Option pricing for a large trader with price impact and liquidity costs

Hyejin Ku*, Hai Zhang

Department of Mathematics and Statistics, York University, Toronto, Canada

1. Introduction

Liquidity risk is considered as the most important risk in finance industry these days. Although there seems no unanimous agreement on the definition of liquidity risk, there has been a growing literature on “illiquidity” which is to model the effect of illiquidity and to solve important problems within the model. Apparently, modeling of liquidity risk is challenging but very important.

In the literature on market liquidity for the underlying asset, there are two approaches for modeling; one is temporary price impact and the other is permanent price impact. The first one is the effect of liquidity cost incurred while changing position as a price-taking trader. Roughly speaking, this arises on short time scales as the result of trading, and can be thought of as a trader having to work through the limit order book to acquire his desired change of holding. The second one is the effect of a large trader on the underlying asset (called feedback effect), caused by the fact that a large trader has just traded some quantity of the asset. By the large trader we mean there is some lasting impact caused by its trading action.

* Corresponding author. 4700 Keele St., Toronto, ON, Canada.
E-mail address: hku@mathstat.yorku.ca (H. Ku).

https://doi.org/10.1016/j.jmaa.2017.10.072
0022-247X/© 2017 Elsevier Inc. All rights reserved.
Models for the effect of liquidity costs without considering permanent price impact have some similarities to those for transaction costs, in the sense that illiquidity adds some extra costs to trade. The effect of illiquidity makes trading more difficult or costly. Many researchers built on the model for market liquidity risk outlined in Cetin et al. [5]. Essentially the spot price of the underlying in the model depends on the size of the block being traded through the stochastic supply curve. For example if one wants to purchase a huge amount of shares of stock, there may not be enough supply at the market price, so one will end up paying above the market price for such a big block of shares. In Cetin et al. [6], authors use strategies with minimal super-replication cost inclusive of liquidity premium to price contingent claims in continuous time setting. Ku et al. [17] derived a partial differential equation which provides discrete time delta hedging strategies whose expected hedging errors approach zero almost surely as the length of the revision interval goes to zero. The equation gives the value of the call from the seller’s point of view. Sorokin and Ku [22] further investigated to minimize the replication errors with inhomogeneous rebalancing times. Zhang and Ku [23] provided a simple model for pricing and hedging options under liquidity risk and jump. In all the literature above, the approach does not take into account the impact on the evolution of underlying asset from the actions of a trader and is suitable for models accounted for small traders.

There also has been a growing literature for large traders. Frey and Stremme [12] modeled the effect of the dynamic hedging strategies on the equilibrium price of the underlying asset and used general aggregate demand reaction function that depends on the trader’s exogenous stochastic income. Also see for example Jarrow [16], Frey [11], Platen and Schweizer [20], Bank and Baum [4], Ku and Zhang [18] and Eksi and Ku [9]. These papers assume trading actions have a lasting effect on the stock price evolution. For the literature on optimal liquidation in which the aim is to unwind an initial position by some fixed time horizon, we refer to Almgren and Chriss [2], Almgren [1], He and Mamaysky [13], and Forsyth [10]. These papers try to liquidate a given initial position optimally by some fixed time. Longstaff [19] considered the optimal portfolio choices in an illiquid market where the trading strategies were assumed to be of bounded variation. The paper of Avellaneda and Lipkin [3] discussed stock pinning on option expiration date and the price impact of delta-hedging.

In this paper, we investigate the option pricing and hedging problem for a large trader considering both temporary price impact and permanent price impact. Specifically, we assume illiquidity will pose some kind of nonlinear transaction cost on trading and a trading action will have a lasting impact on the stock price evolution. A trader will face costs in trying to trade very rapidly. Thus the effect of illiquidity costs depends on the rate of change of holding, rather than the size of change of holding. Rogers and Singh [21], and Forsyth [10] also assume that the effect of illiquidity costs is dependent on the speed of trades. We assume, moreover, there may be a lasting impact on the underlying asset due to this rapid trading action. We use the utility based approach to price European options in a market with liquidity risk. Utility indifference pricing is proven to be a powerful method to price options in the market with friction, such as a market with transaction cost in Hodges and Neuberger [15] and Davis et al. [8] or a market with non-traded assets in Henderson [14]. We apply the utility indifference approach to price European options in illiquid markets.

We study the utility maximization problems, and derive two Hamilton–Jacobi–Bellman equations to characterize the value functions for this optimal control problem. We define the option price to be the difference between the initial wealth of the two utility maximization problems achieving the same expected utility. We use viscosity solutions to characterize HJB equations and prove the existence and uniqueness of solutions of the HJB equations. We provide an example incorporating liquidity risk and permanent price impact, which gives an explicit solution for optimal strategy. We give a detailed discussion of this example and numerical results.

The paper is organized as follows. Section 2 introduces our model and explains the utility maximization problem. Section 3 is devoted to results on existence and uniqueness of solutions for the HJB equations that arise from the optimal control problem. Section 4 provides an example and discusses the numerical results in detail. Section 5 presents some conclusions.
2. The model

We consider a financial market which consists of a risk-free asset and one risky asset \( S \) on a given probability space \( (\Omega, \mathcal{F}, P) \) with a filtration \( \{\mathcal{F}_t : t \geq 0\} \). The price of the risk-free asset (the amount of cash in a bank account) grows at interest rate \( r \).

We define the set of trading strategies to be the set of all \( \{\mathcal{F}_t\} \)-adapted processes with left continuous paths that have right limits. We let \( N_t \) be the number of shares of asset \( S \) held at time \( t \). We shall assume that \( N_t \) to be a finite-variation process, where \( N_t = N_0 + \int_0^t v_s \, d\xi \) and the trading speed can be expressed as

\[
v_t = \frac{dN_t}{dt}.
\]

We restrict the set of trading strategies available to the trader by the condition that a trader cannot change his position extremely fast, i.e., the changes in the number of shares of asset \( S \) held over any time interval never exceed \( M \)-multiple of the length of the time interval. We assume \( v_t \in [-M, M] \), \( M \in \mathbb{R}^+ \) large, for all \( t \in [0, T] \). We note that \( M \) might be determined by market conditions such as the daily trading volume of the asset. We also assume that a trading strategy is allowed if it keeps the wealth (mark-to-market value) bounded below. A trading strategy satisfying this condition is called admissible; the class of admissible trading strategies will be denoted by \( \Gamma = \{v_t : 0 \leq t \leq T\} \).

2.1. The permanent price impact

In this paper, we consider the pricing problem in illiquid markets for a large trader whose trading action will have a lasting impact on the underlying price evolution. The price of the risky asset \( S_t \) follows an \( \{\mathcal{F}_t\} \)-adapted geometric Brownian motion, except the effect of the price impact. The lasting price impact is modeled by imposing a function of the trading speed into the drift term of the risky asset price as follows:

\[
dS_t = (\mu + g(v_t))S_t \, dt + \sigma S_t \, dW_t, \quad t \in [0, T]
\]
(1)

where \( g(\cdot) \) represents the effect of the price impact (possibly identifying as zero, if one wishes). This function is assumed to be continuous and \( g(0) = 0 \). A similar price process has been used in Almgren and Chriss [2].

2.2. The effect of liquidity costs

In illiquid markets, the market provides different prices for buying and selling stock, depending on how many shares she wants to trade, or how rapidly she wants to change the position. Let \( S(t, v_t, \omega)^2 \) be the stock price per share at time \( t \in [0, T] \) that a trader pays/receives for a trading speed \( v_t \in \mathbb{R} \). The actual execution price of the stock to be paid/received is different from the price initially quoted. In practice, if a trader wants to change in her holding with speed \( v_t \) the actual traded price \( S(t, v_t, \omega) \) will not be equal to the market price \( S_t \) due to the effect of illiquidity. More specifically, when \( v_t > 0 \), the stock is purchased and the buying price will be greater than \( S_t \). When \( v_t < 0 \), the stock is sold and the selling price will be less than \( S_t \). We assume the (stochastic) traded price of stock is given by

\[
S(t, v_t, \omega) = f(v_t)S_t,
\]
(2)

\[\text{The authors considered arithmetic rather than geometric Brownian motion.}\]
\[\text{This actual traded price function is similar to the stochastic supply curve \( S(t, x, \omega) \) introduced in Cetin et al. [5] where the price function depends on the order size \( x \) rather than the trading speed.}\]
where \( f(\cdot) \) is a positive, continuous and nondecreasing function with \( f(0) = 1 \). This implies \( S(t, v_t, \omega) \) increases as \( v_t \) increases, which is consistent with the intuition. The faster the buying speed, the higher the average paid price per share. The quicker the selling speed, the lower the average received price per share of stock.

