Dealing with the Inventory Risk: A Solution to the Market Making Problem

Article *in* Mathematics and Financial Economics · May 2011

DOI: 10.1007/s11579-012-0087-0 · Source: arXiv

3 authors, including:

Olivier Gueant
École Nationale de la Statistique et de l'Administration Économique

Joaquin Fernandez-Tapia

Dealing with the Inventory Risk

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

2010

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 [18]).

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 (respectively buyer). Market participants focused on this kind of liquidity-providing activity

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), 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.

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, 15, 21]).

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, 16, 20, 23]).

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 [19]).

Avellaneda and Stoïkov 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 (Theorems 1-2)**.** *The optimal quotes can be expressed as:*

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

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

  where γ is the risk aversion of the agent and where v is a family of strictly positive functions $(v_q)_{q\in\mathbb{Z}}$ solution of the linear system of ODEs $(\mathcal{S})$ that follows:

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

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

  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 3 (asymptotics) and the associated approximation equations)**.**

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

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

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 [22] 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 goes from the description of the model choices that had to be made (section 1), through the main change of variables (section 2), exposes the asymptotics of the obtained dynamics (section 3), its comparative statics (section 4), extends the framework to trends in prices and constraints on the inventory (section 5), finally discusses the model choices that had to be made (section 6) and ends with an application to real data (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 Stoïkov 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 [19]).

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 that small perturbations should be needed.

1 Setup of the model

1.1 Prices and Orders

We consider a market maker operating on a single stock and whose size is small enough to consider price dynamics exogenous. For the sake of simplicity and since we will basically only consider short horizon problems we suppose that the mid-price of the stock moves as a brownian motion:

$$dS_t = \sigma dW_t$$

The market maker under consideration will continuously propose bid and ask quotes denoted respectively S^b_t and S^a_t and will hence buy and sell stocks according to the rate of arrival of aggressive orders at the quoted prices.

His inventory q, that is the (signed) quantity of stocks he holds, is given by $q_t = N^b_t - N^a_t$ where N^b and N^a are the jump processes giving the number of stocks 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 size 1. Arrival rates obviously depend on the prices S^b_t and S^a_t quoted by the market maker and we assume, in accordance with most datasets, that intensities λ^b and λ^a associated to N^b and N^a are of the following form[1]:

$$\lambda^b(s^b, s) = A \exp(-k(s - s^b)) \qquad \lambda^a(s^a, s) = A \exp(-k(s^a - s))$$

This means that the closer to the mid-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:

$$dX_t = S_t^a dN_t^a - S_t^b dN_t^b$$

1.2 The optimization problem

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:

$$\sup_{S^a, S^b} \mathbb{E}\left[-\exp\left(-\gamma(X_T + q_T S_T)\right)\right]$$

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 mid-price evaluation of the (signed) remaining quantity of stocks in the inventory at time T (liquidation at mid-price[2]).

2 Resolution

2.1 Hamilton-Jacobi-Bellman equation

The optimization problem set up in the preceding section can be solved using classical Bellman tools. To this purpose, we introduce a Bellman function u defined as:

$$u(t, x, q, s) = \sup_{S^a, S^b} \mathbb{E}\left[-\exp\left(-\gamma(X_T + q_T S_T)\right) \middle| X_t = x, S_t = s, q_t = q\right]$$

The Hamilton-Jacobi-Bellman equation associated to the optimization problem is then given by the following proposition:

Proposition 1 (HJB). *The Hamilton-Jacobi-Bellman equation for u is:*

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

with the final condition:

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

This equation is not a *simple* 4-variable PDE. Rather, because the inventory is discrete, it is an infinite system of 3-variable PDEs. To solve it, we will use a change of variables that is different from the one used in [5] and transforms the system of PDEs in a system of linear ODEs.

[1] Although this form is in accordance with real data, some authors used a linear form for the intensity functions – see [17] for instance.

[2] We will discuss other hypotheses below.

2.2 Reduction to a system of linear ODEs

In [5], the authors proposed a change of variables to factor out wealth. Here we go further and propose a rather non-intuitive change of variables that allows to write the problem in a linear way.

Proposition 2 (A system of linear ODEs). *Let's consider a family of strictly positive functions $(v_q)_{q\in\mathbb{Z}}$ solution of the linear system of ODEs $(\mathcal{S})$ that follows:*

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

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 (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.*

2.3 Optimal quotes characterization

Theorem 2 (Optimal quotes and bid-ask spread). *Let's consider the solution v of the system $(\mathcal{S})$ as in Theorem 1. Then optimal quotes can be expressed as:*

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

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

Moreover, the bid-ask spread quoted by the market maker, that is $\psi^* = s^{a*}(t,q,s) - s^{b*}(t,q,s)$, is given by:

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

We see that the difference between each quoted price and the mid-price has two components. If we consider the case of the bid quote – the same analysis would be true in the case of the ask quote –, we need to separate the term $-\frac{1}{k}\ln\left(\frac{v_{q+1}(t)}{v_q(t)}\right)$ from the term $\frac{1}{\gamma}\ln\left(1+\frac{\gamma}{k}\right)$. If $\sigma = 0$, then $v_q(t) = \exp(2\eta(T-t))$ defines a solution of the system $(\mathcal{S})$. Hence, the relations $s - s^{b*} = s^{a*} - s = \frac{1}{\gamma}\ln\left(1+\frac{\gamma}{k}\right)$ define the optimal quotes in the "no-volatility" benchmark case[3]. Consequently, in the expression that defines the optimal quotes, the second term corresponds to the "no-volatility" benchmark while the first one takes account of the influence of volatility.

3 Examples and asymptotics

To motivate the asymptotic approximation we provide, and before discussing the way to solve the problem numerically, let us present some graphs to understand the behavior in time and inventory of both the optimal quotes and the bid-ask spread.

[3] Smaller quotes would lead to trade more often with less revenue per trade in a way that is not in favor of the market maker. Symmetrically, larger quotes would lead to more revenue per trade but less trades and the welfare of the market maker would also be reduced.

