Price Manipulation and American Options in Illiquid Markets

Alexandre F. Roch

Research fellow, Caisse de dépôt et placement du Québec Research Chair in Portfolio Management

ESG UQAM

Finance Department

Montreal, CANADA

roch.alexandre_f@uqam.ca

June 12, 2012

Abstract

A setup in which a large trader has sold a certain number of American-type derivatives is considered. The large trader's trades are assumed to have an impact on prices so that he may be tempted to minimize the payoff of the derivative by manipulating the underlying asset. However, the option holders have the right to exercise the option, turning the pricing problem in a two-player stochastic game. It is shown that the solution of this optimization problem can be described as the solution of a double obstacle variational inequality and the optimal strategy for the large trader and the optimal exercise time for the option holder are obtained. The fact that price manipulation may force the option holders to exercise before the maturity implies that regulators must monitor stock prices and options sellers activity continuously. Our theory gives us the precise time at which a market manipulator will start manipulating prices. We also calculate a maximum number of options that can be sold so that no manipulation strategies are possible in terms of quantities such as the bid-ask spread and the price impact factor. We conclude with a sensitivity analysis in which we compare the timing and size of manipulation interventions by the large trader as well as the optimal exercise region for the options holders for different levels of liquidity.

KEYWORDS: American Options, Price Manipulation, Liquidity and Price Impacts, Two-player Stochastic Game

1 Introduction

This study concerns the manipulation of American-type options in a setup in which the seller of the option is large enough for his trades to have an impact on the price. This investor is dubbed 'the large trader'. By setting up an optimization problem, we characterize situations in which it would be optimal for the large trader to manipulate the price of the underlying asset in order to lower the value of the options he has sold. These types of trading strategies would typically be considered illegal price manipulations. Indeed, Kyle and Viswanathan (2008) define an *illegal price manipulation* as a trading strategy for which the trader's intent is "to pursue a scheme that undermines economic efficiency both by making prices less accurate as signals for efficient resource allocation and by making markets less liquid for risk transfer." We assume that the other market participants, in particular the option holders, are unaware of the large trader's intent to manipulate prices so that his trades are construed as genuine demand or supply of liquidity and have an impact on equilibrium prices. Since the goal of an option seller who manipulates prices is to lower the value of the options he has sold, it is clear that this falls under the above definition of price manipulation. This large trader is taking liquidity away from the market by making transactions that bear no informational value and that deliberately lowers the pricing accuracy of the market.

There is a large academic literature on the topic of price manipulations in derivative contracts. Most studies consider futures contracts, as in the works of Dutt and Harris (2005), Kumar and Seppi (1992), Pirrong (1993), (2001). The main manipulation strategy in this context is "punching the close" in which a large trader will drive up the settlement price of a cash-settled futures contract by buying large quantities of the underlying asset. The problem of price manipulations in relation to options markets has been less studied in the literature. Jarrow (1994) gives conditions under which price manipulations are not possible and computes option prices in this situation, and Horst and Naujokat (2011) solves the multi-player price manipulation optimization problem for the case of European-style options. There is a number of papers, for instance Bank and Baum (2004), Çetin et al. (2004), and Roch (2011), in which the hedging problem in illiquid markets is considered, however these types of strategies would typically not be considered illegal manipulation strategies as defined above. Furthermore, these authors only consider European-style options.

To date, there are no papers which have been written on the hedging or the manipulation of American-type options in illiquid markets due to the added difficulty of the optimal exercise time which may depend on the large trader's strategy, turning the optimization problem in a two-player stochastic game. Yet, this issue is particularly relevant to regulators given that single stocks on which options are written are typically less liquid than futures underlying securities such as bonds, stock indexes and commodities. Furthermore, as opposed to European-type options or futures, the fact that price manipulation may force the option holders to exercise at any time before the maturity implies that regulators must monitor stock prices and options sellers activity continuously. Our theory gives us the precise time at which a market manipulator will start manipulating prices in order to lower the value of American options. The manipulation strategy will depend on the number of options sold, the strike, the current level of liquidity and the volatility of the underlying asset.

A special case of stochastic games called Dynkin games, or two-player stochastic stopping games, have been studied from a theoretical perspective by Cvitanic and Karatzas (1996), Touzi and Vieille (2002), Kallsen and Kühn (2004) and Hamadène and Hassini (2005), to name a few. For more general stochastic games in which players can also control the state variables, the known results are typically restricted to settings in which players control the drift of the processes, as done by Friedman (1972) and Hamadène (2006), or both the drift and the volatility coefficient, as in the studies of Karatzas and Sudderth (2006) and Rainer (2007). However, in the setup in which we are interested, the large trader controls the size and the timing of the jumps. Thus, it is not possible to use these results in this setting. This work extends the known theoretical results to this special kind of stochastic controls, and applies the results to the pricing of American options.

The results in this paper bear resemblance to the pricing theory of American options in the Black-Scholes model. In the complete, price-taker setting of Black and Scholes, it has been well known since the works of Karatzas (1988) and Bensoussan (1984) that American-type derivatives can be described by a problem of optimal stopping such that the fair price of an option of maturity T is the supremum over all stopping times τ on $[0,T]$ of the discounted expected payoff at time τ under the risk neutral measure. In particular, the price is given by the Snell envelope of the discounted payoff process. From this, one can write a variational inequality, as done by Bensoussan and Lions (1978), from which it is possible to design a numerical scheme by finite differences methods. Similar methods are used in this paper.

Similar to the continuous auction model of Kyle (1985), or as in the large trader models of Roch (2011) and

Liu and Yong (2005), we assume that the observed underlying stock price process is of the form

$$dS_t \quad = \quad dS_t^0 + \lambda dX_t, \tag{1.1}$$

in which X is the trading strategy used by the large trader. The constant λ is inversely proportional to the market depth for this asset. (When λ increases, the market is less liquid) Here S^0 is an exogenously given stochastic process, representing the stochastic price process which *would have been observed if the large trader did not trade*. This price process is not observable since in general X will not be zero. S^0 includes the informational value of the asset and possibly the trading noise. On the other hand, the equilibrium price is affected by the trading of the large trader which bears no informational value and makes the asset price diverge from its equilibrium value. Consequently, S is the observed price process and includes the price impacts from the strategy X. The liquidity parameter λ is assumed to be positive and constant. The extension to a stochastic parameter λ is rather straightforward. However to keep the low dimensionality of the problem, it is assumed to be a constant in this work. The parameter gives the size of the impact of the trade dX_t at time t. We assume however that the large trader faces other liquidity costs and transaction costs, so that he may not freely manipulate prices. The derivatives are sold to a number of 'small traders' whose trades do not have an impact on prices. However, the option holders have the right to exercise the option at any point in time they desire. On the other hand, the large trader has an incentive to manipulate the price to make the option out-of-the-money, but due to market frictions he will only do so when the benefits outweigh the liquidity costs.

In our large trader setting, the large trader cannot affect prices continuously due to fixed transaction costs and liquidity constraints. As a result, his trading strategy has to be in the form of an impulse control, i.e

$$X_t \quad = \quad \sum_{i\geq 1} \xi_i \mathbf{1}_{\{\delta_i \leq t\}} \quad (t \geq 0) \tag{1.2}$$

with $(\delta_i)_{i\geq 1}$ a sequence of stopping times in $[0,T]$ and ξ_i is $\mathscr{F}_{\delta_i}-$ measurable for all $i \geq 1$. The goal is then to find the instants in time when the large trader will intervene and the size of the interventions.

In Section 2, the model and the assumptions are given and the optimization problem for the large trader is introduced in Section 3. In Section 4, the optimal strategies of the large trader and the option holders are obtained. Conditions to rule out manipulation opportunities are derived in section 5 and the maximum number of options that can be sold so that no manipulation strategies are possible in terms of quantities such as the bid-ask spread and the price impact factor is calculated. Section 6 concludes with a numerical scheme based on the associated variational inequality and some numerical results.

2 The model

Let $T > 0$ denote the maturity of a claim and let $(\Omega, \mathscr{F}, \mathbb{F} = (\mathscr{F}_t)_{0\le t\le T}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions and on which a Brownian motion W is defined.

The large trader has the possibility of trading the risky asset and has a money market account with constant risk-free rate, denoted by r. We assume that there is a liquidity cost associated to a transaction for the large trader as a function of the size of the transaction and the time. In other words, the price of x shares of the asset at time t is given by $xS^X_{t-} + \tilde{C}(x,t)$, in which $\tilde{C}(x,t)$ is a positive function for all x and t. On the other hand, the proceeds from the sale of x shares is given by $xS^X_{t-} - \tilde{C}(-x,t)$. In other words, the cash flow associated to a transaction of size x, regardless of the sign of x, is given by $-xS^X_{t-} - \tilde{C}(x,t)$. Note that at the time τ_n of the n-th transaction, the marginal price is not yet affected by the impact of this trade. As a result, the effective marginal price appearing in the above expression for the price of a transaction is given by $S^X_{t-} = S^0_t + \lambda X_{t-}$. The use of the left-limit is simply a convention. Since the difference between S^X_t and S^X_{t-} is a function of ΔX_t, the price of x shares could be written in terms of S^X_t by appropriately modifying the cost function $\tilde{C}(x,t)$. A typical cost function is of the form $\tilde{C}(\xi,t) = c + a|\xi| + \frac{1}{2}b|\xi|^2$. Here, c can be interpreted as a fixed transaction cost, a as one half of the bid-ask spread and b the liquidity premium, i.e. the marginal cost per share for the purchase of x shares of the asset.

Let $\varepsilon > 0$ and denote by $\mathscr{T}_t$ the set of $\mathbb{F}$-stopping time with values between t and T, and $\mathbb{T}^\varepsilon_t = \{(\delta_i)_{i\ge1} : \delta_i \in \mathscr{T}_t$ for each $i \ge 1$, such that $\delta_{i+1} \wedge T \ge (\delta_i + \varepsilon) \wedge T$ for all $i \ge 1\}$ the set of ordered sequences of stopping times in the time interval $[t,T]$ with lag time of at least ε between two stopping times, up to time T. Let A be a finite subset of $\mathbb{R}$. For $t_0 \in [0,T]$, the set of admissible strategies for the large trader is defined as

$$\begin{aligned}\mathscr{A}_{t_0} \quad = \quad & \{(X_t)_{t\ge t_0} | X_t = \sum_{i\ge1} \xi_i \mathbf{1}_{\{t_0\le\delta_i\le t\}} \text{ in which } (\delta_i)_{i\ge1} \in \mathbb{T}^\varepsilon_{t_0,T} \\ & \qquad\qquad \text{and } \xi_i \text{ is } \mathscr{F}_{\delta_i}-\text{measurable such that } \xi_i \in A \text{ for each } i \ge 1\}.\end{aligned}$$

If $X \in \mathscr{A}$, then $\Delta X_t = \xi_i$ on $\{t = \delta_i\}$. Thus, ξ_i denotes the size of the trade by the large trader at time δ_i. Equivalently, X_t is the number of shares of the risky asset owned by the large trader at time t. Each pair (ξ_i, δ_i) is called an intervention. The set A gives the transaction sizes that are possible.

