

The cost of illiquidity and its effects on hedging

L. C. G. Rogers* and Surbjeet Singh
University of Cambridge and University of Bath

August 30, 2007

1 Introduction.

After credit risk, liquidity risk is probably the next most important risk faced by the finance industry; and yet the study of liquidity is far less advanced. This may be in part due to the fact that there is no agreed definition of what liquidity is, even in qualitative terms; everyone would agree that the effect of illiquidity is to make it difficult or costly to trade large volumes of the underlying asset in small times, but there are different approaches to modelling this.

One modelling philosophy is that trading large amounts moves the price, and the papers of [7], [8], [16], [15], [18] are examples of this viewpoint, where the stock price responds instantaneously to the amount of the stock held by a single large trader. Though such models have a flavour of liquidity, we regard them rather as models of price impact effects. There is evidence that the actions of a large trader can influence the price of the underlying (see, for example, [12] and [10]). A large trader can try to ‘corner the market’. One way to do this in a commodity market is to take a huge long futures position and at the same time buy up the underlying commodity. As expiry approaches the investors who are short the futures contract may find that there is not enough supply to meet their demand and hence the price is pushed up. One such alleged case of this was the activities of the Hunt brothers in the silver market in 1979-1980. Their trading caused the price to increase from \$9 per ounce to \$50 per ounce. Another way to ‘corner the market’, this time involving shares, is to buy up a large supply and then lend some to investors who want to go short. When these shares are sold on the market the agent buys them up and then calls in the short shares. Since the agent has limited the supply of shares by buying a large amount, this pushes the price up. These types of manipulation involve taking huge positions in the underlying and documented examples of this type of activity have shown large price increases followed by a crash.

Though interesting in their own right, such price impact models have drawbacks which make them unsuitable as models of liquidity. One of these is the ‘free round trip’ phenomenon, discussed in [18]; if the large agent rapidly sells and then buys back a large amount of stock, he can force the price instantaneously to drop, and if this round trip is not costly (as is the case in some studies), then the large agent could make profits by selling down-and-out calls,

*Corresponding author: Statistical Laboratory, University of Cambridge, Wilberforce Road, Cambridge CB3 0WB, UK; L.C.G.Rogers@statslab.cam.ac.uk

and subsequently knocking them out by a round trip. Another problem with such price impact models is that they typically present the solution to a hedging problem in feedback form, exhibiting the hedge as a function of time, and current stock price - but if the initial portfolio is not at the exactly correct value, it is not clear how it is to be moved to that value. A more serious problem with such models is that if the actions of *one* agent may affect the price, then logically the actions of *all* agents may affect the price, and the resulting analysis of the inter-related behaviour of the many agents in the market becomes impossibly cumbersome; see, for example, [9] for a partial analysis of such a setup. Moreover, a price-impact story does not explain why *small* agents should be affected by illiquidity, which they surely are.

Our viewpoint here is that the effect of illiquidity is a *cost*, but that illiquidity does not affect the price of the underlying asset. It affects the price at which an agent will trade the asset, however, reflecting the depth of the limit order book. The faster an agent wants to buy (sell) the asset, the deeper into the limit order book he will have to go, and higher (lower) will be the price for the later units of the asset bought (sold). However, once a rapid transaction is completed, we suppose that the limit order book quickly fills up again, and that the rapid transaction has no lasting effect on the price of the underlying. What transpires is that the impact of liquidity modelled in this fashion is like a transaction cost, but not one which is *proportional* to the amounts traded, which is the assumption of the traditional proportional transaction cost model (see [14], [6]). This modelling approach was presented in the preprint [17]; the references [5] and [19] explore further aspects of the model introduced there. The preprint of Isaenko [11] proposes a dynamic with some similarities to ours, but the way in which the illiquidity costs scale with stock price is very different.

The paper of Almgren & Chriss [1] is close in spirit to what we do here. They model both a temporary and a permanent market impact, while we choose only to model the temporary effect, for reasons explained above. We consider the trading of only a single asset, but the objective we consider is more general, in that we aim to *hedge any European option*. Almgren & Chriss in effect treat the problem of optimising the quadratic utility of terminal wealth, working in discrete time with a random-walk asset dynamic.