3.1 Numerical examples

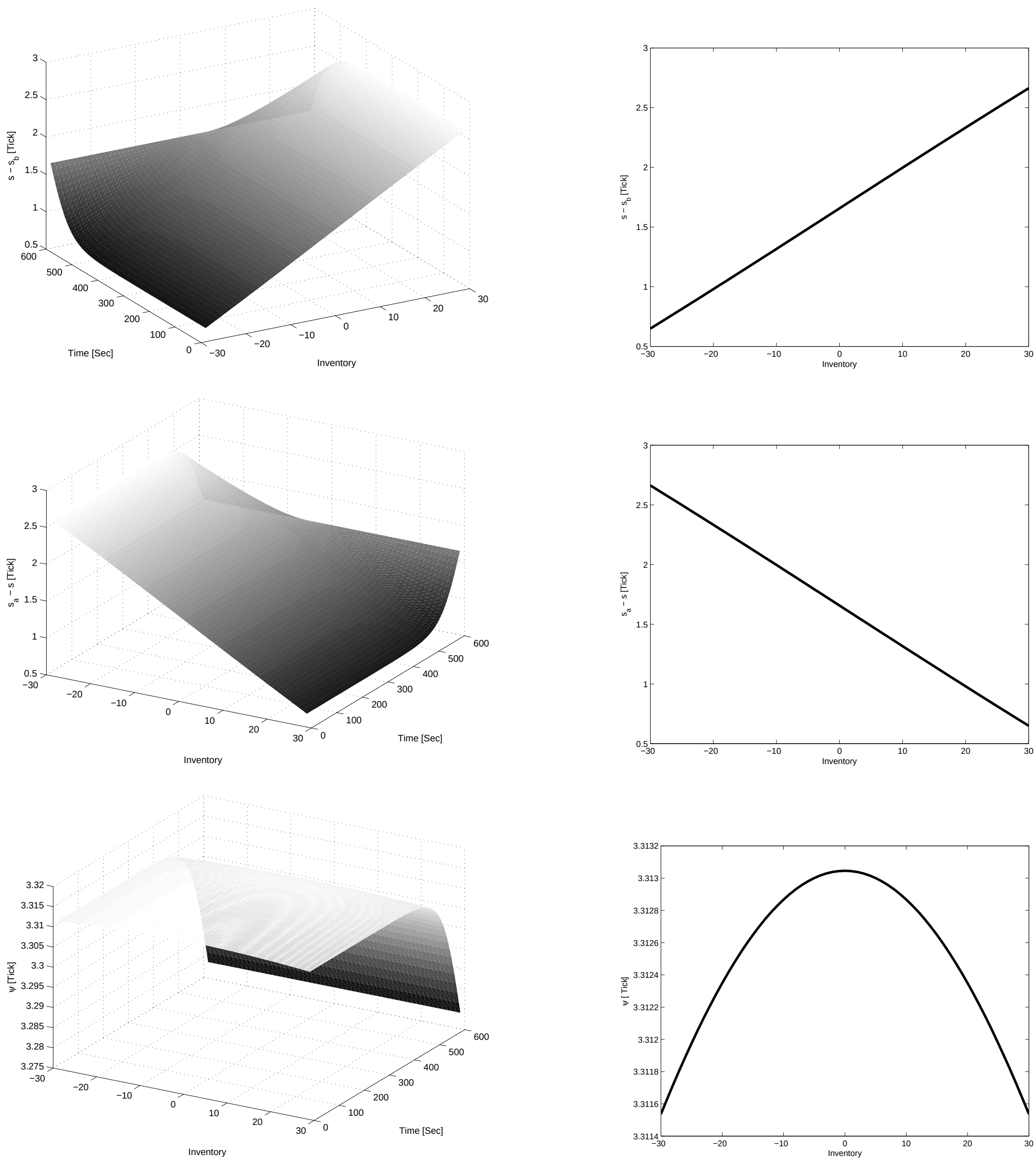


Figure 1: Left: Behavior of the optimal quotes and bid-ask spread with time and inventory. Right: Behavior of the optimal quotes and bid-ask spread with inventory, at time $t = 0$. $\sigma = 0.3 \quad \mathrm{Tick} \cdot \mathrm{s}^{-1/2}$, $A = 0.9 \quad \mathrm{s}^{-1}$, $k = 0.3 \quad \mathrm{Tick}^{-1}$,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

3.2 Asymptotics

In [5], the authors propose a heuristic approximation for the bid-ask spread. Namely they propose to approximate $\psi^*(t, q)$ by $\gamma\sigma^2(T - t) + \frac{2}{\gamma}\ln\left(1 + \frac{\gamma}{k}\right)$. However, as suggested by the graphs exhibited above, the predominant feature is that both the bid-ask spread and the distance between the quotes and the mid-price are rather constant, except near the time horizon T (and in numerical examples, a few minutes are enough to be near the asymptotic values), and certainly not linearly decreasing with time.

In fact, we can prove the existence of an asymptotic behavior and provide semi-explicit expressions for the asymptotic values of the bid-ask spread and the quotes:

Theorem 3 (Asymptotic quotes and bid-ask spread)**.**

$$\forall q \in \mathbb{Z}, \exists \delta_\infty^{b*}(q), \delta_\infty^{a*}(q), \psi_\infty^*(q) \in \mathbb{R}$$

$$\lim_{T\to\infty} s - s^{b*}(0,q,s) = \delta_\infty^{b*}(q)$$

$$\lim_{T\to\infty} s^{a*}(0,q,s) - s = \delta_\infty^{a*}(q)$$

$$\lim_{T\to\infty} \psi^*(0,q) = \psi_\infty^*(q)$$

Moreover,

$$\delta_\infty^{b*}(q) = \frac{1}{\gamma}\ln\left(1+\frac{\gamma}{k}\right) - \frac{1}{k}\ln\left(\frac{f^0_{q+1}}{f^0_q}\right) \qquad \delta_\infty^{a*}(q) = \frac{1}{\gamma}\ln\left(1+\frac{\gamma}{k}\right) + \frac{1}{k}\ln\left(\frac{f^0_q}{f^0_{q-1}}\right)$$

and

$$\psi_\infty^*(q) = -\frac{1}{k}\ln\left(\frac{f^0_{q+1}f^0_{q-1}}{f^{0\,2}_q}\right) + \frac{2}{\gamma}\ln\left(1+\frac{\gamma}{k}\right)$$

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

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

As we have seen in the above numerical examples, only these asymptotic values seem to be relevant in practice. Consequently, we provide an approximation for f^0 that happens to fit the actual figures. This approximation 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 proved[4] 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^0_q to behave as $\exp\left(-\frac{1}{2}\sqrt{\frac{\alpha}{\eta}}q^2\right)$.

This heuristic viewpoint induces an approximation for the optimal quotes and bid-ask spread if we replace f^0_q by $\exp\left(-\frac{1}{2}\sqrt{\frac{\alpha}{\eta}}q^2\right)$:

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

[4] To 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 $\int_{-\infty}^{\infty}\left(\alpha x^2 v(x)^2 + \eta v'(x)^2\right) dx < +\infty$. L_c can be diagonalized and largest eigenvalue of L_c can be shown to be associated to the eigenvector $f(x) = \exp\left(-\frac{1}{2}\sqrt{\frac{\alpha}{\eta}}x^2\right)$.

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

We exhibit below the values of the optimal quotes and the bid-ask spread, both with their associated approximations. Empirically, these approximations for the quotes are satisfactory in most cases and are always 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).