We have constrained the transactions of the large trader to be spaced in time by at least ε units of time to limit the ease with which he can manipulate prices. Indeed, when the cost function $\tilde{C}$ has higher liquidity terms (quadratic terms, for instance), it is cheaper to divide a large transaction into smaller ones and pass them rapidly in time. The possibility to make too many transactions in a given time period would violate the idea that liquidity is not infinite. The ε units of time can then be interpreted as the time required for liquidity to come back to the market after a large transaction.

Let $s > 0$ and $t_0 \in [0,T]$. The unaffected price process S^0 started at $S^0_{t_0} = s$ at time t_0 is defined by

$$dS^0_t \quad = \quad \mu(S^0_t,t)dt + \sigma(S^0_t,t)dW_t \quad (t_0 \le t \le T).$$

The functions μ and σ are Lipschitz in the first argument, uniformly over the second argument, so that the solution of the above SDE exists and is unique. Furthermore, the ratio $\frac{\mu(s,t)-rs}{\sigma(s,t)}$ is assumed to be bounded, and $\sigma(0,t) = 0$ and $\sigma(s,t) > 0$ for $s > 0$ and $t \in [0,T]$. The Black and Scholes model is obtained by letting $\mu(s,t) = \mu s$ and $\sigma(s,t) = \sigma s$ for some positive constants μ and σ. The unaffected price process is the price process that would be observed if the large trader did not trade at all. It does not include the price impacts of the large trader. However, the actual observed price process does and will be defined in terms of the unaffected price process. Let $\mathbb{Q}$ be the equivalent measure under which $B_t = W_t + \int_{t_0}^t \frac{\mu(S^0_u,u)-rS^0_u}{\sigma(S^0_u,u)}du$ is a Brownian motion. All expectations from now on are with respect to $\mathbb{Q}$. Under $\mathbb{Q}$ the discounted unaffected price process, given by $e^{-rt}S^0_t$ is a martingale.

For a strategy $X \in \mathscr{A}_{t_0}$, the *actual observed marginal price process* of the asset is defined as

$$S^X_t = S^0_t + \lambda X_t \quad (t_0 \le t \le T),$$

with $S^X_{t_0} = S^0_{t_0} = s$. More specifically, S^X_t represents the observed marginal price after the trade at time t when there is one.

The large trader's strategy is self-financing, so that the value of the money market account at time T associated to a strategy $X \in \mathscr{A}_t$ for the large trader is defined by

$$Y_T = Y_0 + \int_0^T rY_t dt - \sum_{i\geq 1} \left(\xi_i S^X_{\delta_i -} + \tilde{C}(\xi_i, \delta_i) \right).$$

We can write

$$\begin{aligned} Y_T &= Y_0 + \int_0^T rY_t dt - \sum_{i\geq 1} \left(\xi_i (S^0_{\delta_i} + \lambda X_{\delta_i -}) + \tilde{C}(\xi_i, \delta_i) \right) \\ &= Y_0 + X_0 S_0 + \int_0^T rY_t dt - X_T S^0_T + \int_0^T X_{t-} dS^0_t - \sum_{i\geq 1} \left(\lambda X_{\delta_{i-1}} \Delta X_{\delta_i} + \tilde{C}(\xi_i, \delta_i) \right), \end{aligned}$$

by integration by parts, and the fact that $X_{\delta_i -} = X_{\delta_{i-1}}$ and $\xi_i = X_{\delta_i} - X_{\delta_{i-1}}$. The wealth value W_T of the portfolio at time T is then defined as the sum of the money market account Y_T and the liquidation value of the risky asset, given by $X_T S^0_T - \tilde{C}(-X_T, T)$. It is equal to

$$\begin{aligned} W_T &= Y_0 + X_0 S_0 + \int_0^T rY_t dt + \int_0^T X_{t-} dS^0_t - \tilde{C}(-X_T, T) - \sum_{i\geq 1} \left(\lambda X_{\delta_{i-1}} \Delta X_{\delta_i} + \tilde{C}(\xi_i, \delta_i) \right) \\ &= W_0 + \int_0^T rY_t dt + \int_0^T X_{t-} dS^0_t - C(-X_T, T) - \sum_{i\geq 1} C(\xi_i, \delta_i), \end{aligned} \tag{2.1}$$

also by integration by parts. We also used the fact that $X_0 = 0$ so that $W_0 = Y_0$. In the above expression, $C(\xi, t) = -\frac{1}{2}\lambda\xi^2 + \tilde{C}(\xi, t)$. We call C the effective transaction cost function. It takes into account both the true transaction cost and the effect of price impacts on future prices. The function $\tilde{C}$ should be chosen so that C is positive. This implies that the liquidity cost $\tilde{C}(x, t)$ of moving the stock price by $\lambda\xi$ units from a trade of size ξ should be at least equal to $\frac{1}{2}\lambda\xi^2$ to rule out possible arbitrage opportunities. For instance, in the liquidity risk model of Roch (2011), the effective liquidity cost function is of the form $\frac{1}{2}(1-\Lambda)\lambda\xi^2$, for some positive parameter $\Lambda \leq 1$. For more mathematical tractability, we ignore the transaction cost associated to liquidating the portfolio at time T, i.e. we assume that $C(x, T) = 0$ for all x. This implies that $\tilde{C}(x, T) = \frac{1}{2}\lambda x^2$. Although this is not necessary, it simplifies the problem by reducing the number of state variables by one. We further assume that C is continuous and $\inf_{\xi \in A, t\in[0,T)} C(\xi, t) > 0$. One can think of this infimum as a minimum fixed transaction cost paid at each trade. The fact that this cost is strictly positive implies that the large trader will be restricted from turning around a position at no cost and manipulating prices freely.

Throughout, we use the convention that the infimum of an empty set is equal to T, i.e. $\inf\emptyset = T$.

3 The manipulation problem

We now define the manipulation problem of the options seller as an optimization problem. Suppose a large trader has sold n_0 units of an American-type derivative with payoff function h^0 with maturity T. The payoff function can be equal to $h^0(s) = (s-K)_+$ for a call option, or $h^0(s) = (K-s)_+$ for a put option with K the strike. However, for more generality, h^0 is only assumed to be continuous. We let $h(s,t) = n_0 \times e^{-rt} h^0(s)$ denote the total discounted payoff function associated to the n_0 units sold of the derivative with payoff function h^0.

The final wealth of the large trader is then given by $W_T - n_0 h^0(S^X_{\tau-})$ if X is the chosen strategy and τ the option holder's exercise time.

Remark 3.1 *In the event that the large trader and the option holders intervene at the same time, i.e. $\tau = \delta_i$ for some $i \geq 1$, then the option holders have the priority in the sense that the payoff $h^0(S^X_{\tau-})$ is calculated with the stock price excluding the intervention of the large trader at the exercise time. In particular, an intervention at time T or after will have no consequence. (Because of the convention* $\inf\emptyset = T$*, we can say that the strategy $\delta_i = T$ for all i corresponds to the strategy of no intervention.)*

We assume that the goal of the option seller is to maximize the discounted risk neutral value of his final wealth, i.e. minimize

$$\mathbb{E}_{s,t}\left(-W^*_T + h(S^X_{\tau-},\tau)\right).$$

This expression is equal to

$$\mathbb{E}_{s,t}\left(h(S^X_{\tau-},\tau) + \sum_{i\geq 1} C(\xi_i,\delta_i)\mathbf{1}_{\{\delta_i<T\}}\right),$$

using the martingale property of the integral with respect to $S^{0,*}$ in the expression for W^*_T. Here $\mathbb{E}_{s,t}$ denotes the conditional expectation given $S_t^X = S_t^0 = s$. Throughout, we use the notation

$$J(X,\tau) = h(S_{\tau-}^X, \tau) + \sum_{i\geq 1} C(\xi_i, \delta_i)\mathbf{1}_{\{t\leq\delta_i<\tau\}}.$$

However, the option holders will exercise at the moment τ that maximizes their payoff. The optimization problem can then be written as

$$\inf_{X\in\mathscr{A}_t} \sup_{\tau\in\mathscr{B}_t} \mathbb{E}_{s,t}\left(J(X,\tau)\right).$$

We make the following definition:

Definition 3.2 *Let $s > 0$ and $0 \leq t \leq t' \leq T$. We denote by $q(s,t,t')$ the value of the optimization problem, defined by*

$$q(s,t,t') \quad = \quad \inf_{X\in\mathscr{A}_{t'}} \sup_{\tau\in\mathscr{B}_t} \mathbb{E}_{s,t}\left(J(X,\tau)\right).$$

We also consider the following related optimization problem:

$$u(s,t,t') \quad = \quad \sup_{\tau\in\mathscr{B}_{t'}} \inf_{X\in\mathscr{A}_t} \mathbb{E}_{s,t}\left(J(X,\tau)\right).$$

We write $q(s,t)$ instead of $q(s,t,t)$ when $t' = t$.

It is shown in the following section that these two functions coincide and that the optimal exercise time for the option holders is given by the first time the value of the optimization problem equals the exercise value, as in the classical price taker setting. Similarly, we obtain the large trader's manipulation strategy from a dynamic programming principle. We show that the intervention times can be computed in terms of the value function q and the price process S^0.

3.1 The no arbitrage condition

If there are arbitrage opportunities, the stated optimization problem will make no sense. Fortunately, in this price impact setup there is a well known condition that rules out arbitrage opportunities.

