left[u(T, X_T^{t,x,\nu}, q_T^{t,q,\nu}, S_T^{t,s}) \right]$, we will have that for all controls ν^b and ν^a in \mathcal{A} :

$$\mathbb{E} \left[-\exp \left(-\gamma (X_T^{t,x,\nu} + q_T^{t,q,\nu} S_T^{t,s}) \right) \right] = \mathbb{E} \left[u(T, X_T^{t,x,\nu}, q_T^{t,q,\nu}, S_T^{t,s}) \right] \leq u(t, x, q, s)$$

with equality for $\nu_t^b = \delta^{b*}(t, q_{t-})$ and $\nu_t^a = \delta^{a*}(t, q_{t-})$.

Hence:

$$\sup_{(\nu_t^a)_t, (\nu_t^b)_t \in \mathcal{A}} \mathbb{E} \left[-\exp \left(-\gamma (X_T^{t,x,\nu} + q_T^{t,q,\nu} S_T^{t,s}) \right) \right] = \mathbb{E} \left[-\exp \left(-\gamma (X_T^{t,x,\delta^*} + q_T^{t,q,\delta^*} S_T^{t,s}) \right) \right]$$

and this will give the result.

$$\text{It remains to prove that } \lim_{n \rightarrow \infty} \mathbb{E} \left[u(t_n, X_{t_n^-}^{t,x,\nu}, q_{t_n^-}^{t,q,\nu}, S_{t_n}^{t,s}) \right] = \mathbb{E} \left[u(T, X_T^{t,x,\nu}, q_T^{t,q,\nu}, S_T^{t,s}) \right].$$

First, we have, almost surely, that $u(t_n, X_{t_n^-}^{t,x,\nu}, q_{t_n^-}^{t,q,\nu}, S_{t_n}^{t,s})$ tends to $u(T, X_T^{t,x,\nu}, q_T^{t,q,\nu}, S_T^{t,s})$. Then, in order to prove that the sequence is uniformly integrable we will bound it in L^2 . However, because of the uniform lower bound on v already used early, it is sufficient to bound $\exp \left(-\gamma \left(X_{t_n^-}^{t,x,\nu} + q_{t_n^-}^{t,q,\nu} S_{t_n}^{t,s} \right) \right)$ in L^2 .

But,

$$\begin{aligned} X_{t_n^-}^{t,x,\nu} + q_{t_n^-}^{t,q,\nu} S_{t_n}^{t,s} &= x + qs + \int_t^{t_n} \nu_\tau^a dN_\tau^a + \int_t^{t_n} \nu_\tau^b dN_\tau^b + \sigma \int_t^{t_n} q_\tau^{t,q,\nu} dW_\tau \\ &\geq x + qs - \|\nu_-^a\|_\infty N_T^a - \|\nu_-^b\|_\infty N_T^b + \sigma \int_t^{t_n} q_\tau^{t,q,\nu} dW_\tau \end{aligned}$$

Hence

$$\begin{aligned} &\mathbb{E} \left[\exp(-2\gamma(X_{t_n^-}^{t,x,\nu} + q_{t_n^-}^{t,q,\nu} S_{t_n}^{t,s})) \right] \\ &\leq e^{-2\gamma(x+qs)} \mathbb{E} \left[\exp(2\gamma\|\nu_-^a\|_\infty N_T^a) \exp(2\gamma\|\nu_-^b\|_\infty N_T^b) \exp\left(-2\gamma\sigma \int_t^{t_n} q_\tau^{t,q,\nu} dW_\tau\right) \right] \\ &\leq e^{-2\gamma(x+qs)} \mathbb{E} \left[\exp(6\gamma\|\nu_-^a\|_\infty N_T^a) \right]^{\frac{1}{3}} \mathbb{E} \left[\exp(6\gamma\|\nu_-^b\|_\infty N_T^b) \right]^{\frac{1}{3}} \mathbb{E} \left[\exp\left(-6\gamma\sigma \int_t^{t_n} q_\tau^{t,q,\nu} dW_\tau\right) \right]^{\frac{1}{3}} \end{aligned}$$

Now, since the intensity of each point process is bounded, the point processes have a Laplace transform and the first three terms of the product are finite (and independent of n). Concerning the last term, because $|q_\tau^{t,q,\nu}|$ is bounded by Q , we know (for instance applying Girsanov's theorem and Hölder's inequality) that:

$$\mathbb{E} \left[\exp\left(-6\gamma\sigma \int_t^{t_n} q_\tau^{t,q,\nu} dW_\tau\right) \right] \leq \exp(3\gamma^2\sigma^2(t_n - t)Q^2) \leq \exp(3\gamma^2\sigma^2Q^2T)$$

Hence, the sequence is bounded in L^2 , then uniformly integrable, and we have:

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[u(t_n, X_{t_n^-}^{t,x,\nu}, q_{t_n^-}^{t,q,\nu}, S_{t_n}^{t,s}) \right] = \mathbb{E} \left[u(T, X_T^{t,x,\nu}, q_T^{t,q,\nu}, S_T^{t,s}) \right] = \mathbb{E} \left[u(T, X_T^{t,x,\nu}, q_T^{t,q,\nu}, S_T^{t,s}) \right]$$

We have proved that u is the value function and that δ^{b*} and δ^{a*} are optimal controls. \square

Proof of Theorem 5:

The demonstration of the result and the characterization of f^0 is the same as in Theorem 3 once it is noticed that the counterpart of L is now $(M + 2\eta I)^{-1}$. \square

References

- [1] Aurélien Alfonsi and Alexander Schied. Optimal Execution and Absence of Price Manipulations in Limit Order Book Models. *Social Science Research Network Working Paper Series*, November 2009.
- [2] R. F. Almgren and N. Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, 3(2):5–39, 2000.
- [3] Y. Amihud and H. Mendelson. Dealership Market. Market-Making with Inventory. *Journal of Financial Economics*, 8:31–53, 1980.
- [4] Y. Amihud and H. Mendelson. Asset pricing and the bid-ask spread. *Journal of Financial Economics*, 17(2):223–249, 1986.
- [5] Marco Avellaneda and Sasha Stoikov. High-frequency trading in a limit order book. *Quantitative Finance*, 8(3):217–224, 2008.
- [6] E. Bacry, S. Delattre, M. Hoffmann, and J. F. Muzy. Modeling microstructure noise with mutually exciting point processes. January 2011.
- [7] L. Benston Robert and J. George. Determinants of bid-asked spreads in the over-the-counter market. *Journal of Financial Economics*, 1(4):353–364, 1974.
- [8] Dimitris Bertsimas and Andrew W. Lo. Optimal control of execution costs. *Journal of Financial Markets*, 1(1):1–50, 1998.
- [9] Bruno Bouchard, Ngoc-Minh Dang, and Charles-Albert Lehalle. Optimal control of trading algorithms: a general impulse control approach. *SIAM J. Financial Mathematics*, 2011.
- [10] K. J. Cohen, S. F. Maier, R. A. Schwartz, and D. K. Whitcomb. Market makers and the market spread: A review of recent literature. *Journal of Financial and Quantitative Analysis*, 14(04):813–835, 1979.
- [11] Kalman J. Cohen, Steven F. Maier, Robert A. Schwartz, and David K. Whitcomb. Transaction Costs, Order Placement Strategy, and Existence of the Bid-Ask Spread. *The Journal of Political Economy*, 89(2):287–305, 1981.
- [12] Rama Cont, Arseniy Kukanov, and Sasha Stoikov. The Price Impact of Order Book Events. *Social Science Research Network Working Paper Series*, November 2010.
- [13] M. B. Garman. Market microstructure. *Journal of Financial Economics*, 3(3):257–275, 1976.
- [14] O. Guéant, C. A. Lehalle, and J. Fernandez-Tapia. Optimal Execution with Limit Orders. *Working paper*, 2011.
- [15] T. Hendershott and A.J. Menkveld. Price pressures. *Manuscript, VU University Amsterdam*, 2009.
- [16] T. S. Y. Ho and R. G. Macris. Dealer bid-ask quotes and transaction prices: An empirical study of some AMEX options. *Journal of Finance*, pages 23–45, 1984.
- [17] T. S. Y. Ho and H. R. Stoll. The dynamics of dealer markets under competition. *Journal of Finance*, 38(4):1053–1074, 1983.
- [18] Thomas Ho and Hans R. Stoll. Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial Economics*, 9(1):47–73, March 1981.
- [19] Charles-Albert Lehalle, Olivier Guéant, and Julien Razafimanana. High Frequency Simulations of an Order Book: a Two-Scales Approach. In F. Abergel, B. K. Chakrabarti, A. Chakraborti, and M. Mitra, editors, *Econophysics of Order-Driven Markets*, New Economic Windows. Springer, 2010.

- [20] A. Madhavan and S. Smidt. An analysis of changes in specialist inventories and quotations. *The Journal of Finance*, 48(5):1595–1628, 1993.
- [21] Albert J. Menkveld. High Frequency Trading and The New-Market Makers. *Social Science Research Network Working Paper Series*, December 2010.
- [22] E. Miltenstein and H. Schleef. The optimal pricing policy of a monopolistic market-maker in the equity market. *Journal of Finance*, pages 218–231, 1983.
- [23] M. O’Hara and G. S. Oldfield. The microeconomics of market making. *Journal of Financial and Quantitative Analysis*, 21(04):361–376, 1986.
- [24] Gilles Pagès, Sophie Laruelle, and Charles-Albert Lehalle. Optimal split of orders across liquidity pools: a stochastic algorithm approach. Technical report, 2009.
- [25] R. Roll. A simple implicit measure of the effective bid-ask spread in an efficient market. *Journal of Finance*, 39(4):1127–1139, 1984.