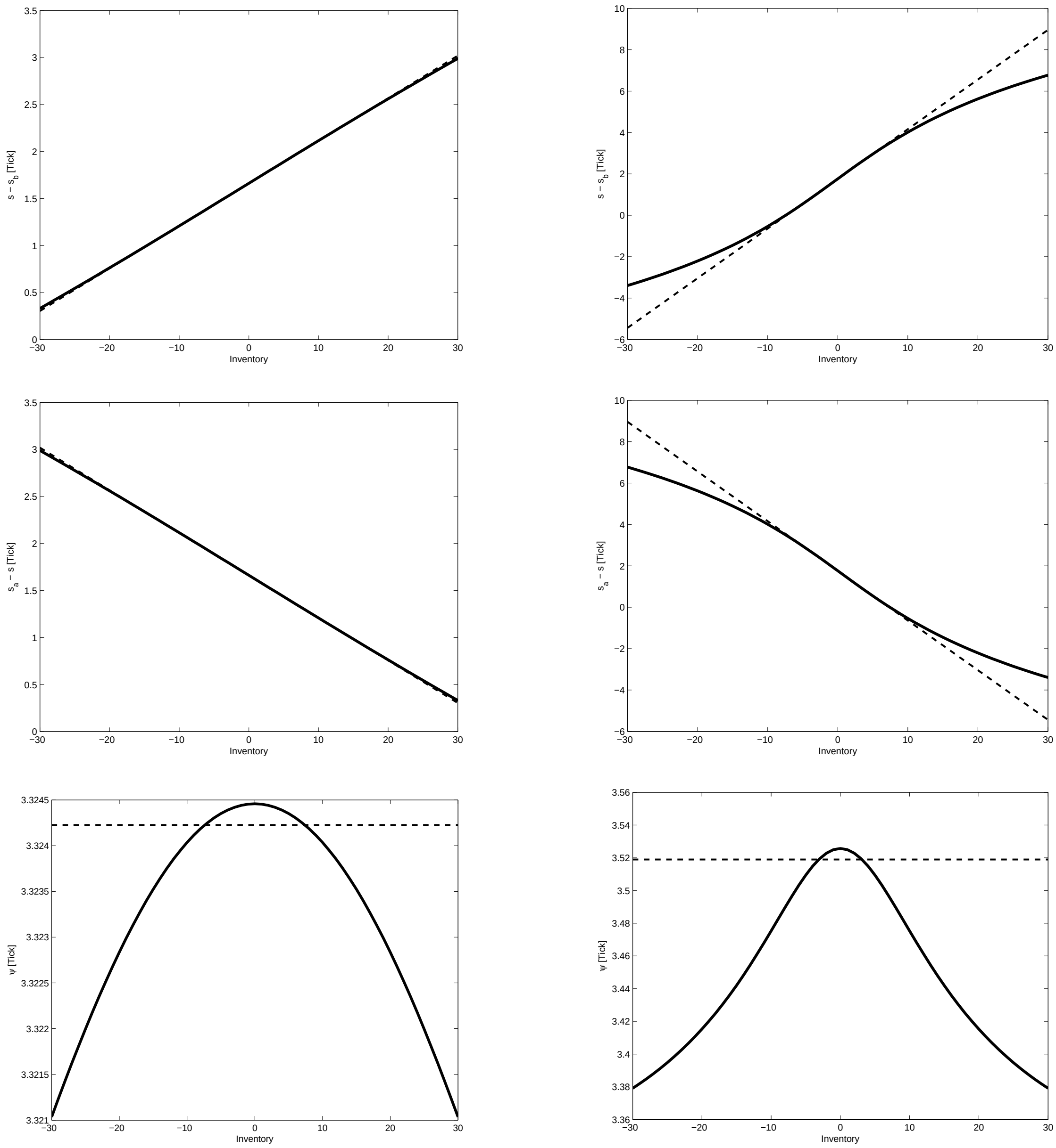


Figure 2: Asymptotic behavior of optimal quotes and the bid-ask spread (bold line). Approximation (dotted line). Left: $\sigma = 0.4 \quad \text{Tick}\cdot \text{s}^{-1/2}$, $A = 0.9 \quad \text{s}^{-1}$, $k = 0.3 \quad \text{Tick}^{-1}$, $\gamma = 0.01 \quad \text{Tick}^{-1}$, $T = 600 \quad s$. Right: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

4 Comparative statics

Before starting with the comparative statics, we rewrite the approximations done 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:

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

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

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

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:

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

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.

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

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

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.

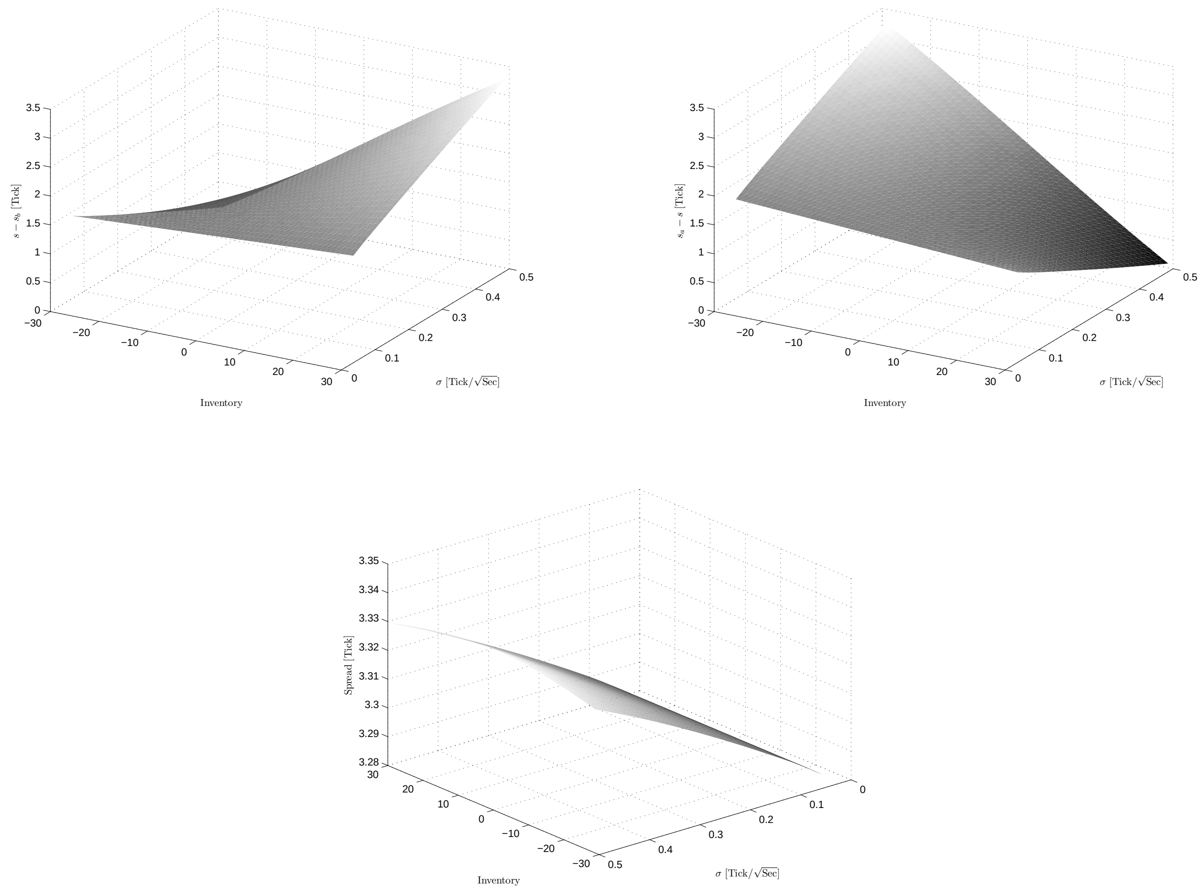
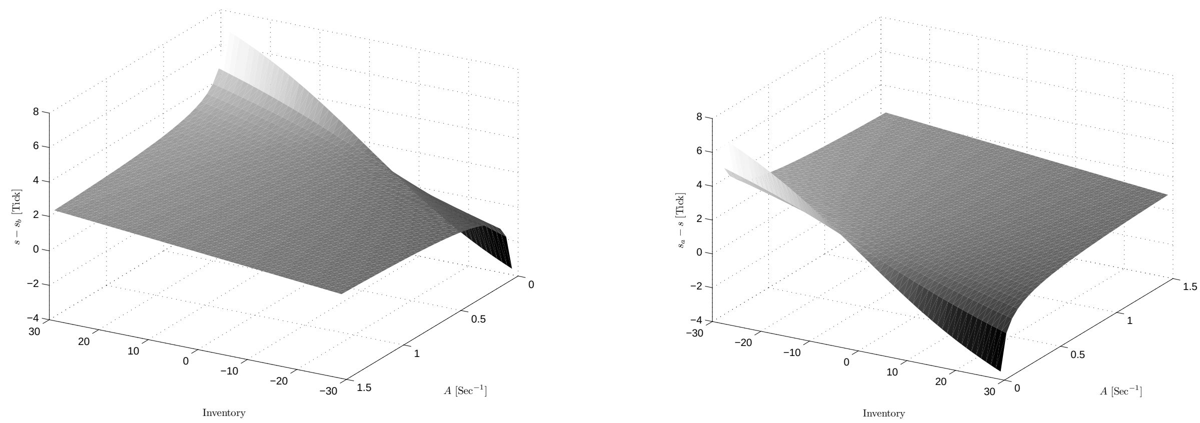


Figure 3: Asymptotic optimal quotes and bid-ask spread for different inventories and different values for the volatility σ. $A = 0.9 \quad \mathrm{s}^{-1}$, $k = 0.3 \quad \mathrm{Tick}^{-1}$, $\gamma = 0.01 \quad \mathrm{Tick}^{-1}$, $T = 600 \quad s$.