Theorem 3.3 *If there exists an equivalent measure $\mathbb{Q}$ under which the discounted* unaffected *process defined as $S_t^{0,*} = e^{-rt} S_t^0$ is a martingale, then there are no arbitrage opportunities.*

Proof: From Equation 2.1, we deduce that the discounted value of the terminal wealth can be written as

$$e^{-rT} W_T = W_0 + \int_t^T X_{u-}(dS_u^0 - rS_u^0 du) - \sum_{i \geq 1} C(\xi_i, \delta_i) e^{-r\delta_i} \tag{3.1}$$

$$W_T^* = W_0^* + \int_t^T X_{u-} dS_u^{0,*} - \sum_{i \geq 1} C(\xi_i, \delta_i) \tag{3.2}$$

in which $W_t^* = e^{-rt} W_t$ and $S_t^{0,*} = e^{-rt} S_t^0$ and $C(x,t) = e^{-rt} C(x,t)$ is the discounted effective transaction cost function. Let $\mathbb{Q}$ be the equivalent measure under which the discounted *unaffected* process $S^{0,*}$ is a martingale. Consider a trading strategy with $W_0 = 0$. Since $E_{\mathbb{Q}} W_T^* \leq -\sum_{i \geq 1} E_{\mathbb{Q}} C(\xi_i, \delta_i) \leq 0$, we can conclude that either $W_T \equiv 0$ or $\mathbb{Q}(W_T^* < 0) > 0$. Since $\mathbb{Q}$ is equivalent to $\mathbb{P}$ we also have that either $W_T \equiv 0$ or $\mathbb{Q}(W_T^* < 0) > 0$. As a result, this strategy is not an arbitrage strategy. ■

Given the assumptions we have made regarding the functions μ and σ, there exists an equivalent measure $\mathbb{Q}$ under which $B_t = W_t + \int_{t_0}^t \frac{\mu(S_u^0, u) - rS_u^0}{\sigma(S_u^0, u)} du$ is a Brownian motion, and $S^{0,*}$ is a martingale with the following dynamic:

$$dS_t^{0,*} = e^{-rt} \sigma(S_t^0, t) dB_t.$$

4 Characterization of the value function

In this section, the optimal strategy X^* of the large trader and the optimal exercise time τ^* of the option holders are characterized in terms of the value function. Furthermore, it is shown that both value functions of the optimization problems coincide and are equal to the risk neutral expectation of the large trader's final wealth $\mathbb{E}J(X^*, \tau^*)$.

We define the following operator:

$$\mathscr{H}[f](s,t) = \max\left(\inf_{x\in A}\left(f(s+\lambda x,t,(t+\varepsilon)\wedge T)+C(x,t)\right),h(s,t)\right)$$

for any $f:\mathbb{R}^+\times[0,T]\times[0,T]\to\mathbb{R}$. $q(s,t_1,t_2)$ is the value function of the optimization problem at time t_1, given that the large trader cannot trade before time t_2. Indeed, in the mean time the option holder is allowed to exercise. If he does not exercise until t_2, the valuation process continues from the value $q(S^0_{t_2},t_2,t_2)$ at time t_2. On the other hand, $\mathscr{H}[q](s,t)$ is the minimal value above the intrinsic value that the large trader can obtain for the value function q at time t by manipulating the stock price, valued at time t at $S^0_t = s$, given that he may not bring the value lower than the intrinsic value $h(s,t)$ of the option at time t because otherwise the option holders would exercise and have the priority. We thus see that at an intervention the large trader is in fact minimizing over the strategies that start ε units of time in the future given the the option holders may exercise in the mean time. These principles are referred to as the dynamic programming principle and will be rigorously demonstrated in the following section.

The following proposition follows from Proposition 3.6 and 2.2.1 of Hamadène and Hassani (2005). It is used to obtain the dynamic programming principle.

Proposition 4.1 *Let $f^0:\mathbb{R}_+\times[0,T]$ be a continuous function. Define*

$$\overline{f}(s,t) = \inf_{\delta\in\mathscr{T}_t}\sup_{\tau\in\mathscr{T}_t}\mathbb{E}_{s,t}\left(h(S^0_\tau,\tau)\mathbf{1}_{\{\tau\le\delta\}}+\mathscr{H}[f^0](S^0_\delta,\delta)\mathbf{1}_{\{\tau>\delta\}}\right)$$

and

$$\underline{f}(s,t) = \sup_{\tau\in\mathscr{T}_t}\inf_{\delta\in\mathscr{T}_t}\mathbb{E}_{s,t}\left(h(S^0_\tau,\tau)\mathbf{1}_{\{\tau\le\delta\}}+\mathscr{H}[f^0](S^0_\delta,\delta)\mathbf{1}_{\{\tau>\delta\}}\right).$$

Then $\overline{f}\equiv\underline{f}=f$ in which

$$f(s,t) = \mathbb{E}_{s,t}\left(h(S^0_{\sigma^*},\sigma^*)\mathbf{1}_{\{\sigma^*\le\delta^*\}}+\mathscr{H}[f^0](S^0_{\delta^*},\delta^*)\mathbf{1}_{\{\sigma^*>\delta^*\}}\right)$$

in which

$$\begin{aligned} \sigma^* &= \inf\{t \le u \le T : f(S_u^0, u) = h(S_u^0, u)\} \\ \delta^* &= \inf\{t \le u \le T : f(S_u^0, u) = \mathscr{H}[f^0](S_u^0, u)\}. \end{aligned}$$

Furthermore, f is continuous.

For $X \in \mathscr{A}_t$ and $t \in [0,T]$, let

$$I_X(t) = \begin{cases} (\delta_i + \varepsilon) \wedge T, & \text{if } t \in [\delta_i, (\delta_i + \varepsilon) \wedge T] \text{ for some } i \ge 1 \text{ ;} \\ t, & \text{otherwise.} \end{cases} \tag{4.1}$$

In words, $I_X(t)$ denotes the next time after t, the large trader can make an intervention (given that an intervention at time T is without consequences).

Our main result is the following dynamic programming principle for the value functions q and u.

Theorem 4.2 *Let $s > 0$ and $0 \le t \le t' \le T$ Then, $u(s,t,t') = q(s,t,t')$. Furthermore,*

$$\begin{aligned} q(s,t) &= \sup_{\tau \in \mathscr{T}_t} \inf_{\delta \in \mathscr{T}_t} \mathbb{E}_{s,t} \left(h(S_\tau^0, \tau) \mathbf{1}_{\{\tau \le \delta\}} + \mathscr{H}[q](S_\delta^0, \delta) \mathbf{1}_{\{\tau > \delta\}} \right) && (4.2) \\ &= \inf_{\delta \in \mathscr{T}_t} \sup_{\tau \in \mathscr{T}_t} \mathbb{E}_{s,t} \left(h(S_\tau^0, \tau) \mathbf{1}_{\{\tau \le \delta\}} + \mathscr{H}[q](S_\delta^0, \delta) \mathbf{1}_{\{\tau > \delta\}} \right) && (4.3) \\ &= \mathbb{E}_{s,t} \left(h(S_{\tau^*}^{X^*}, \tau^*) + \sum_{i \ge 1} C(\xi_i^*, \delta_i^*) \mathbf{1}_{\{\delta_i^* < \tau^*\}} \right) && (4.4) \end{aligned}$$

in which the optimal strategies $X^* \in \mathscr{A}_t$ and $\tau^* \in \mathscr{T}_t$ are given by

$$\begin{aligned} \delta_1^* &= \inf\{t \le u \le T : q(S_u^0, u) = \mathscr{H}[q](S_u^0, u)\} \\ \delta_2^* &= \inf\{\delta_1^* + \varepsilon \le u \le T : q(S_u^0 + \lambda \xi_1^*, u) = \mathscr{H}[q](S_u^0 + \lambda \xi_1^*, u)\} \\ &\quad \dots \\ \delta_i^* &= \inf\{\delta_{i-1}^* + \varepsilon \le u \le T : q(S_u^0 + \lambda \sum_{j=1}^{i-1} \xi_j^*, u) = \mathscr{H}[q](S_u^0 + \lambda \sum_{j=1}^{i-1} \xi_j^*, u)\} && (4.5) \\ &\quad \dots \end{aligned}$$

and $\xi_i^* = argmin_{x\in A} q(S^{X^*}_{\delta_i^*-} + \lambda x, \delta_i^*, (\delta_i^* + \varepsilon)\wedge T) + C(x, \delta_i^*)$, *for* $i \geq 1$, *and*

$$\tau^* = \inf\{t \leq u \leq T : q(S_u^{X^*}, u, I_{X^*}(u)) = h(S_{u-}^{X^*}, u)\}. \tag{4.6}$$

Furthermore, for $t' \geq t$*, we have the following dynamic programming principle:*

$$q(s,t,t') \quad = \quad \sup_{\tau\in\mathscr{T}_{t,t'}} \mathbf{E}_{s,t}\left(h(S_\tau^0,\tau)\mathbf{1}_{\{\tau<t'\}} + q(S_{t'}^0,t')\mathbf{1}_{\{\tau=t'\}}\right) \tag{4.7}$$

in which $\mathscr{T}_{t,t'}$ *is the set of stopping times between* t *and* t'*.*

The proof can be found in the Appendix.

The dynamic programming principle shows that v can be computed from its future values. Furthermore, it gives the optimal time δ_1^* for the large trader's first intervention that minimizes the expected value of v. The size of the first optimal intervention is obtained by minimizing the value $q(S_\delta^0 + \lambda x, \delta_1^*, (\delta_1^* + \varepsilon)\wedge T)$ over x, as done by the operator $\mathscr{H}$.

Since the option price in a frictionless setting is given by

$$q^0(s,t) = \sup_{\tau\in\mathscr{B}_t} \mathbb{E}_{s,t}\left(J(0,\tau)\right),$$

it is greater than the value function

$$q(s,t) = \inf_{X\in\mathscr{A}_t} \sup_{\tau\in\mathscr{T}_t} \mathbb{E}_{s,t}\left(J(X,\tau)\right)$$

for all s and t. The exercise regions correspond to the sets of values of s such that $q(s,t)$ and $q^0(s,t)$ are equal to $h(s,t)$. In particular, the exercise region in the no price impact setting is larger. Since the large trader's strategy is to lower the value of the option, the price of the asset will hit the exercise region earlier in a price impact setting. It is therefore more optimal for the option holders to exercise their option before they would in a market with no price impacts ($\lambda = 0$).

5 Favorable conditions for price manipulations

It will be more difficult for the large trader to manipulate the market when the underlying asset is liquid or when the transaction cost is high. The following proposition gives a conditions under which there will not be any intervention by the large trader in terms of the delta of the American option in a infinitely-liquid, non-manipulated market (a Black-Scholes price, for instance).