Another related paper is that of Longstaff [13], who proposes that the holding of stock must be a finite-variation process with bounded derivative, which may be thought of as a special case of the model we propose. Models which feature illiquidity costs include [3] and [4]. However, in both these references transaction costs can be completely avoided by following a continuous trading strategy of bounded variation. This seems rather unrealistic. The paper of Bakstein & Howison [2] shares a number of features with ours; one main difference is that it leads to price impact effects, which we are trying to eliminate.

The layout of the remainder of the paper is as follows. Section 2 explains how the dynamics that we shall take as fundamental arise from consideration of the order book. Next, Section 3 deals with the hedging of a European-style option in the Black-Scholes world in the presence of illiquidity costs. The Hamilton-Jacobi-Bellman equation that arises from the associated optimal control problem can be solved in almost closed form. The numerical solution requires us just to solve four parabolic PDEs of Black-Scholes type. We then consider approximate solutions in Section 4; this shows that there is a simple hedging rule which approximates the optimal rule, and as a by-product we learn the order of magnitude of the losses due to illiquidity. Next in Section 5 we explore some numerical solutions, comparing the optimal solution with the approximate solution. Finally, Section 6 concludes.

2 Modelling illiquidity: motivation.

Let us start with a simple *static* model for the order book. We imagine that there is some ‘mid price’ S , and that there is an order book of quotes distributed either side of the mid price S , with density $\rho(x)$ of quotes at relative¹ price x . If an agent wishes to acquire h units of the asset, he will have to buy up through the order book, to relative price s defined by

$$h = \int_1^s \rho(x) dx,$$

which will cost

$$S \int_1^s x\rho(x) dx.$$

Having done this trade, the mid-price quickly returns to S , so the book value of what he has just bought is hS and he will record a loss of

$$(2.1) \quad Sl(h) \equiv S \int_1^s x\rho(x) dx - hS = S \int_1^s (x-1)\rho(x) dx.$$

Notice that the same equations hold whether $s > 1$ or $s < 1$, so that $l \geq 0$, and

$$\frac{dl}{dh} = \frac{(s-1)\rho(s)}{\rho(s)} = (s-1) \geq -1$$

is increasing, and therefore the function l is convex, with slope at least -1 .

EXAMPLES. (i) A natural example would be to take

$$(2.2) \quad l(x) = \frac{e^{\varepsilon x} - 1}{\varepsilon} - x$$

where $\varepsilon > 0$ is fixed.

(ii) Taking

$$l(x) = \varepsilon |x|$$

models a proportional transaction cost as in [14], [6].

(iii) We could take

$$\begin{aligned} l(x) &= 0 & (|x| \leq a) \\ &= \infty & (|x| > a) \end{aligned}$$

to model a situation where any trade up to a certain size would be allowed; this is in effect the situation in [13].

We extend this to a dynamic setting by firstly supposing that trading takes place in discrete intervals of time of length Δt , during each of which there is an order book with density $\rho(x)dx\Delta t$

¹We assume that the only effect of the value of mid-price on the problem is through proportionality; otherwise, high-priced assets would be subject to more (or less) liquidity costs than low-priced assets, which is not an effect we consider realistic.

at price x relative to the prevailing mid-price S_t . As before, an agent who wishes to acquire $h\Delta t$ units of the asset in that time period will book a loss of $S_t l(h)\Delta t$.

After such a trade, parts of the order book will have been swept out; as a result, we may expect that the mid-price will have been moved by the sale/purchase. However, this is a price-impact effect, and, as we explained earlier, we do not intend to model this, as this would make the price dynamics not just history-dependent, but also dependent on the past trading decisions of a potentially large number of agents. One justification for this simplifying assumption is that the quotes swept out would quickly be replaced if the asset were reasonably liquid, which is the situation we are most interested in.

Given this description of the effect of liquidity in a Δt -trading-interval formulation, it is not hard to see how we are going to model the cost of illiquidity in continuous time. We shall suppose that the number H_t of units of the risky asset held at time t should be differentiable, with derivative h_t , and the wealth dynamics are summarised as

$$(2.3) \quad dw_t = r_t w_t dt + H_t(dS_t - r_t S_t dt) - S_t l(h_t) dt$$

$$(2.4) \quad dH_t = h_t dt$$

where r_t is the riskless rate, and S_t is the asset price at time t , given exogenously, and not affected by the trades of any agents. The only part of the dynamics (2.3) which is not entirely conventional is the final term $-S_t l(h_t) dt$ representing the cost of illiquidity. Concerning the function l , we shall make the following assumption:

ASSUMPTION A: the function l is convex and non-negative, $l(0) = 0$.