2.3. No arbitrage in the model

As explained in Introduction, a large trader will incur liquidity costs during the trading, and at the same time, her trading action will place a feedback effect on the stock price evolution. We now consider the model including both the effect of liquidity costs and the permanent price impact. To apply the utility indifference pricing method, we must verify that this large trader model is arbitrage-free.

Recall \( N_t \) is the number of shares of asset \( S \) held at time \( t \). If the trading speed for the time interval \([t, t+h]\) is \( v_t \), then we have \( N_{t+h} = N_t + v_t h \). We also let \( B_t \) be the amount of money in the risk-free asset at time \( t \) and denote the portfolio value (the mark-to-market value) by \( Y_t \) i.e.,

\[
Y_t = N_t S_t + B_t.
\]

**Definition 2.1.** Under this large trader model, the portfolio \( Y_t \) is self-financing if it satisfies

\[
dY_t = N_t dS_t + rB_t dt + (1 - f(v_t)) S_t v_t dt.
\]

Observe that if \( f(v) = 1 + \alpha v \) (\( \alpha \geq 0 \)), then the self-financing condition gives

\[
dY_t = N_t dS_t + rB_t dt - \alpha v^2 S_t dt.
\]

Also if \( f(v) = 1 \), we have the usual self-financing condition without liquidity costs.

**Theorem 2.1. (No arbitrage in the large trader model)** Let \( S_t \) be given by the SDE:

\[
dS_t = (\mu + g(v_t)) S_t dt + \sigma S_t dW_t
\]

and the self-financing condition is assumed. Then the model does not admit arbitrage.

**Proof.** The evolutions of the discounted stock price and portfolio value are

\[
d(e^{-rt} S_t) = (\mu + g(v_t) - r) e^{-rt} S_t dt + \sigma e^{-rt} S_t dW_t,
\]

and

\[
d(e^{-rt} Y_t) = N_t d(e^{-rt} S_t) + (1 - f(v_t)) e^{-rt} S_t v_t dt.
\]

Let \( v_t \) be the trading speed corresponding to an admissible trading strategy \((N_t, B_t) \). Since \( g(\cdot) \) is continuous, it is obvious that

\[
E \left[ \exp \left( \frac{1}{2} \int_0^T \left( \frac{\mu + g(v_t) - r}{\sigma} \right)^2 dt \right) \right] < \infty,
\]

and the Novikov condition is satisfied. By Girsanov’s theorem, there exists an equivalent probability measure \( Q \) under which the process
$W_t = W_t + \int_0^t \frac{\mu + g(v_s) - r}{\sigma} ds, \quad t \in [0, T]$

is a Brownian motion. The Radon–Nikodym derivative reads

$$\frac{dQ}{dP} = \exp \left( - \int_0^T \frac{\mu + g(v_t) - r}{\sigma} dW_t - \int_0^T \left( \frac{\mu + g(v_t) - r}{\sigma} \right)^2 dt \right).$$

Under the probability measure $Q$, we have

$$d(e^{-rt}S_t) = \sigma e^{-rt}S_t d\hat{W}_t,$$
$$d(e^{-rt}Y_t) = N_t \sigma e^{-rt}S_t d\hat{W}_t + (1 - f(v_t))e^{-rt}S_tv_t dt,$$

and

$$e^{-rT}Y_T = Y_0 + \int_0^T N_t \sigma e^{-rt}S_t d\hat{W}_t + \int_0^T (1 - f(v_t))e^{-rt}S_tv_t dt.$$

Since $f(\cdot)$ is a nondecreasing function with $f(0) = 1$, we have

$$(1 - f(v_t))e^{-rt}S_tv_t \leq 0.$$

Suppose there exists an arbitrage strategy such that $Y_0 \leq 0$ and

$$P(Y_T \geq 0) = 1 \quad \text{and} \quad P(Y_T > 0) > 0.$$  \hfill (4)

Then $P(Y_T < 0) = 0$, and by the equivalence of $P$ and $Q$,

$$Q(Y_T < 0) = 0.$$ \hfill (5)

We note

$$E^Q \left[ \int_0^T (N_t \sigma e^{-rt}S_t)^2 dt \right] \leq E^Q \left[ \int_0^T \sigma^2 \sup_t \{N_t^2\} (e^{-rt}S_t)^2 dt \right] \leq \sigma^2 \bar{N} \int_0^T E^Q \left[ (e^{-rt}S_t)^2 \right] dt \leq \bar{N} S_0^2 (e^{\sigma^2 T} - 1) < \infty$$

where $\bar{N} = \max\{ (N_0 - MT)^2, (N_0 + MT)^2 \}$, and the integral $\int_0^T N_t \sigma e^{-rt}S_t d\hat{W}_t$ is a martingale under $Q$. Then taking the expectation of $e^{-rT}Y_T$,

$$E^Q[e^{-rT}Y_T] = Y_0 + E^Q \left[ \int_0^T (1 - f(v_t))e^{-rt}S_tv_t dt \right] \leq Y_0 \leq 0 \quad \text{because} \quad (1 - f(v_t))e^{-rt}S_tv_t \leq 0 \text{ for every } t.$$  \hfill (6)

because $1 - f(v_t) e^{-rt} S_t v_t \leq 0$ for every $t$. From (5) and (6), one has
\[ Q(Y_T > 0) = 0. \]

By the equivalence of \( P \) and \( Q \), it implies
\[ P(Y_T > 0) = 0, \]
which contradicts (4). Therefore, we conclude that the model does not admit an arbitrage strategy. \( \square \)

2.4. Utility indifference pricing for European options

The idea of utility maximization approach to pricing a European option is as follows. The utility indifference price for a contingent claim \( C \) is the price at which a trader is indifferent (in the sense that her expected utility under optimal trading is unchanged) between receiving \( p \) now to pay a claim \( C_T \) at time \( T \) and receiving nothing and having no obligation.

Assume the trader has initially \( B_0 \) units of risk-free asset (amount in a bank account). If the trader writes \( n \) units of a European option for price \( p \), she will receives \( n \) multiple of \( p \) at time 0 for writing the option, and she needs to buy or sell stock to maximize expected utility of wealth she will obtain after fulfilling the obligation for the option \( C_T \) at maturity \( T \). The trader will try to maximize her expected utility of the wealth at maturity \( T \) even if she does not take the short position for the option.

Let \( U : \mathbb{R} \to \mathbb{R} \) be a utility function. The utility function is assumed to be a concave, strictly increasing and twice continuously-differentiable function. First, we consider the expected utility maximization for final wealth with the option obligation. Assume that at time \( t \), the stock price is \( S_t \) and the trader holds \( B_t \) units in the bank account and \( N_t \) shares of stock. The large trader controls the trading speed \( v_t \) to adjust her stock position, so \( v_t \) is the control variable.