Proposition 5.1 *Suppose the function h^0 is convex and the function C has a quadratic form: $C(\xi,t) = c + a|\xi| + \frac{1}{2}b|\xi|^2$, for some non-negative constants $a,b,c \geq 0$. Let Δ_t be the delta of the American option with payoff function h^0 in a infinitely-liquid, non-manipulated market ($\lambda = 0$), and let n_0 be the number of options the large trader has sold. The optimal strategy of the large trader is not to intervene if the following condition is satisfied:*

$$\lambda^2 n_0^2 \Delta_t^2 + a\lambda n_0 \Delta_t - bC(\frac{-\lambda n_0 \Delta - a}{b}) < 0.$$

Proof: Suppose the function C has a quadratic form:

$$C(\xi,t) \quad = \quad c + a|\xi| + \frac{1}{2}b|\xi|^2.$$

Using the convexity of the value function, it is known that

$$q(s+\lambda\xi,t) + C(\xi,t) \geq q(s,t) + \lambda q_s(s,t)\xi + C(\xi,t).$$

The function on the right of the inequality attains its minimum at the point $\xi = \frac{-\lambda q_s(s,t)-a}{b}$. If the condition $\lambda q_s(s,t)\frac{\lambda q_s(s,t)+a}{b} < C(\frac{-\lambda q_s(s,t)-a}{b})$ is satisfied then

$$\begin{aligned} q(s+\lambda\xi,t) + C(\xi,t) &\geq q(s,t) + \lambda q_s(s,t)\xi + C(\xi,t) \\ &\geq q(s,t) - \lambda q_s(s,t)\frac{\lambda q_s(s,t)+a}{b} + C(\frac{\lambda q_s(s,t)+a}{b},t) > q(s,t) \end{aligned}$$

for all $\xi \neq 0$. In this case, the minimum is attained at $\xi = 0$ and it is not optimal to intervene for the large trader. In this case, the value function $q = u$ is the price of the American option in a infinitely-liquid, non-manipulated market, and $n_0\Delta = q_s(s,t)$. ■

When $C(\xi,t) = c + a|\xi| + \frac{1}{2}b|\xi|^2$, the value of c is given by the fixed transaction cost, a represents one half of he bid ask-spread and b is the second-order liquidity cost, i.e. the marginal cost per share for the purchase of x shares of the asset. Recall that the sensitivity of the option value in terms of the asset price is given by Δ_t. For an intervention to be optimal, the large trader must be able to lower the value of the option by more monetary units than the cost of that transaction. This explain the significance of the delta in the conditions for no interventions.

The following corollary gives a simplified condition for no manipulation opportunities.

Corollary 5.2 *Suppose the function C has a quadratic form: $C(\xi,t) = c + a|\xi| + \frac{1}{2}b|\xi|^2$, for some positive constants $a,b,c > 0$. If $n_0 < \frac{\sqrt{a^2+2bc}}{\lambda}$ then there are no manipulation opportunities.*

Proof: If $n_0 < \frac{\sqrt{a^2+2bc}}{\lambda}$ then $\frac{1}{2}\lambda^2 n_0^2 \Delta_t^2 < \frac{1}{2}a^2 + bc$ since $|\Delta_t| < 1$. Hence,

$$\lambda^2 n_0^2 \Delta_t^2 + a\lambda n_0 \Delta_t - bC(\frac{-\lambda n_0 \Delta - a}{b}) = \frac{1}{2}\lambda^2 n_0^2 \Delta_t^2 - \frac{1}{2}a^2 - bc - |\lambda n_0 \Delta_t + a| < 0.$$

■

6 Numerical results

Expression (4.3) allows us to write down a partial differential equation for q. Indeed, by Theorem 6.2 and Proposition 6.3 of Hamadène and Hassani (2005), it is known that v is the unique viscosity solution of the following double obstacle variational inequality:

$$\min\{q(s,t) - h(s,t), \max\{-\frac{\partial q(s,t)}{\partial t} - \frac{1}{2}\sigma(s,t)^2 \frac{\partial^2 q(s,t)}{\partial s^2}, q(s,t) - \mathscr{H}[q](s,t)\}\} = 0 \tag{6.1}$$

for $(s,t) \in \mathbb{R}_+ \times [0,T)$ with $q(s,T) = h(s,T)$ for $s \in \mathbb{R}_+$.

Recall the definition of $\mathscr{H}$:

$$\mathscr{H}[q](s,t) = \max\left(\inf_{x\in A}\left(q(s+\lambda x, t, (t+\varepsilon)\wedge T) + C(x,t)\right), h(s,t)\right).$$

For fixed $t' \in [0,T]$, the function $q(s,t,t')$ is a viscosity solution of

$$\min\{q(s,t,t') - h(s,t), -\frac{\partial q(s,t,t')}{\partial t} - \frac{1}{2}\sigma(s,t)^2 \frac{\partial^2 q(s,t,t')}{\partial s^2}\}\} = 0 \tag{6.2}$$

for $(s,t) \in \mathbb{R}_+ \times [0,t')$ with $q(s,t',t') = q(s,t')$ for $s \in \mathbb{R}_+$.

The value of $q(s,(t+\varepsilon)\wedge T)$ for all $s \in \mathbb{R}_+$ is needed to compute $q(s,t,(t+\varepsilon)\wedge T)$, which is also equal to the value of an American-type claim with payoff h strictly prior to $(t+\varepsilon)\wedge T$, and payoff $q(s,(t+\varepsilon)\wedge T)$ at time $(t+\varepsilon)\wedge T$. In particular, for $t \in [T-\varepsilon,T]$, $q(s,t,(t+\varepsilon)\wedge T)$ is equal to the price of a regular American option since $q(s,(t+\varepsilon)\wedge T) = h(s,T)$ in this case. The two obstacles in (6.1) are then explicitly known for $t \in [T-\varepsilon,T]$, so one can compute the value of $q(s,t)$ on $\mathbb{R}_+ \times [T-\varepsilon,T]$ with (6.1). This observation suggests to proceed backwards in time with the following algorithm to compute numerically the value of the function $q(s,t)$:

1. For $t \in (T-\varepsilon,T]$, compute $q(s,t)$ from Equation (6.1) with the two obstacles h and $\mathscr{H}[q^0]$.

2. Set $n = 2$.

3. For $t \in (T-n\varepsilon, T-(n-1)\varepsilon]$,

   - compute $q(s,t,t+\varepsilon)$ from Equation (6.2) with $t' = t+\varepsilon$,
   - compute $\mathscr{H}[q](s,t) = (\min_{x\in A} q(s+\lambda x,t,t+\varepsilon) + C(x,t)) \vee h(s,t)$,
   - and compute q from Equation (6.1) with the two obstacles h and $\mathscr{H}[v]$.

4. Repeat Step 3 for $n+1$.

In order to increase the accuracy of the numerical results, an implicit finite differences method is used to compute $q(s,t,t+\varepsilon)$. Furthermore, to speed up the calculations, the minimum in the expression for $\mathscr{H}[q](s,t)$ is only computed if the slope of q is sufficiently large at this point. Given the results in the previous section, manipulation is possible only when $q_s(s,t)\frac{q_s(s,t)+a}{b} < C(\frac{-q_s(s,t)-a}{b})$.

6.1 Sensitivity Analysis

Figure 1 and 2 present the size of the price impact in terms of moneyness and time-to-maturity. Here, we consider a put option for which moneyness is defined as strike divided by asset price. We have used the following effective transaction cost function:

$$\tilde{C}(\xi,t) \quad = \quad 10+5|\xi|+0.005|\xi|^2$$

with $r = 1\%$, $\sigma(S,t) = 0.25S$, $S_0 = 100$ and $\lambda = 0.0014$. This leads to an effective cost function given by

$$C(\xi,t) \quad = \quad 10+5|\xi|+0.0043|\xi|^2.$$

In both cases, a position of 5000 put options is considered. Note that in both cases, the condition in Proposition 5.1 is not satisfied, so we may expect to find that market manipulation is optimal for the large trader. Indeed, in the case of put options price manipulations happen close to the exercise region since this is where the slope of the put option is the largest. In the case of a put option, this slope is equal to 1. As a result, letting $\Delta_t = 1$ is Proposition 5.1, we find that the smallest value of λ that leads to manipulation is 0.00109.

In Figure 1, $\varepsilon = 1$ hour is used, whereas in Figure 2 $\varepsilon = 1$ day. The exercise region frontier is also plotted. We observe that the large trader has a major impact on option prices near the frontier. This is to be expected since when the price is close to the exercise region, the large trader has an incentive to bring the price to the exercise region to force the exercise. Indeed, the main strategy for the large trader is to force the option holders to exercise sooner than they would if trades had no impact on the underlying asset prices. We also observe that the large trader's impact quickly diminishes as we move away from the exercise region.

The important conclusion we should draw from this numerical experiment is that the large trader will wait until shortly before the exercise time to start manipulating the price process. Indeed, to simplify the analysis, we can first think of European put options. For this type of options, the payoff is calculated from the terminal asset price S_T^X. As a result, early interventions to lower the price of the underlying will be less effective then later ones if the liquidity cost remains mainly unchanged through time. Both types of intervention will have the same impact on prices, however it is only desirable to manipulate the price when the option terminates in the money, a condition which can only be known at the maturity in the case of European options. We can then expect optimal manipulation to take place close to the maturity for this type of options. The same reasoning can be applied to the case of American options. The only difference is that the payoff is computed at a (random) exercise time τ. However, the closer the price is to the boundary of the exercise region, the more likely it is that the options will be exercised. On the other hand, the large trader has to compare the high liquidity cost of making one large transaction close to the exercise region with the lower liquidity cost of placing smaller orders spaced in time by at least ε units of time farther away from the exercise region. Indeed, given a transaction cost C with a quadratic (or higher) term, it is typically cheaper to divide a large transaction into many smaller ones. However, given that transactions have to spaced in time, a series of smaller transactions has to be started sooner than a large transaction when the price of the underlying reaches the exercise region. Consequently, the large trader has to weigh in the probability that the option will not be exercised (defeating the purpose of intervening) with the added benefit of dividing up a large transaction. Comparing Figure 1 and 2, we see that it is optimal to make some interventions farther away from the exercise region when ε is larger (Figure 2).

Figure 3 and 4 present a cross-section of the price impact for different values of λ, using the same cost function and initial parameters as above and $\varepsilon = 1$ hour for Figure 3, and $\varepsilon = 1$ day in Figure 4. We observe a number of things. The larger the price impact coefficient λ, the larger and the more clustered are the manipulation transactions. Furthermore, for large λ, the price impact is done closer to the exercise region and a moneyness coefficient of 1.