REMARKS. (i) Note that we do *not* assume that $l' \geq -1$, a restriction which appears natural in view of the way we derived l from the order book description. This is because we shall presently wish to work with $l(x) = \frac{1}{2}\varepsilon x^2$ for reasons of tractability.

(ii) Note that we do *not* claim that there is some weak convergence argument which takes us from the discrete-time model for the cost of liquidity to (2.3), (2.4) as $\Delta t \downarrow 0$; there may well be, but any such argument is tangential to our purpose. The discussion of this Section is offered by way of motivation for (2.3), (2.4), which we take as the axiomatic basis of the main part of the paper. Any reader who is affronted by such informality can simply cover up the whole of this Section before equations (2.3), (2.4), and take those as the starting point of the paper.

3 An optimal control problem and its solution.

The introduction of illiquidity costs makes the market incomplete, so perfect replication is no longer possible. In this Section, we shall study the effect of illiquidity on the hedging of a European-style option when the illiquidity costs are small.

In the Black-Scholes model, the asset dynamics are given by

$$(3.1) \quad dS_t = S_t(\sigma dW_t + \mu dt),$$

and the riskless rate r_t is assumed equal to a constant. The presence of discounting is an inessential complication to the notation, so we shall immediately assume that we are working with discounted asset prices, that is, we shall take $r = 0$. We shall also assume that $\mu = 0$, for

three reasons. Firstly, in the perfectly liquid Black-Scholes world, under the pricing measure, this is the value we would use; if illiquidity costs were small, we should be close to this situation. Secondly, if we are to use any value for μ other than the riskless rate $r = 0$, we would in practice have to estimate this value, which is notoriously hard to estimate with any accuracy or stability. Thirdly, the analysis is already hard enough with this assumption. Under these assumptions, the wealth dynamics become simply

$$(3.2) \quad \begin{aligned} dw_t &= H_t dS_t - S_t l(h_t) dt \\ &= d(H_t S_t) - S_t f(h_t) dt, \end{aligned}$$

$$(3.3) \quad dS_t = \sigma S_t dW_t,$$

$$(3.4) \quad dH_t = h_t dt,$$

where $f(h) \equiv h + l(h)$.

We shall consider the hedging of a European option which expires at time T and pays $G(S_T)$ at that time. The Black-Scholes value of the option at time $t < T$ is $q(t, S_t)$, where q solves the Black-Scholes initial-value PDE

$$(3.5) \quad \mathcal{L}q = 0, \quad q(T, \cdot) = G(\cdot),$$

where

$$(3.6) \quad \mathcal{L} \equiv \frac{1}{2} \sigma^2 S^2 \frac{\partial^2}{\partial S^2} + \frac{\partial}{\partial t}.$$

In the absence of illiquidity costs, the unique time-0 price of this option would be $q(0, S_0)$, and given that initial wealth, the option could be perfectly replicated by using the self-financing portfolio which holds $\theta(t, S_t) \equiv q_S(t, S_t)$ units of the stock at time t .

Once illiquidity costs are introduced, as modelled in Section 2, it is not possible to hold the Black-Scholes portfolio $H_t = \theta(t, S_t)$ at all times t , since this is a process of infinite variation. We must therefore choose a portfolio H which optimises some criterion, which must be chosen to express our two objectives, namely, to finish up close to the payoff $G(S_T)$ of the option, and not to incur large illiquidity costs on the way. We must also specify what assumption will be made about the portfolio at time T , which will typically hold non-zero amounts of the asset. For this, we shall simply suppose that the asset can be sold at spot with no illiquidity losses; in practice, any money owing on the option would be paid out of cash, and the position in the asset could be unwound sufficiently slowly that no significant illiquidity costs would be incurred.