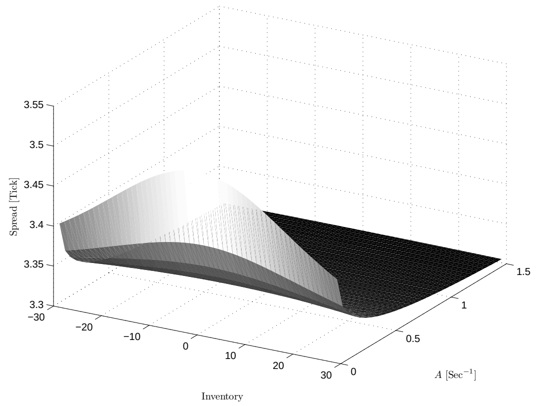


Figure 4: Asymptotic optimal quotes and bid-ask spread for different inventories and different values of A. $\sigma = 0.3 \quad \mathrm{Tick} \cdot \mathrm{s}^{-1/2}$, $k = 0.3 \quad \mathrm{Tick}^{-1}$, $\gamma = 0.01 \quad \mathrm{Tick}^{-1}$, $T = 600 \quad s$.

4.3 Dependence on k

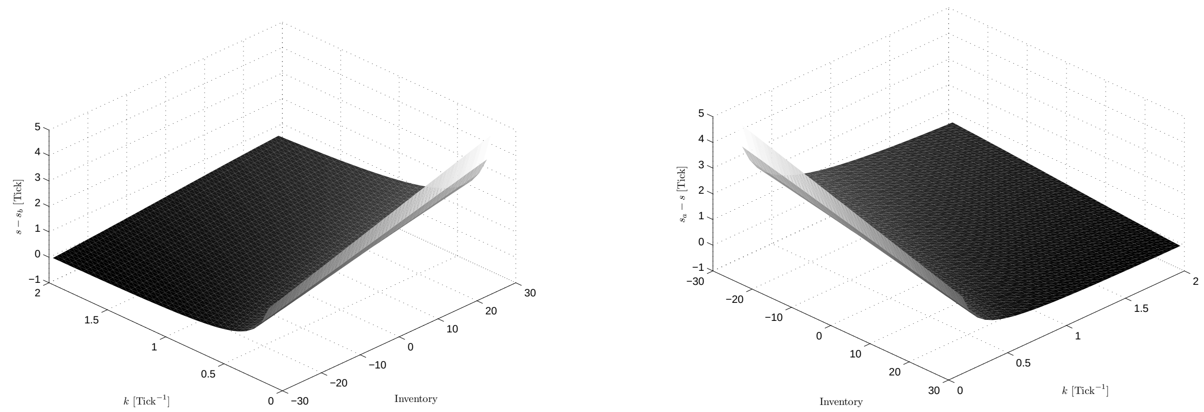
From the above approximations we expect δ_∞^{b*} to be decreasing in k for q greater than some negative threshold. Below this threshold we expect it to be increasing. Similarly we expect δ_∞^{a*} to be decreasing in k for q smaller than some positive threshold. Above this threshold we expect it to be 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, trades occur closer to the mid price. For this reason, and in absence of inventory risk, the optimal quotes have to get closer to the mid-price. 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:



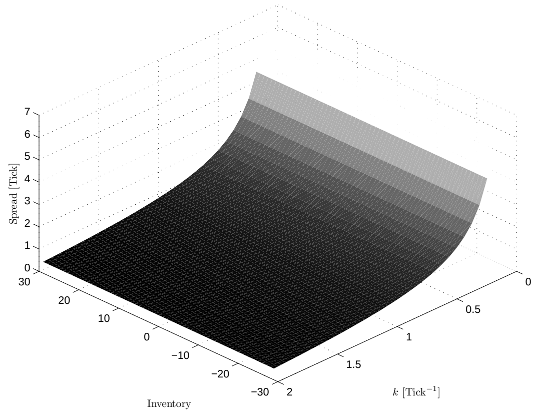
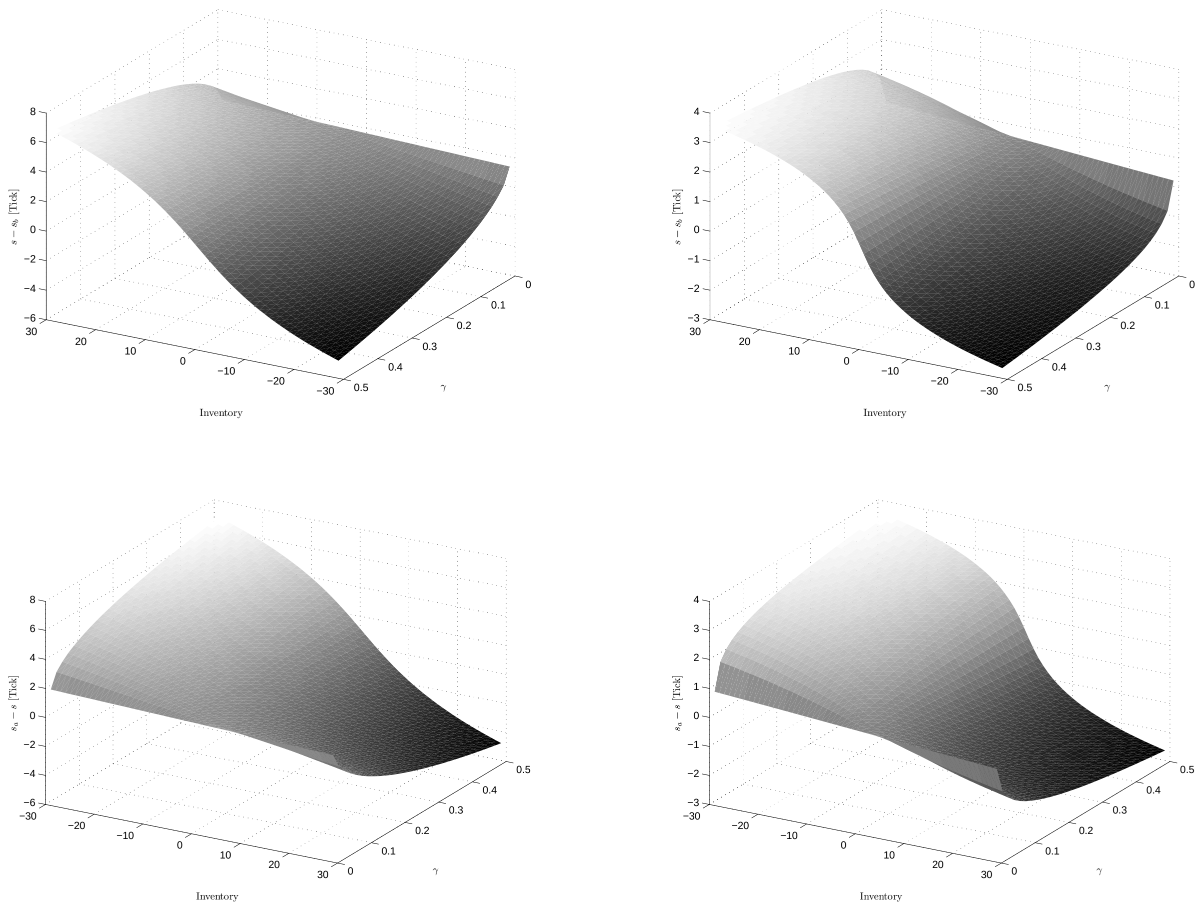


Figure 5: Asymptotic optimal quotes and bid-ask spread for different inventories and different values of k. $\sigma = 0.3 \quad \text{Tick} \cdot \text{s}^{-1/2}$, $A = 0.9 \quad \text{s}^{-1}$, $\gamma = 0.01 \quad \text{Tick}^{-1}$, $T = 600 \quad s$.