In Figure 5 and 6, we have plotted the exercise region for different values of λ. Larger values of λ lead to exercise regions with higher values of the asset price and less convexity in time. As expected, we notice that for λ less than 0.00109, there are no price impacts and the exercise region is the same as the exercise region of the Black-Scholes model.

7 Conclusion

This study concerns the pricing of American-type options in an illiquid setting in which a large trader has sold a number of options and attempts to manipulate the underlying in his favor to minimize the risk-neutral value of the payoff. The optimal exercise time for the option holders is calculated as the first time the risk neutral value of the option equals its intrinsic value. In order to limit the large trader's impact on prices, a constraint that trades by the large trader must be spaced in time by at least ε units of time and incur liquidity costs have been added. As a result, when an intervention is optimal the large trader is in fact minimizing the risk neutral value of the option ε units of time in the future given that the option holders may exercise in the mean time. By the dynamic programming principle, the large trader will intervene when the value of the option equals this minimized future value.

We have also numerically calculated the optimal exercise region. An important observation is that when ε is small, the large trader waits until the asset price is closer to the exercise region than he would for larger values of ε, and makes smaller interventions. The large trader manipulates prices to force the option holders to exercise sooner than under frictionless conditions. The large trader will only attempt to manipulate prices when he has a high certainty that it will have a significant impact of the value of the option. Overall, the main finding is that the manipulation has the largest impact close to the exercise region and the optimal exercise time of the option holders is significantly sooner than in a non-manipulated market.

A Appendix

This appendix contains the proof of Theorem 4.2.

For $X \in \mathscr{A}_t$, define $U(s,t,X) = \sup_{\tau \in \mathscr{T}_t} \mathbf{E}_{s,t} J(X,\tau)$. Recall that J is defined as

$$J(X,\tau) = h(S^X_{\tau-},\tau) + \sum_{i \geq 1} C(\xi_i,\delta_i)\mathbf{1}_{\{t \leq \delta_i < \tau\}}.$$

Since U is the Snell envelope of J for fixed X, the dynamic programming principle applies and for all $\theta \in \mathscr{T}_t$

$$U(s,t,X) = \sup_{\tau} \mathbf{E}_{s,t}\left(J(X,\tau)\mathbf{1}_{\{\tau \leq \theta\}} + U(S^X_\theta,\theta,X)\mathbf{1}_{\{\tau > \theta\}}\right).$$

We can then write $q(s,t,t') = \inf_{X \in \mathscr{A}_{t'}} U(s,t,X)$.

We start by proving the following two inequalities:

For all $\eta > 0$, there exists $X^\eta \in \mathscr{A}_{t'}$ such that for all $\theta \in \mathscr{T}_t$

$$q(s,t,t') \geq \sup_{\tau \in \mathscr{T}_t} \mathbf{E}_{s,t} \Big(J(X^\eta, \tau) \mathbf{1}_{\{\tau \leq \theta\}} + \Big(q(S_\theta^{X^\eta}, \theta, I_{X^\eta}(\theta)) \mathbf{1}_{\{\tau > \theta\}} + C(\xi_{i_\theta}, \theta) \Big) \Big) - \eta, \tag{A.1}$$

and for all $X \in \mathscr{A}_{t'}$ and $\theta \in \mathscr{T}_t$

$$q(s,t,t') \leq \sup_{\tau \in \mathscr{T}_t} \mathbf{E}_{s,t} \left(J(X,\tau) \mathbf{1}_{\{\tau \leq \theta\}} + \left(q(S_\theta^X, \theta, I_X(\theta)) + C(\xi_{i_\theta}, \theta) \right) \mathbf{1}_{\{\tau > \theta\}} \right), \tag{A.2}$$

in which $i_\theta = i$ if $\delta_i = \theta$ (we define $\xi_{i_\theta} = 0$ if $\theta \neq \delta_i$ for any i).

Proof of (A.1): Let $X \in \mathscr{A}_{t'}$ and $\theta \in \mathscr{T}_t$. Since $U(S_\theta^X, \theta, X) \geq q(S_\theta^X, \theta, I_X(\theta)) + C(\xi_{i_\theta}, \theta)$ on $\{\tau > \delta\}$, we find

$$U(s,t,X) \geq \sup_{\tau \in \mathscr{T}_t} \mathbf{E}_{s,t} \left(J(X,\tau) \mathbf{1}_{\{\tau \leq \theta\}} + \left(q(S_\theta^X, \theta, I_X(\theta)) + C(\xi_{i_\theta}, \theta) \right) \mathbf{1}_{\{\tau > \theta\}} \right).$$

In particular, for all $\eta > 0$, there exists $X^\eta \in \mathscr{A}_{t'}$ such that for all θ

$$q(s,t,t') \geq \sup_{\tau \in \mathscr{T}_t} \mathbf{E}_{s,t} \Big(J(X^\eta, \tau) \mathbf{1}_{\{\tau \leq \theta\}} + \Big(q(S_\theta^{X^\eta}, \theta, I_{X^\eta}(\theta)) \mathbf{1}_{\{\tau > \theta\}} + C(\xi_{i_\theta}, \theta) \Big) \Big) - \eta.$$

Proof of (A.2): Let $X \in \mathscr{A}_{t'}$, $\eta > 0$ and $\theta \in \mathscr{T}_t$. There exists $X^\eta \in \mathscr{A}_\theta$ such that $U(S_\theta^X, \theta, X^\eta) \leq q(S_\theta^X, \theta, I_X(\theta)) + C(\xi_{i_\theta}, \theta) + \eta$. Define $\tilde{X}^\eta = X$ on $[0,\theta)$ and $\tilde{X}^\eta = X_\theta + X^\eta$ on $[\theta, T]$. Then,

$$\begin{aligned} q(s,t,t') &\leq U(s,t,\tilde{X}^\eta) = \sup_{\tau \in \mathscr{T}_t} \mathbf{E}_{s,t} \left(J(\tilde{X}^\eta, \tau) \mathbf{1}_{\{\tau \leq \theta\}} + U(S_\theta^X, \theta, \tilde{X}^\eta) \mathbf{1}_{\{\tau > \theta\}} \right) \\ &\leq \sup_{\tau \in \mathscr{T}_t} \mathbf{E}_{s,t} \left(J(X,\tau) \mathbf{1}_{\{\tau \leq \theta\}} + (q(S_\theta^X, \theta, I_X(\theta)) + C(\xi_{i_\theta}, \theta)) \mathbf{1}_{\{\tau > \theta\}} \right) + \eta. \end{aligned}$$

We then obtain (A.2) by sending η to zero.

For $X \in \mathscr{A}_{t'}$, define $\theta = \delta_1$ in (A.2) Then, by taking the infimum over $X \in \mathscr{A}_{t'}$, we find that

$$q(s,t,t') \leq \inf_{X \in \mathscr{A}_{t'}} \sup_{\tau \in \mathscr{T}_t} \mathbf{E}_{s,t} \Big(J(X, \delta_1) \mathbf{1}_{\{\tau \leq \delta_1\}} + (q(S_{\delta_1}^X, \delta_1, (\delta_1 + \varepsilon) \vee T) + C(\xi_1, \delta_1)) \mathbf{1}_{\{\tau > \delta_1\}} \Big).$$

Furthermore, by taking the infimum in (A.1) and sending η to zero, we obtain the following inequality:

$$q(s,t,t') \quad \geq \quad \inf_{X\in\mathscr{A}_{t'}} \sup_{\tau\in\mathscr{T}_t} \mathbf{E}_{s,t}\left(J(X,\delta_1)\mathbf{1}_{\{\tau\leq\delta_1\}} + (q(S^X_{\delta_1},\delta_1,(\delta_1+\varepsilon)\vee T) + C(\xi_1,\delta_1))\mathbf{1}_{\{\tau>\delta_1\}}\right).$$

Combining these two inequalities, we obtain

$$\begin{aligned} q(s,t,t') &= \inf_{\delta_1\in\mathscr{T}_{t'},\xi_1\in\mathscr{F}_{\delta_1}} \sup_{\tau\in\mathscr{T}_t} \mathbf{E}_{s,t}\left(h(S^0_\tau,\tau)\mathbf{1}_{\{\tau\leq\delta_1\}} + (q(S^0_{\delta_1}+\lambda\xi_1,\delta_1,(\delta_1+\varepsilon)\vee T) + C(\xi_1,\delta_1))\mathbf{1}_{\{\tau>\delta_1\}}\right) \\ &= \inf_{\delta_1\in\mathscr{T}_{t'}} \sup_{\tau\in\mathscr{T}_t} \mathbf{E}_{s,t}\left(h(S^0_\tau,\tau)\mathbf{1}_{\{\tau\leq\delta_1\}} + \min_{\xi_1\in\mathscr{F}_{\delta_1}}(q(S^0_{\delta_1}+\lambda\xi_1,\delta_1,(\delta_1+\varepsilon)\vee T) + C(\xi_1,\delta_1))\mathbf{1}_{\{\tau>\delta_1\}}\right). \end{aligned}$$

We now show that

$$q(s,t,t') = \inf_{\delta\in\mathscr{T}_{t'}} \sup_{\tau\in\mathscr{T}_t} \mathbf{E}_{s,t}\left(h(S^0_\tau,\tau)\mathbf{1}_{\{\tau\leq\delta\}} + h(S^0_\delta,\delta)\vee\min_{\xi\in A}(q(S^0_\delta+\lambda\xi,\delta,(\delta+\varepsilon)\vee T) + C(\xi,\delta))\mathbf{1}_{\{\tau>\delta_1\}}\right).$$

Define

$$U_1(s,t,\delta) = \sup_{\tau\in\mathscr{T}_t} \mathbf{E}_{s,t}\left(h(S^0_\tau,\tau)\mathbf{1}_{\{\tau\leq\delta\}} + \min_{\xi\in A}(q(S^0_\delta+\lambda\xi,\delta,(\delta+\varepsilon)\vee T) + C(\xi,\delta))\mathbf{1}_{\{\tau>\delta\}}\right).$$

Then

$$U_1(s,t,\delta) \leq \sup_{\tau\in\mathscr{T}_t} \mathbf{E}_{s,t}\left(h(S^0_\tau,\tau)\mathbf{1}_{\{\tau\leq\delta\}} + h(S^0_\delta,\delta)\vee\min_{\xi\in A}(q(S^0_\delta+\lambda\xi,\delta,(\delta+\varepsilon)\vee T) + C(\xi,\delta))\mathbf{1}_{\{\tau>\delta\}}\right).$$