We define
\[ J^w(t, N, S, B, v) = \mathbb{E}\left\{ U\left( N_T S_T + B_T - nC_T \right) \right\} \bigg| N_t = N, S_t = S, B_t = B \]
where \( C_T \) is the payoff of the European option at maturity \( T \) and \( n \) is the number of the options sold. The value function of the trader with the option obligation is given by
\[ V^w(t, N, S, B) = \max_{v \in \Gamma} J^w(t, N, S, B, v) \]

Next we consider the expected utility maximization of final wealth without option obligation. Define
\[ J(t, N, S, B, v) = \mathbb{E}\left\{ U\left( N_T S_T + B_T \right) \right\} \bigg| N_t = N, S_t = S, B_t = B \]
The value function of the trader without having option obligation is given by
\[ V(t, N, S, B) = \max_{v \in \Gamma} J(t, N, S, B, v) \]
Let the initial stock price at time 0 is \( S_0 \). Then the utility indifference price for the option at time \( t = 0 \) is the real number \( p \) satisfying the following equation
\[ V^w(0, N_0, S_0, B_0 + np) = V(0, N_0, S_0, B_0) \]
Clearly, the utility indifference price \( p \) depends on her prior exposure \((N_0, S_0, B_0)\) and utility function \( U(\cdot) \).
From (1) and (2), we have the following dynamics of state variables
\[
\begin{align*}
\frac{dN_t}{dt} &= v_t dt \\
\frac{dS_t}{dt} &= (\mu + g(v_t))S_t dt + \sigma S_t dW_t \\
\frac{dB_t}{dt} &= rB_t dt - f(v_t)S_t v_t dt
\end{align*}
\]

From the definitions of value functions and the dynamics of state variables, it is evident that the dynamic programming principle yields the same HJB equation for these two value functions. The only difference between the two value functions is that they have different terminal conditions. By a familiar argument, the value function \( V^w(t, N, S, B) \) and \( V(t, N, S, B) \) are expected to satisfy the following HJB equation:
\[
0 = \max_{v \in [-M, M]} \left\{ \frac{\partial W}{\partial t} + v \frac{\partial W}{\partial N} + (\mu + g(v))\frac{\partial W}{\partial S} - f(v)S \frac{\partial W}{\partial B} + rB \frac{\partial W}{\partial B} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 W}{\partial S^2} \right\}
\]
which can be rewritten as
\[
0 = \max_{v \in [-M, M]} \left\{ v \frac{\partial W}{\partial N} - f(v)S \frac{\partial W}{\partial B} + (\mu + g(v))\frac{\partial W}{\partial S} \right\} + \frac{\partial W}{\partial t} + \frac{\partial W}{\partial B} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 W}{\partial S^2}
\]
with the terminal condition \( W(T, N, S, B) = U(NS + B - nC_T) \) for \( V^w(t, N, S, B) \), and \( W(T, N, S, B) = U(NS + B) \) for \( V(t, N, S, B) \).

3. Value functions of the optimal control problem

In this section, we show the value function of our stochastic control problem is a unique viscosity solution of HJB equation that arises from the optimal control problem in Section 2. This result is implied by obtaining the existence and uniqueness of the solutions for a nonlinear second-order PDE. The notion of viscosity solutions was introduced by Crandall and Lions. For a general view of the theory, we refer to the user’s guide by Crandall et al. [7]. We consider a nonlinear second-order PDE of the form
\[
-\frac{\partial W(t, x)}{\partial t} + H(x, D_x W(t, x), D_x^2 W(t, x)) = 0
\]
where \((t, x) \in (0, T] \times \mathcal{D}\) and \(H(x, p, Z)\) is continuous mapping from \(\mathcal{D} \times \mathbb{R}^N \times S_N \rightarrow \mathbb{R}\), where \(S_N\) denotes the set of symmetric \(N \times N\) matrices.

For the PDE in this paper, \(H(x, p, Z)\) has the following specific form:
\[
H(x, p, Z) = -\max_{v \in \mathbb{R}} \left\{ [v, (\mu + g(v))x_2, -vf(v)x_2 - rx_3] \cdot [p_1, p_2, p_3]^T \right\} - \frac{\sigma^2}{2} (0, x_2, 0)^T Z (0, x_2, 0)^T
\]
where \(x = (x_1, x_2, x_3) \in \mathcal{D}\), \(p = (p_1, p_2, p_3) \in \mathbb{R}^3\), and \(Z \in S_3\). \(\mathcal{D}\) denotes a locally compact subset of \(\mathbb{R}^3\).

**Theorem 3.1.** The value function \( V^w(t, N, S, B) \) is a viscosity solution of
\[
-\frac{\partial W}{\partial t} - \max_v \left\{ v \frac{\partial W}{\partial N} - f(v)S \frac{\partial W}{\partial B} - rB \frac{\partial W}{\partial B} + (\mu + g(v))S \frac{\partial W}{\partial S} \right\} - \frac{\sigma^2 S^2}{2} \frac{\partial^2 W}{\partial S^2} = 0 \quad (7)
\]
on \([0, T] \times \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}\).
Proof. (i) Let $x = (N, S, B)$ and $\mathcal{D} = \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}$. We first prove that $V^w(t, x)$ is a viscosity subsolution of (7) on $[0, T) \times \mathcal{D}$. For this, we need to show that for all smooth function $\phi(t, x)$, such that $V^w(t, x) - \phi(t, x)$ has a local maximum at $(t_0, x_0) \in (0, T) \times \mathcal{D}$, the following inequality holds:

\[
- \frac{\partial \phi(t_0, x_0)}{\partial t} - \max_v \left\{ v \frac{\partial \phi(t_0, x_0)}{\partial N} - (f(v)S_t v + rB_t)\frac{\partial \phi(t_0, x_0)}{\partial B} + (\mu + g(v))S_t \frac{\partial \phi(t_0, x_0)}{\partial S} \right\} - \frac{\sigma^2 S_t^2}{2} \frac{\partial^2 \phi(t_0, x_0)}{\partial S^2} \leq 0
\]

Without loss of generality, we assume that $V^w(t_0, x_0) = \phi(t_0, x_0)$, and $V^w \leq \phi$ on $(0, T) \times \mathcal{D}$. Suppose that, on the contrary, there exist function $\phi$ and control variable $v_0 \in \Gamma$ where $\Gamma$ is the set of all admissible controls, satisfying the property that there exists an open set $\mathcal{O}(t_0, x_0)$ containing $(t_0, x_0)$ such that $\phi(t_0, x_0) = V^w(t_0, x_0)$ and $\phi(t, x) \geq V^w(t, x)$ for all $(t, x) \in \mathcal{O}(t_0, x_0)$. Then there exists $\theta > 0$ such that

\[
- \frac{\partial \phi(t, x)}{\partial t} - \max_v \left\{ v \frac{\partial \phi(t, x)}{\partial N} - (f(v)S_t v + rB_t)\frac{\partial \phi(t, x)}{\partial B} + (\mu + g(v))S_t \frac{\partial \phi(t, x)}{\partial S} \right\} - \frac{\sigma^2 S_t^2}{2} \frac{\partial^2 \phi(t, x)}{\partial S^2} > \theta
\]

for all $(t, x) \in \mathcal{O}(t_0, x_0)$. Let $\tau$ be the stopping time

\[
\tau = \inf \{ t \in [t_0, T], (t, x) \notin \mathcal{O}(t_0, x_0) \}.
\]

Then, for $t_0 \leq t \leq \tau$ and fixed $v \in \Gamma$, we have

\[
J(t_0, N_{t_0}, S_{t_0}, B_{t_0}, v) \leq E_{t_0} [V^w(\tau, S_\tau, N_\tau, B_\tau)] \leq E_{t_0} [\phi(\tau, S_\tau, N_\tau, B_\tau)]
\]

By Dynkin’s formula, we have

\[
E[\phi(t_\tau, S_\tau, N_\tau, B_\tau)] = \phi(t_0, N_{t_0}, S_{t_0}, B_{t_0}) + E_{t_0} \left[ \int_{t_0}^{\tau} \frac{\partial \phi(t, x)}{\partial t} + v \frac{\partial \phi(t, x)}{\partial N} - (f(v)S_t v + rB_t)\frac{\partial \phi(t, x)}{\partial B} + (\mu + g(v))S_t \frac{\partial \phi(t, x)}{\partial S} - \frac{\sigma^2 S_t^2}{2} \frac{\partial^2 \phi(t, x)}{\partial S^2} dt \right] 
\]

\[
\leq \phi(t_0, N_{t_0}, S_{t_0}, B_{t_0}) - E_{t_0} \left[ \int_{t_0}^{\tau} \theta dt \right]
\]

Taking the supremum over all admissible control $v \in \Gamma$, we have

\[
\phi(t_0, N_{t_0}, S_{t_0}, B_{t_0}) = V^w(t_0, N_{t_0}, S_{t_0}, B_{t_0}) = \max_{v \in \Gamma} J(t_0, N_{t_0}, S_{t_0}, B_{t_0}, v)
\]

\[
\leq \phi(t_0, N_{t_0}, S_{t_0}, B_{t_0}) - E_{t_0} \left[ \int_{t_0}^{\tau} \theta dt \right]
\]

This contradicts the fact that $\theta > 0$. Therefore $V^w(t, N, S, B)$ is a viscosity subsolution.