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 mid 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 (as in the "no-volatility" benchmark) 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:



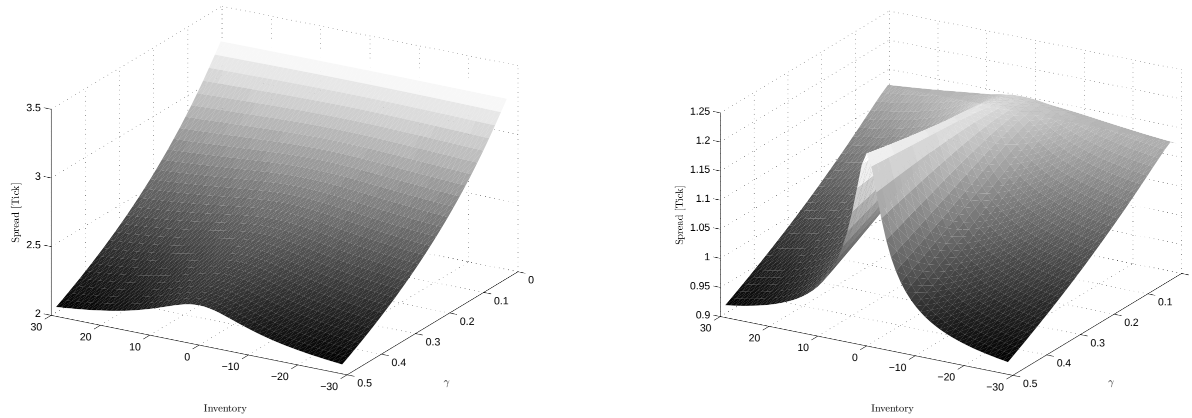


Figure 6: Asymptotic optimal quotes and bid-ask spread for different inventories and different values for the risk aversion parameter γ. Left: $\sigma = 0.3 \quad \text{Tick} \cdot \text{s}^{-1/2}$, $A = 0.9 \quad \text{s}^{-1}$, $k = 0.3 \quad \text{Tick}^{-1}$, $T = 600 \quad s$. Right: $\sigma = 0.6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

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, a model with a penalization term for not having cleared one's inventory and a model with inventory constraints from which all the figures have been drawn.
For each model, we enounce the associated results and some specific points are proved in the appendix. However, the general proofs are not repeated since they can be derived from adaptations of the proofs of the initial model.

5.1 Trend in prices

In the preceding setting, we supposed that the mid-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's consider a family of functions $(v_q)_{q\in\mathbb{Z}}$ solution of the linear system of ODEs that follows:*

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

with $v_q(T) = 1$, 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 can be expressed as:

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

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

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,

$$\lim_{T\to\infty} s - s^{b*}(0,q,s) = \frac{1}{\gamma}\ln\left(1+\frac{\gamma}{k}\right) - \frac{1}{k}\ln\left(\frac{f^0_{q+1}}{f^0_q}\right)$$

$$\lim_{T\to\infty} s^{a*}(0,q,s) - s = \frac{1}{\gamma}\ln\left(1+\frac{\gamma}{k}\right) + \frac{1}{k}\ln\left(\frac{f^0_{q}}{f^0_{q-1}}\right)$$

$$\lim_{T\to\infty} \psi^*(0,q) = -\frac{1}{k}\ln\left(\frac{f^0_{q+1}f^0_{q-1}}{f^{0\,2}_q}\right) + \frac{2}{\gamma}\ln\left(1+\frac{\gamma}{k}\right)$$

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

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

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

$$\delta^{b*}_\infty(q) \simeq \frac{1}{\gamma}\ln\left(1+\frac{\gamma}{k}\right) + \left[-\frac{\mu}{\gamma\sigma^2} + \frac{2q+1}{2}\right]\sqrt{\frac{\sigma^2\gamma}{2kA}\left(1+\frac{\gamma}{k}\right)^{1+\frac{k}{\gamma}}}$$

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

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

5.2 Inventory liquidation below mid price

In the initial model we imposed a terminal condition based on the assumption that the market maker liquidates his inventory at mid-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:

$$\sup_{S^a,S^b} \mathbb{E}\left[-\exp\left(-\gamma(X_T + q_T S_T - \phi(|q_T|))\right)\right]$$

where $\phi(\cdot) \geq 0$ is an increasing function with $\phi(0) = 0$ that represents the penalization term modeling the incurred cost at the end of the period for not having cleared the inventory[5].

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's consider a family of functions $(v_q)_{q\in\mathbb{Z}}$ solution of the linear system of ODEs that follows:*

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

[5]For analytical reasons we supposed that this penalization term does not depend on S_T. However, nothing prevents this penalization term to depend on price through S_0 for instance, a rather acceptable hypothesis in a short horizon perspective.

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

Then, optimal quotes can be expressed as:

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

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

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 term $\phi(|q|)$ *and is the same as in the initial model.*

5.3 Introduction of inventory constraints

Another possible setting is to consider explicitly in the model that the market maker cannot have too large an inventory. This is interesting by itself but it also provides numerical methods to solve the problem and all the graphs presented above have been made using this model that approximates the general one when the inventory limits are large.

5.3.1 The model

In this model, we introduce limits on the inventory. This means that once the agent holds a certain amount Q of shares, he does not propose an ask quote until he sells some of his shares. Symmetrically, once the agent is short of Q shares, he does not short sell anymore before he buys a share.

In modeling terms, it means that the Hamilton-Jacobi-Bellman equation of the problem is the following:

$\forall q \in \{-(Q-1), \dots, 0, \dots, Q-1\}$,

$$0 = \partial_t u(t,x,q,s) + \frac{1}{2}\sigma^2 \partial^2_{ss} u(t,x,q,s)$$

$$+ \sup_{s^b} \lambda^b(s^b,s)\left[u(t,x-s^b,q+1,s) - u(t,x,q,s)\right]$$

$$+ \sup_{s^a} \lambda^a(s^a,s)\left[u(t,x+s^a,q-1,s) - u(t,x,q,s)\right]$$

for $q = Q$ we have:

$$0 = \partial_t u(t,x,Q,s) + \frac{1}{2}\sigma^2 \partial^2_{ss} u(t,x,Q,s)$$

$$+ \sup_{s^a} \lambda^a(s^a,s)\left[u(t,x+s^a,Q-1,s) - u(t,x,Q,s)\right]$$

and symmetrically, for $q = -Q$ we have:

$$0 = \partial_t u(t,x,-Q,s) + \frac{1}{2}\sigma^2 \partial^2_{ss} u(t,x,-Q,s)$$

$$+ \sup_{s^b} \lambda^b(s^b,s)\left[u(t,x-s^b,-Q+1,s) - u(t,x,-Q,s)\right]$$

with the final condition:

$$\forall q \in \{-Q, \dots, 0, \dots, Q\}, \quad u(T, x, q, s) = -\exp\left(-\gamma(x + qs)\right)$$

As in the initial model we can reduce it to a linear system of ODEs. However the linear system associated to this model will be simpler since it involves $2Q+1$ equations only.