For $\tau\in\mathscr{T}_t$, define $\hat{\tau}$ as follows:

$$\hat{\tau} = \begin{cases} \tau, & \text{if } \tau\leq\delta; \\ \delta, & \text{if } \tau>\delta \text{ and } \min_{\xi\in A}(q(S^0_\delta+\lambda\xi,\delta,(\delta+\varepsilon)\vee T) + C(\xi,\delta)) < h(S^0_\delta,\delta); \\ \tau, & \text{otherwise.} \end{cases}$$

From the definition of $\hat{\tau}$, it follows that

$$\begin{aligned}
&\mathbf{E}_{s,t}\left(h(S^0_{\hat{\tau}},\hat{\tau})\mathbf{1}_{\{\hat{\tau}\le\delta\}}+\min_{\xi\in A}(q(S^0_\delta+\lambda\xi,\delta,(\delta+\varepsilon)\vee T)+C(\xi,\delta))\mathbf{1}_{\{\hat{\tau}>\delta\}}\right)\\
&=\mathbf{E}_{s,t}\left(h(S^0_\tau,\tau)\mathbf{1}_{\{\tau\le\delta\}}+h(S^0_\delta,\delta)\vee\min_{\xi\in A}(q(S^0_\delta+\lambda\xi,\delta,(\delta+\varepsilon)\vee T)+C(\xi,\delta))\mathbf{1}_{\{\tau>\delta\}}\right)
\end{aligned}$$

As a result, we can state that

$$q(s,t,t') \quad = \quad \inf_{\delta\in\mathscr{T}_{t'}}\sup_{\tau\in\mathscr{T}_t}\mathbf{E}_{s,t}\left(h(S^0_\tau,\tau)\mathbf{1}_{\{\tau\le\delta\}}+h(S^0_\delta,\delta)\vee\min_{\xi\in A}(q(S^0_\delta+\lambda\xi,\delta,(\delta+\varepsilon)\vee T)+C(\xi,\delta))\mathbf{1}_{\{\tau>\delta\}}\right).$$

This proves (4.2) since

$$\mathscr{H}[q](S^0_\delta,\delta)=\left(\min_{x\in A}\left(q(S^0_\delta+\lambda x,\delta,(\delta+\varepsilon)\wedge T)+C(x,\delta)\right)\right)\vee h(S^0_\delta,\delta).$$

By Proposition 4.1, we can also swap the infimum and the supremum, so that Equation (4.3) is also verified. Furthermore, the optimal stopping times for $q(s,t)$ are given in Proposition 4.1 by

$$\begin{aligned}
\sigma^* &= \inf\{t\le u\le T: q(S^0_u,u)=h(S^0_u,u)\}\\
\delta^* &= \inf\{t\le u\le T: f(S^0_u,u)=\mathscr{H}[f^0](S^0_u,u)\},
\end{aligned}$$

so that the following representation of $q(s,t)$ is satisfied:

$$q(s,t) \quad = \quad \mathbf{E}_{s,t}\left(h(S^0_{\sigma^*},\sigma^*)\mathbf{1}_{\{\sigma^*\le\delta^*\}}+\mathscr{H}[q](S^0_{\delta^*},\delta^*)\mathbf{1}_{\{\sigma^*>\delta^*\}}\right). \tag{A.3}$$

We can now prove Equation (4.7). In this regard, define $V(s,t,t',\tau)=\inf_{\delta\in\mathscr{T}_{t'}}\mathbf{E}_{s,t}\left(h(S^0_\tau,\tau)\mathbf{1}_{\{\tau\le\delta\}}+\mathscr{H}[q](S^0_\delta,\delta)\mathbf{1}_{\{\tau>\delta\}}\right)$. Then,

$$\begin{aligned}
V(s,t,t',\tau) &= \mathbf{E}_{s,t}\left(h(S^0_\tau,\tau)\mathbf{1}_{\{\tau<t'\}}+\mathbf{1}_{\{\tau\ge t'\}}\inf_{\delta\in\mathscr{T}_{t'}}\mathbf{E}_{S^0_{t'},t'}\left(h(S^0_\tau,\tau)\mathbf{1}_{\{\tau\le\delta\}}+\mathscr{H}[q](S^0_\delta,\delta)\mathbf{1}_{\{\tau>\delta\}}\right)\right)\\
&= \mathbf{E}_{s,t}\left(h(S^0_\tau,\tau)\mathbf{1}_{\{\tau<t'\}}+V(S^0_{t'},t',t',\tau)\mathbf{1}_{\{\tau\ge t'\}}\right).
\end{aligned}$$

Then,

$$
\begin{aligned}
q(s,t,t') &= \sup_{\tau\in\mathscr{T}_t} \mathbf{E}_{s,t}\left(h(S^0_\tau,\tau)\mathbf{1}_{\{\tau<t'\}} + V(S^0_{t'},t',t',\tau)\mathbf{1}_{\{\tau\geq t'\}}\right)\\
&= \sup_{\tau\in\mathscr{T}_t} \mathbf{E}_{s,t}\left(h(S^0_\tau,\tau)\mathbf{1}_{\{\tau<t'\}} + \sup_{\tau\in\mathscr{T}_{t'}} V(S^0_{t'},t',t',\tau)\mathbf{1}_{\{\tau\geq t'\}}\right)\\
&= \sup_{\tau\in\mathscr{T}_{t,t'}} \mathbf{E}_{s,t}\left(h(S^0_\tau,\tau)\mathbf{1}_{\{\tau<t'\}} + q(S^0_{t'},t')\mathbf{1}_{\{\tau=t'\}}\right),
\end{aligned}
$$

which proves (4.7). Furthermore, by the Snell envelope principle the value function $q(s,t,t')$ can be written as follows:

$$
q(s,t,t') = \mathbf{E}_{s,t}\left(h(S^0_{\check{\tau}^*},\check{\tau}^*)\mathbf{1}_{\{\check{\tau}^*<t'\}} + q(S^0_{t'},t')\mathbf{1}_{\{\check{\tau}^*=t'\}}\right) \tag{A.4}
$$

in which $\check{\tau}^* = \inf\{t\leq u\leq t' : q(S^0_u,u,t') = h(S^0_u,u)\}$ is the optimal time associated to this optimal stopping problem.

To obtain the optimal strategy we proceed inductively on the intervals $(T-n\varepsilon, T-(n-1)\varepsilon]$, starting with $n=1$.

Let $t\in(T-\varepsilon,T]$. Define $X^*\in\mathscr{A}_t$ and $\tau^*\in\mathscr{T}_t$ by

$$
\delta^*_1 = \inf\{t\leq u\leq T : q(S^0_u,u) = \mathscr{H}[q](S^0_u,u)\}
$$

and $\xi^*_1 = argmin_{x\in A} q(S^0_{\delta^*_1}+\lambda x, \delta^*_1, (\delta^*_1+\varepsilon)\wedge T) + C(x,\delta^*_1)$, and

$$
\tau^* = \begin{cases} \inf\{t\leq u\leq T : q(S^0_u,u) = h(S^0_u,u)\}, & \text{on } \{\inf\{t\leq u\leq T : q(S^0_u,u) = h(S^0_u,u)\}\leq\delta^*_1\};\\ \inf\{\delta^*_1\leq u\leq T : q(S^0_u+\lambda\delta^*_1,u,I_{X^*}(u)) = h(S^0_u+\lambda\delta^*_1,u)\}, & \text{otherwise.}\end{cases}
$$

In this case,

$$
q(s,t) = \inf_{X\in\mathscr{A}_t}\sup_{\tau\in\mathscr{T}_t}\mathbb{E}_{s,t}\left(h(S^X_\tau,\tau) + C(\xi_1,\delta_1)\mathbf{1}_{\{\delta_1<\tau\}}\right).
$$

Furthermore, for all $u \in [t,T]$, $q(s,u,(u+\varepsilon)\wedge T) = q^0(s,u)$ since $(u+\varepsilon)\wedge T = T$ and $q(s,T) = h(s,T)$.

From (4.2) and (4.7), we obtain

$$\begin{aligned}
q(s,t) &= \mathbb{E}_{s,t}\left(h(S^0_{\tau^*},\tau^*)\mathbf{1}_{\{\tau^*\le\delta_1^*\}} + \left(q(S^{X^*}_{\delta_1^*},\delta_1^*,(\delta_1^*+\varepsilon)\wedge T) + C(\xi_1^*,\delta_1^*)\right)\mathbf{1}_{\{\tau^*>\delta_1^*\}}\right) \\
&= \mathbb{E}_{s,t}\left(h(S^{X^*}_{\tau^*-},\tau^*)\mathbf{1}_{\{\tau^*\le\delta_1^*\}} + C(\xi_1^*,\delta_1^*)\mathbf{1}_{\{\tau^*>\delta_1^*\}}\right) \\
&\quad + \mathbb{E}_{s,t}\left(\left(\mathbb{E}_{S^{X^*}_{\delta_1^*},\delta_1^*}\left(h(S^{X^*}_{\tau^*-},\tau^*)\mathbf{1}_{\{\tau^*<(\delta_1^*+\varepsilon)\wedge T\}} + q(S^{X^*}_{(\delta_1^*+\varepsilon)\wedge T},(\delta_1^*+\varepsilon)\wedge T)\mathbf{1}_{\{\tau^*\ge(\delta_1^*+\varepsilon)\wedge T\}}\right)\right)\mathbf{1}_{\{\tau^*>\delta_1^*\}}\right),
\end{aligned}$$

from (A.4). Above, the facts that $S^{X^*}_{\tau^*-} = S^0_{\tau^*}$ on $\{\tau^* \le \delta_1^*\}$ and $q(S^{X^*}_{\delta_1^*},\delta_1^*,(\delta_1^*+\varepsilon)\wedge T) + C(\xi_1^*,\delta_1^*) > h(S^{X^*}_{\delta_1^*},\delta_1^*)$ on $\{\tau^* > \delta_1^*\}$ were used. Using the fact that $q(S^{X^*}_{(\delta_1^*+\varepsilon)\wedge T},(\delta_1^*+\varepsilon)\wedge T) = h(S^{X^*}_T,T)$, we can simplify the above expression to

$$q(s,t) \quad = \quad \mathbb{E}_{s,t}\left(h(S^{X^*}_{\tau^*-},\tau^*) + C(\xi_1^*,\delta_1^*)\mathbf{1}_{\{\tau^*>\delta_1^*\}}\right).$$

Suppose now that for some $n \ge 1$,

$$q(s,t) \quad = \quad \mathbb{E}_{s,t}\left(h(S^{X^*}_{\tau^*-},\tau^*) + \sum_{i=1}^{n} C(\xi_i^*,\delta_i^*)\mathbf{1}_{\{\tau^*>\delta_i^*\}}\right),$$

for all $(s,t) \in \mathbb{R}\times(T-n\varepsilon,T]$.