(ii) Next, we prove $V^w(t, x)$ is a viscosity supersolution of (7) on $[0, T) \times \mathcal{D}$. Given $(t_0, N_{t_0}, S_{t_0}, B_{t_0}) \in (0, T) \times \mathcal{D}$, let $\phi(t, x) \in C^{1,2}$ such that $V^w(t, x) - \phi(t, x)$ has a local minimum in $\mathcal{O}(t_0, x_0)$. Without loss of
generality, we assume that $V^w(t_0, x_0) = \phi(t_0, x_0)$ and $V^w(t, x) \geq \phi(t, x)$ on $O(t_0, x_0)$. Let $\tau$ be the stopping time

$$\tau = \inf \left\{ t \in [t_0, T], (t, x) \notin O(t_0, x_0) \right\}$$

Given $t_0 < t_1 < \tau$, consider the control variable $v_t = v \in \Gamma$ where $v$ is a constant for $t \in [t_0, t_1]$. From the dynamic programming principle, we have

$$V^w(t_0, N_{t_0}, S_{t_0}, B_{t_0}) \geq E_{t_0}[V^w(t_1, N_{t_1}, S_{t_1}, B_{t_1})]$$

Also, we know

$$V^w(t_1, N_{t_1}, S_{t_1}, B_{t_1}) \geq \phi(t_1, N_{t_1}, S_{t_1}, B_{t_1})$$

By Dynkin’s formula

$$E[\phi(t_1, N_{t_1}, S_{t_1}, B_{t_1})] = \phi(t_0, N_{t_0}, S_{t_0}, B_{t_0}) + E_{t_0}\left[ \int_{t_0}^{t_1} \frac{\partial \phi(t, x)}{\partial t} + v \frac{\partial \phi(t, x)}{\partial N} - (f(v)Sv + rB) \frac{\partial \phi(t, x)}{\partial B} + (\mu + g(v))S \frac{\partial \phi(t, x)}{\partial S} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 \phi(t, x)}{\partial S^2} dt \right].$$

From the fact that

$$V^w(t_0, N_{t_0}, S_{t_0}, B_{t_0}) \geq E_{t_0}[V^w(t_1, N_{t_1}, S_{t_1}, B_{t_1})] \geq E_{t_0}[\phi(t_1, N_{t_1}, S_{t_1}, B_{t_1})]$$

and $V^w(t_0, N_{t_0}, S_{t_0}, B_{t_0}) = \phi(t_0, N_{t_0}, S_{t_0}, B_{t_0})$,

$$E_{t_0}\left[ \int_{t_0}^{t_1} \frac{\partial \phi(t, x)}{\partial t} + v \frac{\partial \phi(t, x)}{\partial N} - (f(v)Sv + rB) \frac{\partial \phi(t, x)}{\partial B} + (\mu + g(v))S \frac{\partial \phi(t, x)}{\partial S} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 \phi(t, x)}{\partial S^2} dt \right] \leq 0$$

Divide by $t_1 - t_0$. Using the mean value theorem for integrals and letting $t_1 \to t_0$, we have

$$\frac{\partial \phi(t_0, x_0)}{\partial t} + v \frac{\partial \phi(t_0, x_0)}{\partial N} - (f(v)Sv + rB) \frac{\partial \phi(t_0, x_0)}{\partial B} + (\mu + g(v))S \frac{\partial \phi(t_0, x_0)}{\partial S} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 \phi(t_0, x_0)}{\partial S^2} \leq 0$$

Taking the supremum over $v \in [-M, M]$, we conclude

$$-\frac{\partial \phi(t_0, x_0)}{\partial t} - \max_v \left\{ v \frac{\partial \phi(t_0, x_0)}{\partial N} - (f(v)Sv + rB) \frac{\partial \phi(t_0, x_0)}{\partial B} + (\mu + g(v))S \frac{\partial \phi(t_0, x_0)}{\partial S} \right\} - \frac{\sigma^2 S^2}{2} \frac{\partial^2 \phi(t_0, x_0)}{\partial S^2} \geq 0$$

Thus, $V^w(t, N, S, B)$ is a viscosity supersolution of (7).

From (i) and (ii), $V^w(t, N, S, B)$ is both a viscosity supersolution and subsolution of (7), and hence the proof is completed. \qed

We showed that $V^w(t, N, S, B)$ is a viscosity solution of

$$0 = \max_v \left\{ v \frac{\partial W}{\partial N} - f(v)Sv \frac{\partial W}{\partial B} + (\mu + g(v))S \frac{\partial W}{\partial S} \right\} + \frac{\partial W}{\partial t} + rB \frac{\partial W}{\partial B} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 W}{\partial S^2}.$$
on $[0,T) \times \mathcal{D}$ with the terminal condition
\[ W(T,N,S,B) = U (NS + B - nC_T), \]
and also $V(t, N, S, B)$ is a viscosity solution of
\[ 0 = \max_v \left\{ v \frac{\partial W}{\partial N} - f(v)S v \frac{\partial W}{\partial B} + (\mu + g(v)) S \frac{\partial W}{\partial S} \right\} + \frac{\partial W}{\partial t} + rB \frac{\partial W}{\partial B} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 W}{\partial S^2} \]
on $[0,T) \times \mathcal{D}$ with the terminal condition
\[ W(T,N,S,B) = U (NS + B). \]

In the following we show the value function is the unique viscosity solution of (7). For this, we first present the following comparison principle:

**Theorem 3.2.** Let $V_i(t,x)$ be a upper semicontinuous viscosity subsolution of (7), and let $V_2(t,x)$ be a lower semicontinuous viscosity supersolution of (7). Assume $V_1$ and $V_2$ satisfy the linear growth condition: $|V_i(t,x)| \leq K(1 + |x|)$, $i = 1, 2$, for positive constant $K$. If $V_1(T,x) \leq V_2(T,x)$ for $x \in \mathcal{D}$ and $V_1(t,x) \leq V_2(t,x)$ for $0 < t \leq T$ and $x \in \partial \mathcal{D}$, then $V_1(t,x) \leq V_2(t,x)$ for all $(t,x) \in (0,T) \times \mathcal{D}$.

**Proof.** Step 1: We can rewrite the equation in the following form:
\[ -\frac{\partial W(t,x)}{\partial t} + H(x,D_xW(t,x),D_x^2W(t,x)) = 0 \] (8)
where
\[ H(x,p,Z) = -\max_v \left\{ v \frac{\partial W}{\partial N} - f(v)S v \frac{\partial W}{\partial B} + (\mu + g(v)) S \frac{\partial W}{\partial S} \right\} - \frac{\sigma^2}{2} (0, x_2, 0)Z(0, x_2, 0)^T \]
Denote
\[ \hat{H}(x,p) = -\max_v \left\{ v \frac{\partial W}{\partial N} - f(v)S v \frac{\partial W}{\partial B} + (\mu + g(v)) S \frac{\partial W}{\partial S} \right\} \]
Let $V_1(t,x)$ be a viscosity subsolution of (8). For $\rho > 0$, define
\[ V_1^\rho(t,x) = V_1(t,x) - \frac{\rho}{t+T}, (t,x) \in (0,T) \times \mathcal{D} \]
Then we have
\[ \frac{d}{dt}(-\frac{\rho}{t+T}) = \frac{\rho}{(t+T)^2} > 0 \]
So we can claim $V_1^\rho(t,x)$ is a viscosity subsolution of (8). In fact,
\[ -\frac{\partial V_1^\rho(t,x)}{\partial t} + H(x,D_xV_1^\rho(t,x),D_x^2V_1^\rho(t,x)) \leq -\frac{\rho}{(t+T)^2} \leq -\frac{\rho}{4T^2} \] (9)
Step 2: For any $0 < \delta < 1$ and $0 < \gamma < 1$, define
\[ \Phi(t,x,y) = V_1^\rho(t,x) - V_2(t,y) - \frac{1}{\delta} |x - y|^2 - \gamma e^{T-t}(x^2 + y^2) \]
and

$$\phi(t, x, y) = \frac{1}{\delta} |x - y|^2 + \gamma e^{t-s}(x^2 + y^2)$$

Since $V_1(t, x)$ and $V_2(t, x)$ satisfy the linear growth condition, we have

$$\lim_{|t| \to \infty} \Phi(t, x, y) = -\infty$$

and $\Phi(t, x, y)$ is continuous in $(t, x, y)$. Therefore, $\Phi(t, x, y)$ has a global maximum at a point $(t_\delta, x_\delta, y_\delta)$. Note that

$$\Phi(t_\delta, x_\delta, y_\delta) = V_1^\rho(t_\delta, x_\delta) - V_2(t_\delta, y_\delta) - \frac{1}{\delta} |x_\delta - y_\delta|^2 - \gamma e^{t-s}(x_\delta^2 + y_\delta^2).$$