Proposition 5 (Resolution with inventory limits)**.** *Let's 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

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

the optimal quotes are:

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

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

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

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

Moreover, the asymptotic quotes and bid-ask spread can be expressed as:

$$\delta_\infty^{b*}(q) = \frac{1}{\gamma}\ln\left(1 + \frac{\gamma}{k}\right) - \frac{1}{k}\ln\left(\frac{f^0_{q+1}}{f^0_q}\right) \qquad \delta_\infty^{a*}(q) = \frac{1}{\gamma}\ln\left(1 + \frac{\gamma}{k}\right) + \frac{1}{k}\ln\left(\frac{f^0_q}{f^0_{q-1}}\right)$$

and

$$\psi^*_\infty(q) = -\frac{1}{k}\ln\left(\frac{f^0_{q+1}f^0_{q-1}}{f^{0\,2}_q}\right) + \frac{2}{\gamma}\ln\left(1 + \frac{\gamma}{k}\right)$$

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

5.3.2 Application to numerical resolution

This model based on a slight modification of the initial one leads to a system of linear ODEs whose associated matrix is solely tridiagonal. Hence, for all numerical resolutions we considered this modified problem with the inventory limit Q large enough and the numerical resolution simply boiled down to exponentiate a tridiagonal matrix. The rationale behind this is that the variant of the model under consideration imposes $v_q(t) = 0$ for $|q| > Q$ and for all times. Since the solution of the initial problem is in $\ell^2(\mathbb{Z})$ for $t < T$, this approximation will be valid when Q is large as long as t is far enough from the terminal time T and q not too close to Q.

Another possible method is to compute an eigenvector associated to the smallest eigenvalue of M. As we noticed before, we expect f^0_q to behave as $\exp\left(-\frac{1}{2}\sqrt{\frac{\alpha}{\eta}}q^2\right)$.

Hence it's a better idea to look for g^0 instead of f^0 where $g^0_q = f^0_q \exp\left(\frac{1}{2}\sqrt{\frac{\alpha}{\eta}}q^2\right)$. To this purpose, we replace the spectral analysis of M by the spectral analysis of the tridiagonal matrix DMD^{-1} where D is a diagonal matrix whose terms are $\left(\exp\left(\frac{1}{2}\sqrt{\frac{\alpha}{\eta}}q^2\right)\right)_{q\in\{-Q,\ldots,Q\}}$ and g^0 will be an eigenvector associated to the smallest eigenvalue of DMD^{-1}.

Now, once g^0 has been calculated, the asymptotic values of the optimal quotes and the bid-ask spread are:

$$\delta^{b*}_\infty(q) = \frac{1}{\gamma}\ln\left(1+\frac{\gamma}{k}\right) - \frac{1}{k}\ln\left(\frac{g^0_{q+1}}{g^0_q}\right) + \frac{1}{2k}\sqrt{\frac{\alpha}{\eta}}(2q+1)$$

$$\delta^{a*}_\infty(q) = \frac{1}{\gamma}\ln\left(1+\frac{\gamma}{k}\right) + \frac{1}{k}\ln\left(\frac{g^0_{q}}{g^0_{q-1}}\right) - \frac{1}{2k}\sqrt{\frac{\alpha}{\eta}}(2q-1)$$

and

$$\psi^*_\infty(q) = \frac{2}{\gamma}\ln\left(1+\frac{\gamma}{k}\right) - \frac{1}{k}\ln\left(\frac{g^0_{q+1}g^0_{q-1}}{g^{0\,2}_q}\right) + \frac{1}{k}\sqrt{\frac{\alpha}{\eta}}$$

6 Discussion on the model

6.1 Exogenous nature of prices

In our model, the mid-price is 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 mid-price must be clarified. Indeed, one may consider that, although it has little impact on the market, the market maker can put an order inside the bid-ask spread of the market order book and hence change the mid-price. This would be a misunderstanding of the model since the mid-price is to be considered in the model before any insertion of an order. Hence, the mid-price in this model must be understood as the mid-price of an order book in which our market maker's orders would be removed. More generally, it may be viewed as a generalized mid-price calculated across trading facilities or any reference price for which the hypothesis on orders arrival is a good approximation of reality.

6.2 Dependence on price

What may be counterintuitive at first sight is that the bid-ask spread or any of the two spreads between quoted prices and market mid-price seems not to depend on the mid-price itself. In fact, in our model, the bid-ask spread and the gap between quoted prices and the market mid-price depends on price, though indirectly, through parameters. Prices are indeed hidden in the trading intensity λ, and more specifically into the parameter k. We indeed considered a trading intensity depending on the distance between the quoted prices and the mid-price. Thus, k must depend indirectly on prices to normalize prices differences.

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:

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

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 price[6], 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.

6.4 Constant parameters

Another 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.

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 have to be made to the model and the way backtests have been carried out. Then we present the result on the French stock AXA.

[6]If 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.

7.1 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 easily 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.

7.2 Results

To present the results, we chose to illustrate the case of the French stock AXA on November 2^{nd} 2010.
We first show the evolution of the inventory and we see that this inventory mean-reverts around 0.

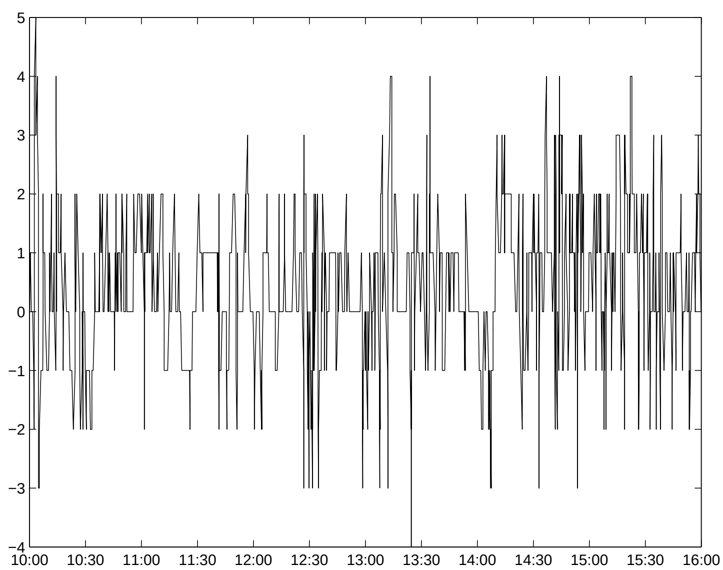


Figure 7: Inventory when the strategy is used on AXA (02/11/2010) from 10:00 to 16:00 with $\gamma = 0.05 \quad \text{Tick}^{-1}$

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 and we can see on the following plot the corresponding evolution of the inventory.