Let $t \in (T-(n+1)\varepsilon, T-n\varepsilon]$. Define

$$\delta_1^* \quad = \quad \inf\{t \le u \le T : q(S^0_u,u) = \mathscr{H}[q](S^0_u,u)\}$$

and $\xi_1^* = argmin_{x\in A} q(S^0_{\delta_1^*}+\lambda x,\delta_1^*,(\delta_1^*+\varepsilon)\wedge T) + C(x,\delta_1^*)$.

Let $\tilde{X}^* \in \mathscr{A}_{(\delta_1^*+\varepsilon)\wedge T}$ be defined by $\tilde{X}_t^* = \xi_1^* + \sum_{i=1}^n \tilde{\xi}_i^* \mathbf{1}_{\{t \leq \tilde{\delta}_i^*\}}$ with

$$
\begin{aligned}
\tilde{\delta}_1^* &= \inf\{(\delta_1^*+\varepsilon)\wedge T \leq u \leq T : q(S_u^0+\lambda\xi_1^*, u) = \mathscr{H}[q](S_u^0, u)\} \\
\tilde{\delta}_2^* &= \inf\{(\tilde{\delta}_1^*+\varepsilon)\wedge T \leq u \leq T : q(S_u^0+\lambda\xi_1^*+\lambda\tilde{\xi}_1^*, u) = \mathscr{H}[q](S_u^0+\lambda\tilde{\xi}_1^*, u)\} \\
&\dots \\
\tilde{\delta}_i^* &= \inf\{(\tilde{\delta}_{i-1}^*+\varepsilon)\wedge T \leq u \leq T : q(S_u^0+\lambda\xi_1^*+\lambda\sum_{j=1}^{i-1}\tilde{\xi}_j^*, u) = \mathscr{H}[q](S_u^0+\lambda\sum_{j=1}^{i-1}\tilde{\xi}_j^*, u)\} \\
&\dots
\end{aligned}
$$

and $\tilde{\xi}_i^* = argmin_{x\in A} q(S_{\tilde{\delta}_i^*-}^{\tilde{X}^*} + \lambda\xi_1^* + \lambda x, \tilde{\delta}_i^*, (\tilde{\delta}_i^*+\varepsilon)\wedge T) + C(x, \tilde{\delta}_i^*)$, for $i \geq 1$.

Define $X^* \in \mathscr{A}_t$ by concatenation as follows:

$$
\begin{aligned}
\delta_{i+1}^* &= \tilde{\delta}_i^*, \text{ and} \\
\xi_{i+1}^* &= \tilde{\xi}_i^*, \text{ for } i \geq 1.
\end{aligned}
$$

Finally, define τ^* as in the statement of the theorem:

$$
\tau^* = \inf\{t \leq u \leq T : q(S_u^{X^*}, u, I_{X^*}(u)) = h(S_{u-}^{X^*}, u)\}.
$$

Then,

$$
\begin{aligned}
q(s,t) &= \mathbb{E}_{s,t}\left(h(S^0_{\tau^*},\tau^*)\mathbf{1}_{\{\tau^*\le\delta_1^*\}} + \left(q(S^{X^*}_{\delta_1^*},\delta_1^*,(\delta_1^*+\varepsilon)\wedge T) + C(\xi_1^*,\delta_1^*)\right)\mathbf{1}_{\{\tau^*>\delta_1^*\}}\right) \\
&= \mathbb{E}_{s,t}\left(h(S^0_{\tau^*},\tau^*)\mathbf{1}_{\{\tau^*\le\delta_1^*\}} + C(\xi_1^*,\delta_1^*)\mathbf{1}_{\{\tau^*>\delta_1^*\}}\right) \\
&\quad + \mathbb{E}_{s,t}\left(\mathbb{E}_{S^{X^*}_{\delta_1^*},\delta_1^*}\left(h(S^{\widetilde{X}^*}_{\tilde{\tau}^*-},\tilde{\tau}^*)\mathbf{1}_{\{\delta_1^*<\tau^*<(\delta_1^*+\varepsilon)\wedge T\}} + q(S^{\widetilde{X}^*}_{(\delta_1^*+\varepsilon)\wedge T},(\delta_1^*+\varepsilon)\wedge T)\mathbf{1}_{\{\tau^*\ge(\delta_1^*+\varepsilon)\wedge T\}}\right)\right) \\
&= \mathbb{E}_{s,t}\left(h(S^0_{\tau^*},\tau^*)\mathbf{1}_{\{\tau^*\le\delta_1^*\}} + \mathbb{E}_{S^{X^*}_{\delta_1^*},\delta_1^*}\left(h(S^{\widetilde{X}^*}_{\tilde{\tau}^*-},\tilde{\tau}^*)\mathbf{1}_{\{\delta_1^*<\tau^*<(\delta_1^*+\varepsilon)\wedge T\}}\right)\right) \\
&\quad + \mathbb{E}_{s,t}\left(q(S^{\widetilde{X}^*}_{(\delta_1^*+\varepsilon)\wedge T},(\delta_1^*+\varepsilon)\wedge T)\mathbf{1}_{\{\tau^*\ge(\delta_1^*+\varepsilon)\wedge T\}}\right) + \mathbb{E}_{s,t}\left(C(\xi_1^*,\delta_1^*)\mathbf{1}_{\{\tau^*>\delta_1^*\}}\right) \\
&= \mathbb{E}_{s,t}\left(h(S^{X^*}_{\tau^*-},\tau^*)\mathbf{1}_{\{\tau^*<(\delta_1^*+\varepsilon)\wedge T\}} + C(\xi_1^*,\delta_1^*)\mathbf{1}_{\{\tau^*>\delta_1^*\}}\right) \\
&\quad + \mathbb{E}_{s,t}\left(\mathbb{E}_{S^{\widetilde{X}^*}_{(\delta_1^*+\varepsilon)\wedge T},(\delta_1^*+\varepsilon)\wedge T}\left(h(S^{\widetilde{X}^*}_{\tau^*-},\tau^*) + \sum_{i=1}^n C(\tilde{\xi}_i^*,\tilde{\delta}_i^*)\mathbf{1}_{\{\tau^*>\tilde{\delta}_i^*\}}\right)\mathbf{1}_{\{\tau^*\ge(\delta_1^*+\varepsilon)\wedge T\}}\right) \\
&= \mathbb{E}_{s,t}\left(h(S^{X^*}_{\tau^*-},\tau^*) + \sum_{i=1}^n C(\xi_i^*,\delta_i^*)\mathbf{1}_{\{\tau^*>\delta_i^*\}}\right),
\end{aligned}
$$

since by definition $\widetilde{X}_t^* = X_t^*$ for $t > (\delta_1^*+\varepsilon)\wedge T$.

■

Bibliography

Bank, P. and Baum, D. (2004). Hedging and portfolio optimization in financial markets with a large trader. *Mathematical Finance*, 14:1–18.

Bensoussan, A. (1984). On the theory of option pricing. *Acta Applicandae Mathematicae*, 2:139–158.

Bensoussan, A. and Lions, J. L. (1978). *Applications des inéquations variationelles en contrôle stochastique*. Dunod, Paris.

Çetin, U. and Rogers, L. (2007). Modelling liquidity effects in discrete time. *Mathematical Finance*, 17:15–29.

Cetin, U., Jarrow, R., and Protter, P. (2004). Liquidity risk and arbitrage pricing theory. *Finance and Stochastics*, 8:311–341.

Cvitanic, J. and Karatzas, I. (1996). Backward stochastic differential equations with reflection and Dynkin games. *The Annals of Probability*, 24(4):2024–2056.

Dutt, H. R. and Harris, L. E. (2005). Position limits for cash-settled derivative contracts. *Journal of Futures Markets*, 25:945–965.

Friedman, A. (1972). Stochastic differential games. *Journal of Differential Equations*, 11:79–108.

Hamadène, S. (2006). Mixed zero-sum stochastic differential game and American game options. *SIAM J. Control Optim.*, 45(2):496–518.

Hamadène, S. and Hassani, M. (2005). BSDEs with two reflecting barriers : the general result. *Probability Theory and Related Fields*, 132:237–264.

Horst, U. and Naujokat, F. (2011). On derivatives with illiquid underlying and market manipulation. *Quantitative Finance*, 11:1051–1066.

Jarrow, R. (1994). Derivative securities markets, market manipulation and option pricing theory. *The Journal of Financial and Quantitative Analysis*, 29(2):241–261.

Kallsen, J. and Kühn, C. (2004). Pricing derivatives of American and game type in incomplete markets. *Finance and Stochastics*, 8:3261–284.

Karatzas, I. (1988). On the pricing of American options. *Applied Mathematics and Optimization*, 17:37–60.

Karatzas, I. and Sudderth, W. D. (2006). Stochastic games of control and stopping for a linear diffusion. In Hsiung, A., Ying, Z., and Zhang, C., editors, *Random Walk, Sequential Analysis and Related Topics: A Festschrift in Honor of Y.S. Chow*, pages 100–117. World Scientific, Singapore.

Kumar, P. and Seppi, D. J. (1992). Futures manipulation with "cash settlement". *Journal of Finance*, 47:1485–1502.

Kyle, A. (1985). Continuous auctions and insider trading. *Econometrica*, 53:1315–1335.

Kyle, A. and Viswanathan, S. (2008). How to define illegal price manipulation. *Amer. Econ. Rev.*, 98:274–279.

Liu, H. and Yong, J. (2005). Option pricing with an illiquid underlying asset market. *Journal of Economic Dynamics and Control*, 29:2125–2156.

Pirrong, S. C. (1993). Manipulation of the commodity futures market delivery process. *Journal of Business*, 66:335–369.

Pirrong, S. C. (2001). Manipulation of cash-settled futures contracts. *Journal of Business*, 74:221–244.

Roch, A. (2011). Liquidity risk, price impacts and the replication problem. *Finance and Stochastics*, 15:399–419.

Touzi, N. and Vieille, N. (2002). Continuous-time Dynkin games with mixed strategies. *SIAM Journal on Control and Optimization*, 41:1037–1088.