In particular,

$$\Phi(t_\delta, x_\delta, x_\delta) + \Phi(t_\delta, y_\delta, y_\delta) \leq 2\Phi(t_\delta, x_\delta, y_\delta)$$

which means

$$V_1^\rho(t_\delta, x_\delta) - V_2(t_\delta, x_\delta) - \gamma e^{t-s}(x_\delta^2 + x_\delta^2)$$

$$+ V_1^\rho(t_\delta, y_\delta) - V_2(t_\delta, y_\delta) - \gamma e^{t-s}(y_\delta^2 + y_\delta^2)$$

$$\leq 2V_1^\rho(t_\delta, x_\delta) - 2V_2(t_\delta, y_\delta) - \frac{2}{\delta} |x_\delta - y_\delta|^2 - 2\gamma e^{t-s}(x_\delta^2 + y_\delta^2).$$

Thus we have

$$\frac{2}{\delta} |x_\delta - y_\delta|^2 \leq [V_1^\rho(t_\delta, x_\delta) - V_1^\rho(t_\delta, y_\delta)] + [V_2(t_\delta, y_\delta) - V_2(t_\delta, y_\delta)]$$

(10)

By the linear growth condition, there exist $K_1, K_2$ such that $|V_1^\rho(t, x)| \leq K_1(1 + |x|)$ and $|V_2(t, x)| \leq K_2(1 + |x|)$. So, exists $C$ such that

$$\frac{2}{\delta} |x_\delta - y_\delta|^2 \leq C(1 + |x_\delta| + |y_\delta|)$$

(11)

We know $\Phi(t_\delta, 0, 0) \leq \Phi(t_\delta, x_\delta, y_\delta)$, which implies

$$\Phi(t_\delta, 0, 0) \leq V_1^\rho(t_\delta, x_\delta) - V_2(t_\delta, y_\delta) - \frac{1}{\delta} |x_\delta - y_\delta|^2 - \gamma e^{t-s}(x_\delta^2 + y_\delta^2)$$

So, we have

$$\gamma e^{t-s}(x_\delta^2 + y_\delta^2) \leq V_1^\rho(t_\delta, x_\delta) - V_2(t_\delta, y_\delta) - \frac{1}{\delta} |x_\delta - y_\delta|^2 - \Phi(t_\delta, 0, 0) \leq 3C(1 + |x_\delta| + |y_\delta|)$$

and

$$\frac{\gamma e^{t-s}(x_\delta^2 + y_\delta^2)}{1 + |x_\delta| + |y_\delta|} \leq 3C$$

Therefore there exists $C_\gamma$ such that

$$|x_\delta| + |y_\delta| \leq C_\gamma$$
It implies that \((x_\delta, y_\delta)\) is bounded by \(C_\gamma\) and there exists a subsequence \((t_\delta, x_\delta, y_\delta)\) which converges to some \((t_0, x_0, y_0)\). By (11), we conclude
\[
\lim_{\delta \to 0} x_\delta = x_0 = y_0 = \lim_{\delta \to 0} y_\delta \quad \text{and} \quad \lim_{\delta \to 0} t_\delta = t_0
\]

Step 3: Suppose that there exists \((\hat{t}, \hat{x}) \in (0, T] \times \mathcal{D}\) satisfying
\[
V_1(\hat{t}, \hat{x}) \geq V_2(\hat{t}, \hat{x})
\]
and we work for a contradiction. Then there is real \(a > 0\) such that
\[
V_1(\hat{t}, \hat{x}) - V_2(\hat{t}, \hat{x}) = 2a
\]
Equation (10) and the semicontinuities of \(V_1^\rho(t, x)\) and \(V_2(t, x)\) give us
\[
\lim_{\delta \to 0} \frac{2}{\delta} |x_\delta - y_\delta|^2 = 0
\]
Letting \(\delta \to 0\), we have
\[
\lim_{\delta \to 0} \Phi(t_\delta, x_\delta, y_\delta) \leq \lim_{\delta \to 0} (V_1^\rho(t_\delta, x_\delta) - V_2(t_\delta, y_\delta)) \leq \lim_{\delta \to 0} \sup(V_1^\rho(t_\delta, x_\delta) - \lim_{\delta \to 0} \inf(V_2(t_\delta, y_\delta)) \leq V_1^\rho(t_0, x_0) - V_2(t_0, x_0)
\]
also
\[
\Phi(t_\delta, x_\delta, y_\delta) \geq \Phi(\hat{t}, \hat{x}, \hat{x}) \geq V_1^\rho(\hat{t}, \hat{x}) - V_2(\hat{t}, \hat{x}) - \gamma e^{T-\hat{t}}(\hat{x}^2 + \hat{x}^2) \geq V_1(\hat{t}, \hat{x}) - V_2(\hat{t}, \hat{x}) - \frac{\rho}{t + T} - \gamma e^{T-\hat{t}}(\hat{x}^2 + \hat{x}^2) \geq 2\tau - \frac{\rho}{t + T} - \gamma e^{T-\hat{t}}(\hat{x}^2 + \hat{x}^2)
\]
When \(\gamma\) and \(\rho\) are small enough, we have
\[
2a - \frac{\rho}{t + T} - \gamma e^{T-\hat{t}}(\hat{x}^2 + \hat{x}^2) \geq a
\]
So, we obtain
\[
a \leq \Phi(t_\delta, x_\delta, y_\delta)
\]
and
\[
a \leq \lim_{\delta \to 0} \Phi(t_\delta, x_\delta, y_\delta) \leq V_1^\rho(t_0, x_0) - V_2(t_0, x_0)
\]
Since \(V_1 \leq V_2\) on \((\{T\} \times \mathcal{D}) \cup ((0, T] \times \partial \mathcal{D})\), we have
By (19) and \((\in (\{T\} \times D) \cup ((0, T] \times \partial D)\) and hence \((t_\delta, x_\delta, y_\delta)\) is a local maximizer of \(\Phi(t, x, y)\).

Step 4: By Theorem A.1, for \(\epsilon > 0\) there exists \(b_{1\delta}, b_{2\delta}, X_\delta, Y_\delta\) such that

\[
(b_{1\delta}, \frac{2}{\rho}(x_\delta - y_\delta) + 2\gamma e^{T-t_\delta} x_\delta, X_\delta) \in \overline{J}^{2, +}_{(0, T) \times D} V_1^{\rho}(t_\delta, x_\delta),
\]

and

\[
(b_{2\delta}, \frac{2}{\rho}(x_\delta - y_\delta) - 2\gamma e^{T-t_\delta} y_\delta, Y_\delta) \in \overline{J}^{2, -}_{(0, T) \times D} V_2(t_\delta, y_\delta),
\]

Equations (9) and (12) imply that there exists \(c > 0\) such that

\[-b_{1\delta} + H(x_\delta, \frac{2}{\rho}(x_\delta - y_\delta) + 2\gamma e^{T-t_\delta} x_\delta, X_\delta) \leq -c\]

and equation (13) implies

\[-b_{2\delta} + H(y_\delta, \frac{2}{\rho}(x_\delta - y_\delta) - 2\gamma e^{T-t_\delta} y_\delta, Y_\delta) \geq 0\]

From equations (14) and (15)

\[b_{1\delta} - b_{2\delta} + H(y_\delta, \frac{2}{\rho}(x_\delta - y_\delta) - 2\gamma e^{T-t_\delta} y_\delta, Y_\delta) - H(x_\delta, \frac{2}{\rho}(x_\delta - y_\delta) + 2\gamma e^{T-t_\delta} x_\delta, X_\delta) \geq c\]