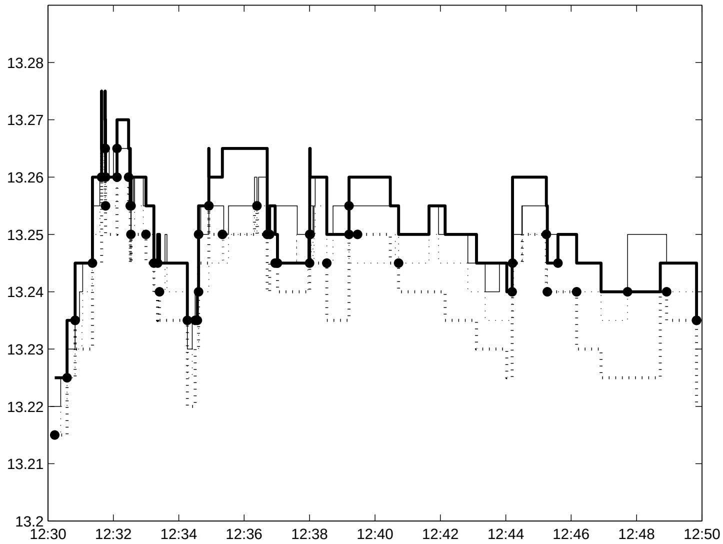


Figure 8: Details for the quotes and trades when the strategy is used on AXA (02/11/2010) with $\gamma = 0.05 \quad \text{Tick}^{-1}$. Thin lines represent the market while bold lines represent the quotes of the market maker. Dotted lines are associated to the bid side while plain lines are associated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

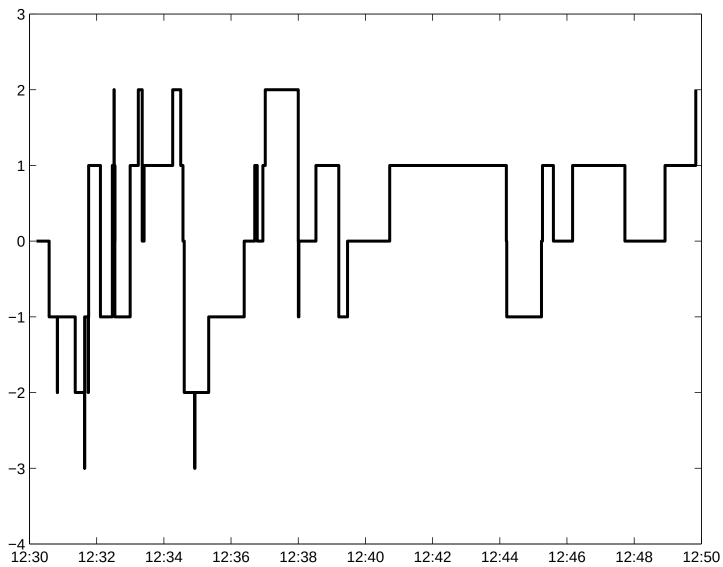


Figure 9: Details for the inventory when the strategy is used on AXA (02/11/2010) with $\gamma = 0.05 \quad \text{Tick}^{-1}$

Conclusion

In this paper we presented a model for the optimal quotes of a market maker. Starting with the model by Avellaneda ans Stoikov [5] we introduced a change of variables[7] that allows to find semi-explicit 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.

[7] In 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:

This is the classical PDE representation of a stochastic control problem with jump processes. □

Proof of Proposition 2 and Theorem 2:

Let's consider a solution $(v_q)_q$ of $(\mathcal{S})$ and introduce $u(t,x,q,s) = -\exp\left(-\gamma(x+qs)\right) v_q(t)^{-\frac{\gamma}{k}}$.

Then:

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

Now, concerning the bid quote, we have:

$$\sup_{s^b} \lambda^b(s^b, s)\left[u(t, x - s^b, q+1, s) - u(t,x,q,s)\right]$$

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

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

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

Hence:

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

and

$$\sup_{s^b} \lambda^b(s^b, s)\left[u(t, x - s^b, q+1, s) - u(t,x,q,s)\right]$$

$$= -\frac{\gamma}{k+\gamma} A \exp(-k(s - s^{b*})) u(t,x,q,s)$$

$$= -\frac{\gamma A}{k+\gamma}\left(1 + \frac{\gamma}{k}\right)^{-\frac{k}{\gamma}} \frac{v_{q+1}(t)}{v_q(t)} u(t,x,q,s)$$

Similarly,

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

and

$$\sup_{s^a} \lambda^a(s^a, s)\left[u(t, x + s^a, q-1, s) - u(t,x,q,s)\right]$$

$$= -\frac{\gamma}{k+\gamma} A \exp(-k(s^{a*} - s)) u(t,x,q,s)$$

$$= -\frac{\gamma A}{k+\gamma}\left(1 + \frac{\gamma}{k}\right)^{-\frac{k}{\gamma}} \frac{v_{q-1}(t)}{v_q(t)} u(t,x,q,s)$$

Hence, putting the three terms together we get:

$$\partial_t u(t,x,q,s) + \frac{1}{2}\sigma^2 \partial^2_{ss} u(t,x,q,s)$$

$$+\sup_{s^b} \lambda^b(s^b,s)\left[u(t,x-s^b,q+1,s) - u(t,x,q,s)\right]$$

$$+\sup_{s^a} \lambda^a(s^a,s)\left[u(t,x+s^a,q-1,s) - u(t,x,q,s)\right]$$

$$= -\frac{\gamma}{k}\frac{\dot{v}_q(t)}{v_q(t)}u + \frac{\gamma^2\sigma^2}{2}q^2 u - \frac{\gamma A}{k+\gamma}\left(1+\frac{\gamma}{k}\right)^{\frac{k}{\gamma}}\left[\frac{v_{q+1}(t)}{v_q(t)} + \frac{v_{q-1}(t)}{v_q(t)}\right]u$$

$$= -\frac{\gamma}{k}\frac{u}{v_q(t)}\left[\dot{v}_q(t) - \frac{k\gamma\sigma^2}{2}q^2 v_q(t) + A\left(1+\frac{\gamma}{k}\right)^{-\left(1+\frac{k}{\gamma}\right)}(v_{q+1}(t)+v_{q-1}(t))\right] = 0$$

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

Proof of Theorem 1 and Theorem 3:

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

Let's 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.

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

Proof:

$$\forall w \in H, \|w\|^2_{\ell^2(\mathbb{Z})} = |w_0|^2 + \sum_{q\in\mathbb{Z}^*} |w_q|^2 \le |w_0|^2 + \frac{1}{\alpha}\|w\|^2_H$$

$$\le (|w_1| + |w_0 - w_1|)^2 + \frac{1}{\alpha}\|w\|^2_H \le \left[\left(\sqrt{\frac{1}{\alpha}} + \sqrt{\frac{1}{\eta}}\right)^2 + \frac{1}{\alpha}\right]\|w\|^2_H$$

□

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:

$$\hat{s}^k_0 = s^k_0, \qquad \forall q \in \mathbb{Z}^*, \hat{s}^k_q = q s^k_q$$

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 $\hat{s}^\infty$ in $\ell^2(\mathbb{Z})$.
Now, we can define $s^\infty \in H$ by the inverse transformation:

$$s^\infty_0 = \hat{s}^\infty_0, \qquad \forall q \in \mathbb{Z}^*, s^\infty_q = \frac{1}{q}\hat{s}^\infty_q$$

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

$$\|s^{k_j} - s^\infty\|^2_{\ell^2(\mathbb{Z})} \le |s^{k_j}_0 - s^\infty_0|^2 + \sum_{q\in\mathbb{Z}^*} \frac{1}{q^2}|\hat{s}^{k_j}_q - \hat{s}^\infty_q|^2$$

$$= |\hat{s}^{k_j}_0 - \hat{s}^\infty_0|^2 + \sum_{q\neq 0, |q|\le N} \frac{1}{q^2}|\hat{s}^{k_j}_q - \hat{s}^\infty_q|^2 + \sum_{|q|>N} \frac{1}{q^2}|\hat{s}^{k_j}_q - \hat{s}^\infty_q|^2$$

The first two terms tend to 0 because $(\hat{s}^{k_j})_{\mathbb{N}}$ weakly converges in $\ell^2(\mathbb{Z})$ towards $\hat{s}^\infty$. 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 $(\|\hat{s}^{k_j} - \hat{s}^\infty\|_{\ell^2(\mathbb{N})})_{j\in\mathbb{N}}$ is a bounded sequence. □

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's 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. □

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 eigenvalue[8] 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 $v(t) = \sum_{k\in\mathbb{N}} \mu^k(t) f^k$.