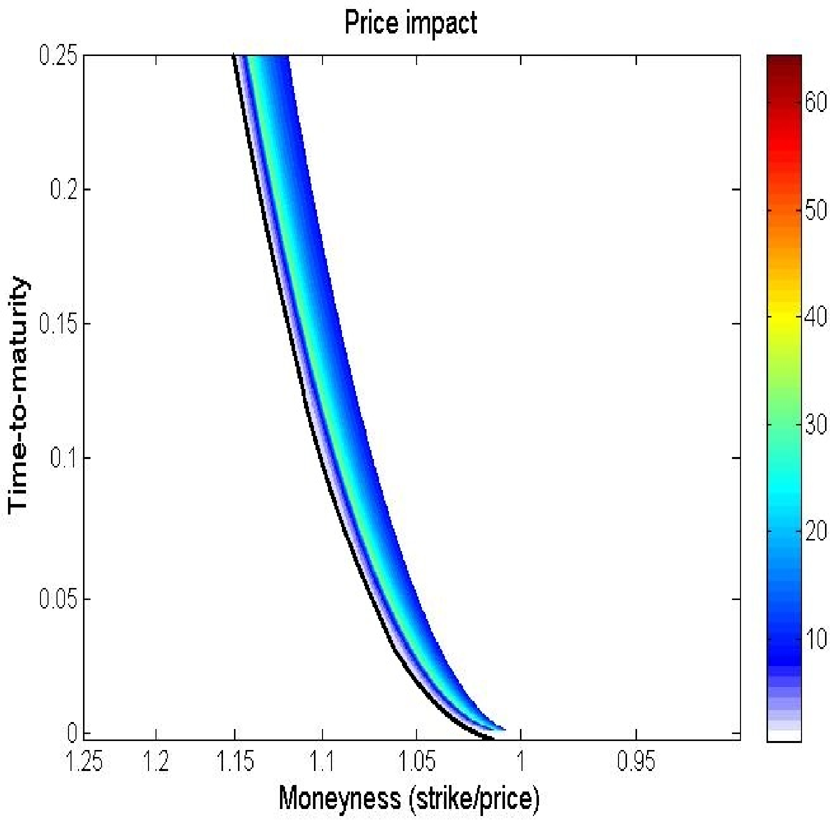

Figure 1: Price impact with ε = 1 hour and $\lambda = 0.0014$. The black line is the exercise region.

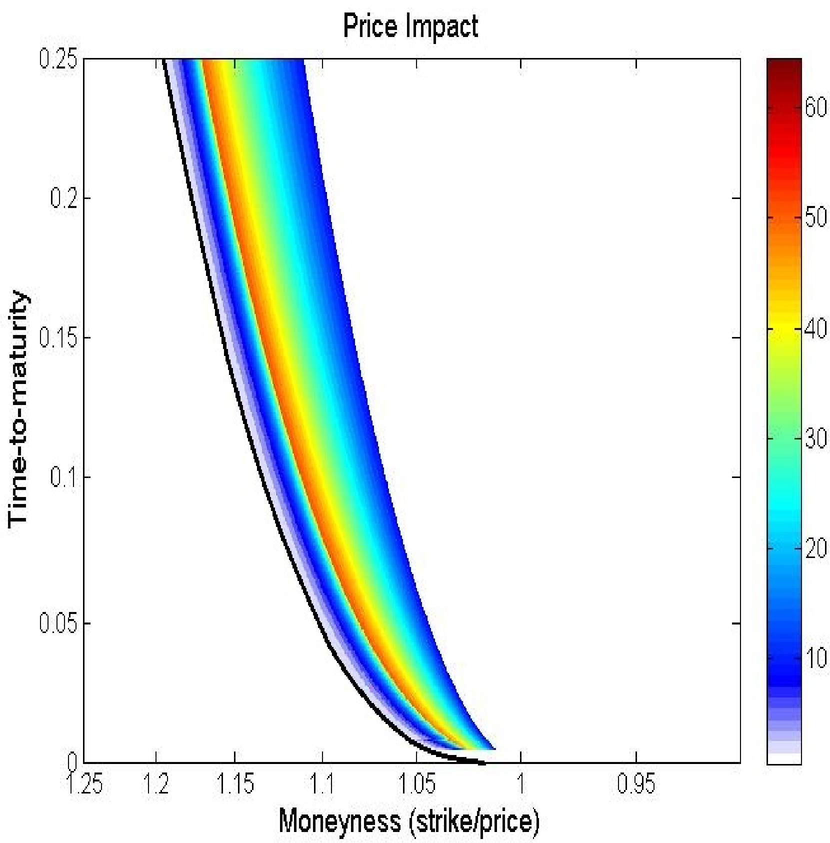

Figure 2: Price impact with ε = 1 day and $\lambda = 0.0014$. The black line is the exercise region.

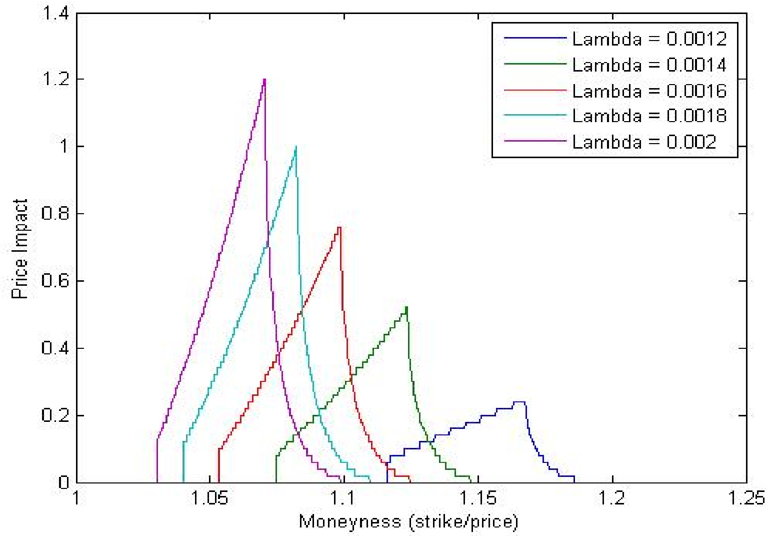

Figure 3: Price impact, with 30 days left before maturity, and ε = 1 hour.

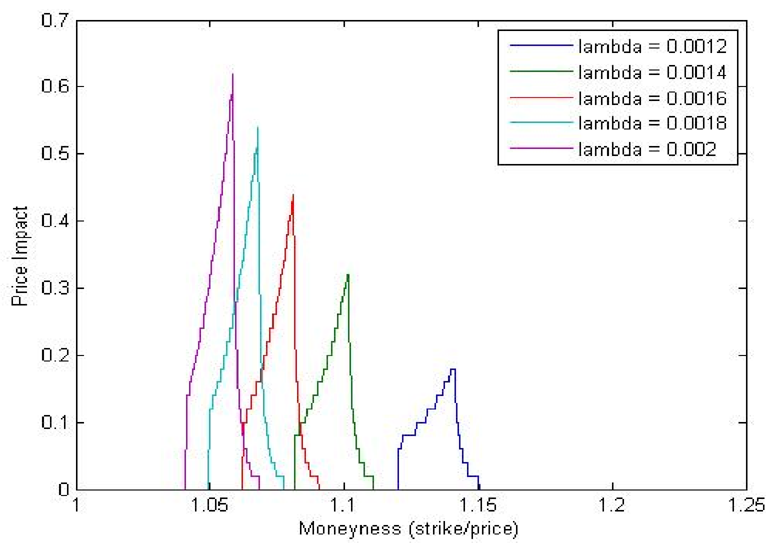

Figure 4: Price impact, with 30 days left before maturity, and ε = 1 day.

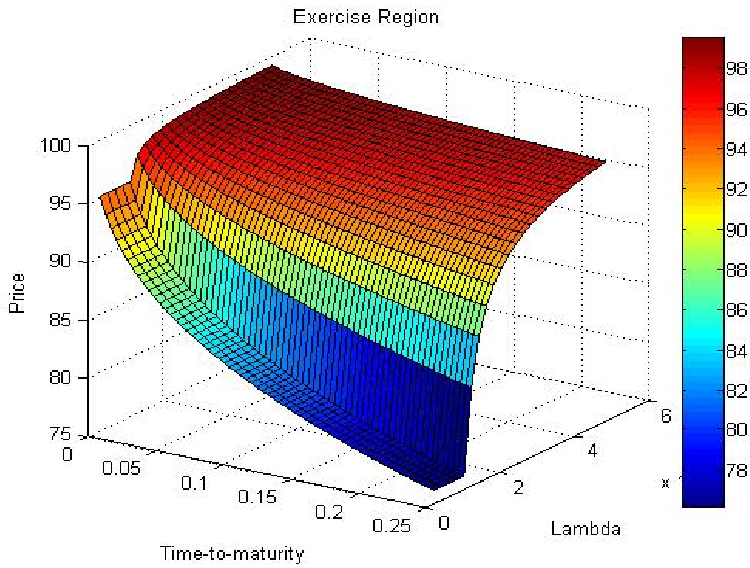

Figure 5: Exercise region with ε = 1 hour.

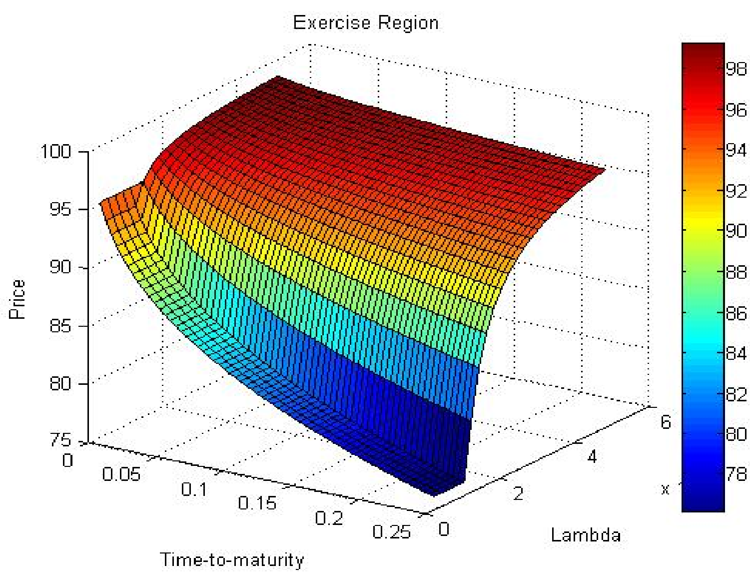

Figure 6: Exercise region with ε = 1 day.