Applying the maximum principle (Theorem A.1) again, we have

\[-\frac{1}{\epsilon} + \|D^2\phi(t_\delta, x_\delta, y_\delta)\| I \leq \frac{X_\delta}{0} \leq D^2\phi(t_\delta, x_\delta, y_\delta) + \epsilon(D^2\phi(t_\delta, x_\delta, y_\delta))^2\]

\[D^2\phi(t_\delta, x_\delta, y_\delta) = \frac{2}{\delta} \left( \begin{array}{ccc} I_3 & -I_3 & 0 \\ -I_3 & I_3 & 0 \\ 0 & 0 & I_3 \end{array} \right) + 2\gamma e^{T-t_\delta} \left( \begin{array}{ccc} I_3 & -I_3 & 0 \\ -I_3 & I_3 & 0 \\ 0 & 0 & I_3 \end{array} \right)\]

and

\[(D^2\phi(\delta))^2 = \frac{8}{\delta^2} \left( \begin{array}{ccc} I_3 & -I_3 & 0 \\ -I_3 & I_3 & 0 \\ 0 & 0 & I_3 \end{array} \right) + \frac{8\gamma e^{T-t_\delta}}{\delta} \left( \begin{array}{ccc} I_3 & -I_3 & 0 \\ -I_3 & I_3 & 0 \\ 0 & 0 & I_3 \end{array} \right) + 4\gamma e^{2(T-t_\delta)} \left( \begin{array}{ccc} I_3 & 0 \\ 0 & I_3 \end{array} \right)\]

We rewrite

\[x_\delta X_\delta y_\delta y_\delta^T = (x_\delta, y_\delta) \left( \begin{array}{ccc} X_\delta & 0 \\ 0 & -Y_\delta \end{array} \right) \left( \begin{array}{c} x_\delta \\ y_\delta \end{array} \right) \]

\[\leq (x_\delta, y_\delta) \left[ \frac{2}{\delta} \left( \begin{array}{ccc} I_3 & -I_3 & 0 \\ -I_3 & I_3 & 0 \\ 0 & 0 & I_3 \end{array} \right) + (2\gamma e^{T-t_\delta} + 4\epsilon e^{2(T-t_\delta)}) \left( \begin{array}{ccc} I_3 & 0 \\ 0 & I_3 \end{array} \right) + \frac{8 + 8\delta e^{T-t_\delta}}{\delta^2} \left( \begin{array}{ccc} I_3 & -I_3 & 0 \\ -I_3 & I_3 & 0 \\ 0 & 0 & I_3 \end{array} \right) \left( \begin{array}{c} x_\delta \\ y_\delta \end{array} \right) \right]\]
Letting $\gamma \to 0$ and $\epsilon = \frac{\delta}{3}$, we have

$$x_\delta X_\delta^T - y_\delta Y_\delta^T \leq \frac{4}{\delta} (x_\delta - y_\delta)^2;$$

$$y_\delta Y_\delta^T - x_\delta X_\delta^T \geq -\frac{4}{\delta} (x_\delta - y_\delta)^2$$

By (16)

$$H(y_\delta, \frac{2}{\delta} (x_\delta - y_\delta) + 2\gamma e^{T-t} y_\delta, X_\delta) - H(x_\delta, \frac{2}{\delta} (x_\delta - y_\delta) + 2\gamma e^{T-t} x_\delta, X_\delta) \geq b_{2\delta} - b_{1\delta} + c$$

$$\hat{H}(y_\delta, \frac{2}{\delta} (x_\delta - y_\delta) - 2\gamma e^{T-t} y_\delta) - \hat{H}(x_\delta, \frac{2}{\delta} (x_\delta - y_\delta) + 2\gamma e^{T-t} x_\delta)$$

$$\geq \gamma e^{T-t} (x_\delta^2 + y_\delta^2) - \frac{4\sigma^2}{2\delta} (x_\delta - y_\delta)^2 + c$$

Letting $\gamma \to 0$

$$\lim_{\delta \to 0} \frac{2}{\delta} |x_\delta - y_\delta|^2 = 0 \quad \text{and from the continuity of } \hat{H}, \text{ and } \lim_{\delta \to 0} x_\delta = x_0 = \lim_{\delta \to 0} y_\delta, \text{ we have}$$

$$0 = \lim_{\delta \to 0} \left[ \hat{H}(y_\delta, \frac{2}{\delta} (x_\delta - y_\delta)) - \hat{H}(x_\delta, \frac{2}{\delta} (x_\delta - y_\delta)) \right] > c$$

which leads to a contradiction. \hfill \Box

Assume that our value function is continuous and satisfies the linear growth condition and the boundary condition: $V(t, N, 0, B) = U (e^{r(T-t)} B)$ for $0 < t \leq T$. Then the uniqueness of viscosity solutions is obtained from Theorem 3.2.

**Corollary 3.1.** The value function $V^w(t, N, S, B)$ is the unique viscosity solution of (7) on $(0, T) \times \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}$ with the terminal condition $W(T, N, S, B) = U (NS + B - nC_T)$ and the boundary condition $W(t, N, 0, B) = U (e^{r(T-t)} B - nC(0))$ for $0 < t \leq T$. Also, the value function $V(t, N, S, B)$ is the unique viscosity solution of (7) on $(0, T) \times \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}$ with the terminal condition $W(T, N, S, B) = U (NS + B)$ and the boundary condition $W(t, N, 0, B) = U (e^{r(T-t)} B)$ for $0 < t \leq T$.

4. Example and numerical experiments

4.1. Example

To illustrate our model, we provide a simple example which is interesting enough to give us the explicit solution. Consider the functions of trading speed $v_t$, $f(v_t) = 1 + \alpha v_t$ and $g(v_t) = \beta v_t$ for $\alpha > 0$ and $\beta > 0$. Then we have the following SDE for market price and actual traded price:

$$dS_t = (\mu + \beta v_t) S_t dt + \sigma S_t dB_t,$$

$$S(t, v_t, \omega) = (1 + \alpha v_t) S_t.$$
In here $\alpha$ is positive and indicates the depth of illiquidity (the parameter for liquidity costs). $\beta$ is also positive and indicates the permanent price impact factor. We substitute for $f(v_t) = 1 + \alpha v_t$ and $g(v_t) = \beta v_t$ in (7). With a little analysis, we have

$$0 = \max_{v \in [-M,M]} \left\{ -\alpha S \frac{\partial W}{\partial B} v^2 + \left( \frac{\partial W}{\partial N} + \beta S \frac{\partial W}{\partial S} - S \frac{\partial W}{\partial B} \right) v \right\} + \mu S \frac{\partial W}{\partial S} + \frac{\partial W}{\partial t}$$

$$+ rB \frac{\partial W}{\partial B} + \frac{\sigma^2 S^2 \partial^2 W}{2 \partial S^2}$$

Note that $\left\{ -\alpha S \frac{\partial W}{\partial B} v^2 + \left( \frac{\partial W}{\partial N} + \beta S \frac{\partial W}{\partial S} - S \frac{\partial W}{\partial B} \right) v \right\}$ is a quadratic function of $v$. Since $W(t, N, S, B)$ is strictly increasing with respect to $B$, so $\frac{\partial W}{\partial B} > 0$. Also, the fact that $S > 0$ and $\alpha > 0$ gets

$$-\alpha S \frac{\partial W}{\partial B} < 0$$

Therefore, the maximum of $\left\{ -\alpha S \frac{\partial W}{\partial B} v^2 + \left( \frac{\partial W}{\partial N} + \beta S \frac{\partial W}{\partial S} - S \frac{\partial W}{\partial B} \right) v \right\}$ is achieved by

$$v^* = \frac{\frac{\partial W}{\partial N} + \beta \frac{\partial W}{\partial S} - \frac{\partial W}{\partial B}}{2\alpha \frac{\partial W}{\partial B}},$$

$M$, or $-M$. Since $2\alpha \frac{\partial W}{\partial B}$ is always positive, the sign of $v^*$ is determined by that of $\frac{\partial W}{\partial N} + \beta \frac{\partial W}{\partial S} - \frac{\partial W}{\partial B}$. There are three possible cases:

Case (i):

$$\frac{\partial W}{\partial B} > \frac{1}{S} \frac{\partial W}{\partial N} + \beta \frac{\partial W}{\partial S}$$

The optimal solution $v^* < 0$ where the maximum is achieved by selling the stock and increasing our holdings in bank account. Marginal utility per dollar of stock holding plus marginal utility on stock price caused by the permanent price impact factor is less than marginal utility per dollar on bank account. To maximize the utility, it is recommended to transfer money from holding stocks to the bank account.