[8] 0 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 $\ell^2(\mathbb{Z})$, $f^k = 0$.

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

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

We must have $\frac{d\mu^k}{dt}(t) = (\frac{1}{\lambda^k} - 2\eta)\mu^k(t)$ and hence, since $\lambda^k \to 0$ 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 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

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

Hence, $\forall k \in \mathbb{Z}, \langle f^k, v(T)\rangle_{\ell^2(\mathbb{Z})} = 0 \implies \forall k \in \mathbb{Z}, \forall t \in [0,T], \quad \langle f^k, v(t)\rangle_{\ell^2(\mathbb{Z})} = 0$ 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 $\forall q \in \mathbb{Z}, v_q(0) \underset{T\to\infty}{\sim} \langle 1, f^0\rangle_{H',H} f^0_q \exp\left((2\eta - \frac{1}{\lambda^0})T\right)$ so that the result is proved with

$$\delta^{b*}_\infty(q) = \frac{1}{\gamma}\ln\left(1+\frac{\gamma}{k}\right) - \frac{1}{k}\ln\left(\frac{f^0_{q+1}}{f^0_q}\right) \qquad \delta^{a*}_\infty(q) = \frac{1}{\gamma}\ln\left(1+\frac{\gamma}{k}\right) + \frac{1}{k}\ln\left(\frac{f^0_q}{f^0_{q-1}}\right)$$

and

$$\psi^*_\infty(q) = -\frac{1}{k}\ln\left(\frac{f^0_{q+1}f^0_{q-1}}{f^{0\,2}_q}\right) + \frac{2}{\gamma}\ln\left(1+\frac{\gamma}{k}\right)$$

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's consider the following characterization of λ^0 and ofthe 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's 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. □

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:

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

This representation guarantees that $v > 0$.

This ends the proof of Theorems 1 and 3 □

Proof of Proposition 3:

If $dS_t = \mu dt + \sigma dW_t$ then the (HJB) equation becomes:

$$\text{(HJB)} \qquad 0 = \partial_t u(t,x,q,s) + \mu \partial_s u(t,x,q,s) + \frac{1}{2}\sigma^2 \partial^2_{ss} u(t,x,q,s)$$

$$+ \sup_{s^b} \lambda^b(s^b,s) \left[u(t,x-s^b,q+1,s) - u(t,x,q,s) \right]$$

$$+ \sup_{s^a} \lambda^a(s^a,s) \left[u(t,x+s^a,q-1,s) - u(t,x,q,s) \right]$$

Using the same change of variables as in the proof of Proposition 2, we can consider a family of strictly positive functions $(v_q)_{q\in\mathbb{Z}}$ is solution of the linear system of ODEs that follows:

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

with $v_q(T) = 1$, $\alpha = \frac{k}{2}\gamma\sigma^2$, $\beta = k\mu$ and $\eta = A(1+\frac{\gamma}{k})^{-(1+\frac{k}{\gamma})}$.
Then $u(t,x,q,s) = -\exp\left(-\gamma(x+qs)\right) v_q(t)^{-\frac{\gamma}{k}}$ is a solution of (HJB) and the final condition is satisfied.

Also, as in the proof of Proposition 2, the optimal quotes are given by:

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

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

and the bid-ask spread follows straightforwardly.

To prove the counterparts of Theorem 1 and Theorem 3, we have to write:

$$\begin{aligned} \dot{v}_q(t) &= (\alpha q^2 - \beta q) v_q(t) - \eta \left(v_{q-1}(t) + v_{q+1}(t) \right) \\ &= \left[\alpha \left(q - \frac{\beta}{2\alpha} \right)^2 v_q(t) - \eta \left(v_{q-1}(t) - 2v_q(t) + v_{q+1}(t) \right) \right] - \left(2\eta + \frac{\beta^2}{4\alpha} \right) v_q(t) \end{aligned}$$

Hence the operator L and the Hilbert space H are modified but the results are the same *mutatis mutandis.* □

Proof of Proposition 4:

In this setting, the Bellman function is defined by:

$$u(t,x,q,s) = \sup_{S^a,S^b} \mathbb{E}\left[-\exp\left(-\gamma(X_T + q_T S_T - \phi(q_T))\right) \middle| X_t = x, q_t = q, S_t = s \right]$$

The (HJB) equation is:

$$\text{(HJB)} \qquad 0 = \partial_t u(t,x,q,s) + \frac{1}{2}\sigma^2 \partial^2_{ss} u(t,x,q,s)$$

$$+ \sup_{s^b} \lambda^b(s^b,s) \left[u(t,x-s^b,q+1,s) - u(t,x,q,s) \right]$$

$$+ \sup_{s^a} \lambda^a(s^a,s) \left[u(t,x+s^a,q-1,s) - u(t,x,q,s) \right]$$

with $u(T,x,q,s) = -\exp\left(-\gamma(x+qs-\phi(q))\right)$.

Using the same change of variables as in the proof of Proposition 2, we can consider a family of strictly positive functions $(v_q)_{q\in\mathbb{Z}}$ is solution of the linear system of ODEs that follows:

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

with $v_q(T) = e^{-k\phi(q)}$, $\alpha = \frac{k}{2}\gamma\sigma^2$ and $\eta = A(1+\frac{\gamma}{k})^{-(1+\frac{k}{\gamma})}$.
Then $u(t,x,q,s) = -\exp\left(-\gamma(x+qs)\right) v_q(t)^{-\frac{\gamma}{k}}$ is a solution of (HJB) and the terminal condition is satisfied.

Also, as in the proof of Proposition 2, the optimal quotes are given by:

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

and the bid-ask spread follows straightforwardly. □

Proof of Proposition 5:

The demonstration is the same as for Proposition 2, with the same change of variables. The only difference here is that the system of ODEs for $(v_q)_q$ can be solved explicitly using a matrix exponentiation. The characterization of f^0 is the same as in Proposition 2 once it is noticed that the counterpart of L is now $(M+2\eta I)^{-1}$. □

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. 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.

[16] T. S. Y. Ho and H. R. Stoll. The dynamics of dealer markets under competition. *Journal of Finance*, 38(4):1053–1074, 1983.

[17] Thomas Ho and Hans R. Stoll. Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial Economics*, 9(1):47–73, March 1981.

[18] Charles-Albert Lehalle, Olivier Guéant, and Julien Razafinimanana. 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.

[19] Albert J. Menkveld. High Frequency Trading and The New-Market Makers. *Social Science Research Network Working Paper Series*, December 2010.

[20] E. Mildenstein and H. Schleef. The optimal pricing policy of a monopolistic market-maker in the equity market. *Journal of Finance*, pages 218–231, 1983.

[21] M. O'Hara and G. S. Oldfield. The microeconomics of market making. *Journal of Financial and Quantitative Analysis*, 21(04):361–376, 1986.

[22] Gilles Pagès, Sophie Laruelle, and Charles-Albert Lehalle. Optimal split of orders across liquidity pools: a stochatic algorithm approach. Technical report, 2009.

[23] R. Roll. A simple implicit measure of the effective bid-ask spread in an efficient market. *Journal of Finance*, 39(4):1127–1139, 1984.