Case (ii):

$$\frac{\partial W}{\partial B} < \frac{1}{S} \frac{\partial W}{\partial N} + \beta \frac{\partial W}{\partial S}$$

The optimal solution $v^* > 0$ where the maximum is achieved by buying the stock and decreasing our holdings in bank account. Marginal utility per dollar of stock holding plus marginal utility on stock price caused by the permanent price impact factor is greater than marginal utility per dollar on bank account. To maximize the utility, it is recommended to transfer cash from bank account to stock holdings.

Case (iii):

$$\frac{\partial W}{\partial B} = \frac{1}{S} \frac{\partial W}{\partial N} + \beta \frac{\partial W}{\partial S}$$

The optimal solution $v^* = 0$ where the maximum is achieved by doing nothing. Marginal utility per dollar of stock holding plus marginal utility on stock price caused by the permanent price impact factor is equal to marginal utility per dollar on bank account. There is no transaction needed.

If a value function is defined in the 4-dimensional space $(t, N, S, B)$, the optimization problem is a free boundary problem. At a fixed time $t$, the above result suggests that the state space can be divided into buy
and sell regions by a surface. On the surface, the trading speed is 0 and there is no transaction. The buy region is characterized by

\[
\frac{\partial W}{\partial B} - \frac{1}{S} \frac{\partial W}{\partial N} < \beta \frac{\partial W}{\partial S},
\]

and the sell region is characterized by

\[
\frac{\partial W}{\partial B} - \frac{1}{S} \frac{\partial W}{\partial N} > \beta \frac{\partial W}{\partial S}.
\]

We now consider the exponential utility function given by

\[
U(x) = 1 - \exp(-\lambda x)
\]

where the index of risk aversion is \(-\frac{U''(x)}{U'(x)} = \lambda\), independent of the trader’s wealth. The integral version of state variable \(B_T\) is written as

\[
B_T = B_t \exp\left( r(T-t) \right) - \int_t^T e^{r(T-u)} f(v_u) S_u v_u du.
\]

Then

\[
V(t, N, S, B) = \max_{v \in \Gamma} \left\{ 1 - \exp\left( -\lambda (N_T S_T + B_T) \right) \mid N_t = N, S_t = S, B_t = B \right\}
\]

\[
= 1 - \min_{v \in \Gamma} \left\{ \exp\left( -\lambda \left( N_T S_T + B \exp\left( r(T-t) \right) - \int_t^T e^{r(T-u)} f(v_u) S_u v_u du \right) \right) \right\}
\]

\[
= 1 - \exp\left( -\lambda B \exp\left( r(T-t) \right) \right) Q(t, N, S)
\]

where \(Q(t, N, S)\) is a continuous function in \(N\) and \(S\), and defined by \(Q(t, N, S) = 1 - V(t, N, S, 0)\). Then we have

\[
0 = \max_{v \in [-M,M]} \left\{ -\alpha S \lambda \exp(r(T-t))Q(t, N, S)v^2 + \left( \frac{\partial Q}{\partial N} + \beta S \frac{\partial Q}{\partial S} - S \lambda \exp(r(T-t))Q(t, N, S) \right) v \right\}
\]

\[
+ \mu S \frac{\partial Q}{\partial S} + \frac{\partial Q}{\partial t} - B \lambda \exp(r(T-t))r + r B \lambda \exp(r(T-t))Q(t, N, S) \frac{\partial Q}{\partial B} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 Q}{\partial S^2}
\]

\[
Q(T, N, S) = 1 - \exp(-\lambda NS)
\]

and letting \(Q'(t, N, S) = 1 - V'(t, N, S, 0)\),

\[
Q'(T, N, S) = 1 - \exp(-\lambda (NS - nC_T)).
\]

Note that the term in the above PDE

\[
-\alpha S \lambda \exp(r(T-t))Q(t, N, S)v^2 + \left( \frac{\partial Q}{\partial N} + \beta S \frac{\partial Q}{\partial S} - S \lambda \exp(r(T-t))Q(t, N, S) \right) v
\]

is a quadratic function of \(v\). The maximum is achieved by

When we are assuming \( \alpha = 0.0001 \) and \( \beta = 0.001 \). When \( S_0 = 100 \), \( r = 0 \), and \( T = 0.1 \) years. We have observed \( \lambda = 0.000001 \) and \( n = 1000 \). When \( S_0 \) is fixed, \( V^w \) and \( V \) are functions of bank account \( B_0 \) and number \( N_0 \) of shares held. \( Q(t, N, S) \) is the function of \( t, N, S \). We have the following explicit formula for the utility indifference price \( p \):

\[
p = \frac{1}{n\lambda \exp(rT)} \log \frac{Q^w(0, 0, S)}{Q(0, 0, S)}.
\]

In this case, we observe the utility indifference price \( p \) is independent of the trader’s initial wealth.

### 4.2. Numerical experiments

In this section, we discuss the numerical solution to the example and present some results. We compute the utility indifference price of a European call with strike 100 and observe interesting properties. In the numerical experiments, the parameter values that we used are initial stock price \( S_0 = 100 \), \( \mu = 0.05 \), \( r = 0 \), and \( T = 0.1 \) years. We assume \( \lambda = 0.000001 \) and \( n = 1000 \). When \( S_0 \) is fixed, \( V^w \) and \( V \) are functions of bank account \( B_0 \) and number \( N_0 \) of shares held. \( \beta \) is large, a large trader has more influence on the stock price evolution. This can be interpreted as the trader may have the power to manipulate the stock price, to some extent, to maximize his utility. Table 1 and 2 provide a comparison of the option prices for different \( \alpha \), \( \beta \) and number \( n \) of the options written by the trader.

When time \( t \) is fixed, the optimal solution is independent of \( B \) in the case of the exponential utility. Fig. 2 shows the optimal trading speed \( v^* \) for a fixed \( t \). The parameters in this computation are initial stock
Table 1
Utility indifference price for different $\alpha$, $\beta$, and $\sigma$ when $\lambda = 0.00001$.

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>$\alpha = 0.0000$</th>
<th>$\alpha = 0.0001$</th>
<th>$\alpha = 0.0005$</th>
<th>$\alpha = 0.0010$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma = 0.2$</td>
<td>$\sigma = 0.3$</td>
<td>$\sigma = 0.2$</td>
<td>$\sigma = 0.3$</td>
<td>$\sigma = 0.2$</td>
</tr>
<tr>
<td>0.0000</td>
<td>2.8511</td>
<td>4.1832</td>
<td>2.8579</td>
<td>4.2324</td>
</tr>
<tr>
<td>0.0001</td>
<td>2.7527</td>
<td>3.6915</td>
<td>2.7873</td>
<td>4.1424</td>
</tr>
<tr>
<td>0.0002</td>
<td>2.6425</td>
<td>3.1454</td>
<td>2.6789</td>
<td>3.8436</td>
</tr>
<tr>
<td>0.0005</td>
<td>2.3285</td>
<td>1.7884</td>
<td>2.3658</td>
<td>2.6175</td>
</tr>
</tbody>
</table>

Table 2
Utility indifference price for different $\alpha$, $\beta$, and $\sigma$ when $\lambda = 0.00002$.

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>$\alpha = 0.0000$</th>
<th>$\alpha = 0.0001$</th>
<th>$\alpha = 0.0005$</th>
<th>$\alpha = 0.0010$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma = 0.2$</td>
<td>$\sigma = 0.3$</td>
<td>$\sigma = 0.2$</td>
<td>$\sigma = 0.3$</td>
<td>$\sigma = 0.2$</td>
</tr>
<tr>
<td>0.0000</td>
<td>2.9446</td>
<td>4.4286</td>
<td>2.9492</td>
<td>4.4364</td>
</tr>
<tr>
<td>0.0001</td>
<td>2.8961</td>
<td>4.3850</td>
<td>2.9400</td>
<td>4.4279</td>
</tr>
<tr>
<td>0.0002</td>
<td>2.8388</td>
<td>4.3259</td>
<td>2.9097</td>
<td>4.3972</td>
</tr>
<tr>
<td>0.0005</td>
<td>2.6713</td>
<td>4.1517</td>
<td>2.7582</td>
<td>4.2395</td>
</tr>
</tbody>
</table>

Table 3
Utility indifference price for different $\alpha$, $\beta$, and $n$ with $M = 1000$.

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>$\alpha = 0.00001$</th>
<th>$\alpha = 0.00005$</th>
<th>$\alpha = 0.00010$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = 500$</td>
<td>$n = 1000$</td>
<td>$n = 2000$</td>
<td>$n = 500$</td>
</tr>
<tr>
<td>0.0000</td>
<td>2.8180</td>
<td>2.8584</td>
<td>2.9452</td>
</tr>
<tr>
<td>0.0001</td>
<td>2.8144</td>
<td>2.8580</td>
<td>2.9388</td>
</tr>
<tr>
<td>0.0002</td>
<td>2.8112</td>
<td>2.8369</td>
<td>2.8863</td>
</tr>
<tr>
<td>0.0005</td>
<td>2.7182</td>
<td>2.6186</td>
<td>2.4406</td>
</tr>
</tbody>
</table>

Table 4
Utility indifference price with different $\alpha$, $\beta$, and $n$ when $M = 100$.

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>$\alpha = 0.00001$</th>
<th>$\alpha = 0.00005$</th>
<th>$\alpha = 0.00010$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = 500$</td>
<td>$n = 1000$</td>
<td>$n = 2000$</td>
<td>$n = 500$</td>
</tr>
<tr>
<td>0.0000</td>
<td>2.8200</td>
<td>2.8634</td>
<td>2.9508</td>
</tr>
<tr>
<td>0.0001</td>
<td>2.8222</td>
<td>2.8641</td>
<td>2.9498</td>
</tr>
<tr>
<td>0.0002</td>
<td>2.8198</td>
<td>2.8592</td>
<td>2.9441</td>
</tr>
<tr>
<td>0.0005</td>
<td>2.8040</td>
<td>2.8424</td>
<td>2.9268</td>
</tr>
</tbody>
</table>

Fig. 2. Optimal trading speed at fixed time.

price $S_0 = 100$, $\mu = 0.05$, $r = 0$, and $T = 0.1$ years. We assume $\lambda = 0.00001$, $\alpha = 0.0001$, $\beta = 0.00001$ and $n = 500$. Knowing the optimal trading speed, Fig. 3 illustrates the trading regions, in which it is divided into
Fig. 3. Buy and sell regions.

buy and sell regions by a smooth curve. On the curve, the trading speed is 0, and there is no transaction. Above the curve, we have

\[
\frac{1}{S} \frac{\partial Q}{\partial N} + \beta \frac{\partial Q}{\partial S} < \lambda Q(t, N, S)
\]

which corresponds to the buy region. Also, below the curve

\[
\frac{1}{S} \frac{\partial Q}{\partial N} + \beta \frac{\partial Q}{\partial S} > \lambda Q(t, N, S)
\]

which corresponds to the sell region.

4.3. An extended large trader model

In our large trader model, we assumed that its trading action has some lasting impact on the evolution of the stock price, so that the drift term has an impact from its trades. In this subsection, we extend it to the model in which trading speed \( v_t \) has an effect on both the drift and volatility terms of the asset price process. In most cases, large trader’s trading action (of both buying and selling) increases the volatility of the trading asset, so it is quite natural to assume that the volatility increases when \( v_t^2 \) (or \(|v_t|\)) increases. We consider a simple extended large trader model with \( f(v_t) = 1 + \alpha v_t \), \( g(v_t) = \beta v_t \) and \( \sigma(v_t) = \sqrt{\sigma^2 + \gamma v_t^2} \), where \( \alpha > 0 \), \( \beta > 0 \) and \( \gamma > 0 \). Then

\[
dS_t = (\mu + \beta v_t)S_t dt + \sqrt{\sigma^2 + \gamma v_t^2} S_t dB_t, \\
S(t, v_t, \omega) = (1 + \alpha v_t)S_t,
\]

where \( M \leq v_t \leq M \), \( \alpha \) indicates the depth of illiquidity, and \( \beta \) and \( \gamma \) indicate the feedback effect index on drift and volatility, respectively. With a similar analysis as before, we have the following HJB equations:

\[
0 = \max_v \left\{ -\alpha S \frac{\partial W}{\partial B} v^2 + \left( \frac{\partial W}{\partial P} + \beta S \frac{\partial W}{\partial S} - S \frac{\partial W}{\partial B} \right) v + \frac{\gamma v^2 S^2}{2} \frac{\partial^2 W}{\partial S^2} + \frac{\partial W}{\partial t} + rB \frac{\partial W}{\partial B} \right\}
\]

(18)

Table 5 shows the option price when \( \gamma = 5 \times 10^{-8} \) with different values of \( \alpha \) and \( \beta \). When a large trader has the influence on volatility of the stock price, option prices are relatively higher than those when the
Table 5
Utility indifference price when $\gamma = 5 \times 10^{-8}$ with different $\alpha$ and $\beta$.

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>$\alpha = 0.00001$</th>
<th>$\alpha = 0.00005$</th>
<th>$\alpha = 0.00010$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n = 500$</td>
<td>$n = 1000$</td>
<td>$n = 2000$</td>
</tr>
<tr>
<td>0.00000</td>
<td>2.8196</td>
<td>2.8619</td>
<td>2.9490</td>
</tr>
<tr>
<td>0.00001</td>
<td>2.8190</td>
<td>2.8614</td>
<td>2.9483</td>
</tr>
<tr>
<td>0.00002</td>
<td>2.8164</td>
<td>2.8571</td>
<td>2.9443</td>
</tr>
<tr>
<td>0.00005</td>
<td>2.7854</td>
<td>2.8219</td>
<td>2.9129</td>
</tr>
</tbody>
</table>

large trader has no influence on the volatility. This result is illustrated by comparing the option prices in Table 3 and Table 5.

5. Conclusion

In this paper, we investigated option valuation based on utility maximization for a large trader in a market with liquidity risk. We considered two effects of illiquidity; the cost of illiquidity and permanent price impact benefits. In illiquid markets, trading action will incur liquidity costs, but at the same time, the trader can have influence on the stock price evolution and gain benefits from the permanent price impact by choosing the optimal strategy. Thus, the option price, in some sense, is determined by these two contradicting phenomena. When both the permanent impact function and the liquidity cost function are linear in the trading speed, the optimal solution is computed explicitly, and moreover, the state space can be characterized and divided into the buy and sell regions.

Acknowledgment

The research was partially supported by Natural Sciences and Engineering Research Council of Canada, Discovery grant No. 504316.

Appendix A

Theorem A.1. (Crandall, Lions and Ishii) For $i = 1, 2$, let $D_i$ be locally compact subsets of $\mathbb{R}^N$, and $D = D_1 \times D_2$, let $u_i$ be upper semicontinuous in $(0, T) \times D_i$, and $J^{2,+}_{(0, T) \times D_i} u_i(t, x)$ the parabolic superjet of $u_i(t, x)$, and $\phi$ be twice continuously differentiable in a neighborhood of $(0, T) \times D$. Set

$$\omega(t, x_1, x_2) = u_1(t, x_1) + u_2(t, x_2) - \phi(t, x_1, x_2)$$

for $(t, x_1, x_2) \in (0, T) \times D$, and suppose $(\tilde{t}, \tilde{x}_1, \tilde{x}_2)$ is a local maximum of $\omega$ relative to $(0, T) \times D$. Moreover, assume that there is an $r > 0$ such that for every $M > 0$ there exists a $C$ such that for $i = 1, 2$

$$b_i \leq C \text{ whenever } (b_i, q_i, X_i) \in J^{2,+}_{(0, T) \times D_i} u_i(t, x)$$

$$|x_i - \tilde{x}_i| + |t - \tilde{t}| \leq r \text{ and } |u_i(t, x_i)| + |q_i| + \|X_i\| \leq M$$

Then for each $\varepsilon > 0$ there exists $X_i \in S(N)$ such that

(i)

$$(b_i, D_x \phi(\tilde{t}, \tilde{x}), X_i) \in J^{2,+}_{(0, T) \times D_i} u_i(\tilde{t}, \tilde{x}) \text{ for } i = 1, 2$$
(ii)
\[-\frac{1}{\varepsilon} + \|D^2\phi(\hat{x})\| I \leq \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \leq D^2\phi(\hat{x}) + \varepsilon(D^2\phi(\hat{x}))^2\]

(iii)
\[b_1 + b_2 = \frac{\partial\phi(t, \hat{x}_1, \hat{x}_2)}{\partial t}\]

where for a symmetric matrix $A$, $\|A\| := \sup\{x^T A x : \|x\| \leq 1\}$.

References