Suppose we hold H_0 units of the asset at time 0, and x_0 in cash. By following the (differentiable adapted) hedging portfolio process $(H_t)_{0 \leq t \leq T}$, the value at time T of the hedge so constructed will be

$$H_0 S_0 + x_0 + \int_0^T H_t dS_t \equiv \xi,$$

and we will have incurred illiquidity costs

$$\int_0^T S_t l(h_t) dt$$

on the way. We therefore propose to minimise the objective

$$\begin{aligned}
\Phi_0 &= \frac{1}{2}E(\xi - G(S_T))^2 + E \int_0^T S_t l(h_t) dt \\
&= \frac{1}{2}(x_0 + H_0 S_0 - q(0, S_0))^2 + E \int_0^T \frac{1}{2}(H_t - \theta(t, S_t))^2 \sigma^2 S_t^2 dt + E \int_0^T S_t l(h_t) dt \\
(3.7) \quad &\equiv \frac{1}{2}(x_0 + H_0 S_0 - q(0, S_0))^2 + \Phi,
\end{aligned}$$

which penalises both the illiquidity costs incurred, and the mean-squared hedging error. Other criteria could be considered; for example, in [17] the utility-indifference price was used. The criterion we use here has the twin virtues of simplicity and reasonable tractability.

TECHNICAL REMARKS. (i) We require a polynomial bound on the Black-Scholes portfolio θ ; we shall assume that

$$(3.8) \quad \text{for some } \gamma > 0, C > 0: \quad |\theta(t, S)| \leq C(1 + S^\gamma) \quad \forall t \in [0, T], S > 0.$$

Notice that this implies that

$$(3.9) \quad E \int_0^T \theta(t, S_t)^2 S_t^2 dt < \infty.$$

(ii) We likewise have to restrict slightly the class of possible controls, and we shall suppose always that $h \in \mathcal{H}$, where

$$(3.10) \quad \mathcal{H} \equiv \{h : E \int_0^T H_u^2 S_u^2 du < \infty\}.$$

This simply requires that our attempt at replicating an L^2 contingent claim should also be restricted to lie in L^2 , a completely natural condition.

Since the values of initial cash x_0 and holding of asset H_0 will usually be given, our goal is to minimise the objective Φ . For this, we define the value function

$$(3.11) \quad V(t, H, S) \equiv \inf_{h \in \mathcal{H}} E \left[\int_t^T \frac{1}{2}(H_u - \theta(u, S_u))^2 \sigma^2 S_u^2 du + \int_t^T S_u l(h_u) du \mid H_t = H, S_t = S \right],$$

and a familiar argument shows that V solves the Hamilton-Jacobi-Bellman (HJB) equation

$$(3.12) \quad \inf_{h \in \mathcal{H}} \left[V_t + hV_H + \frac{1}{2}\sigma^2 S^2 V_{SS} + \frac{1}{2}\sigma^2 S^2 (H - \theta(t, S))^2 + S l(h) \right] = 0.$$

Typically, it is impossible to get very far with an HJB equation; we have to content ourselves with results on existence and uniqueness; or estimates on the solution; or numerical solution. However, in this case, because of the simple assumed form of the objective, we can get a long way if we suppose that

$$(3.13) \quad l(h) = \frac{1}{2}\varepsilon h^2.$$

where ε will be thought of as a small parameter². The minimisation in (3.12) can be carried out explicitly, by taking

$$(3.14) \quad h = -\frac{V_H}{\varepsilon S},$$

giving the non-linear second-order PDE

$$(3.15) \quad \mathcal{L}V + \frac{1}{2}\sigma^2 S^2 (H - \theta)^2 - \frac{V_H^2}{2\varepsilon S} = 0,$$

where we shall omit the arguments of $\theta(t, S)$ unless there is need to state them explicitly. Now the assumed quadratic form of the objective and illiquidity loss function leads us to guess that (3.15) may be solved by a *quadratic* function

$$(3.16) \quad V(t, H, S) = a(t, S)H^2 + b(t, S)H + c(t, S),$$

and this guess will turn out to be correct. Substituting this form into (3.15) gives a quadratic in H , and equating the coefficients to zero gives us three equations for a , b , and c :

$$(3.17) \quad \mathcal{L}a + \frac{1}{2}\sigma^2 S^2 - \frac{2a^2}{\varepsilon S} = 0,$$

$$(3.18) \quad \mathcal{L}b - \sigma^2 S^2 \theta - \frac{2ab}{\varepsilon S} = 0,$$

$$(3.19) \quad \mathcal{L}c + \frac{1}{2}\sigma^2 S^2 \theta^2 - \frac{b^2}{2\varepsilon S} = 0.$$

The first equation (3.17) is a non-linear PDE, but given the solution to that, it is then straightforward to solve (3.18), then (3.19). We therefore focus on (3.17), and show that it has a unique non-negative solution satisfying suitable boundedness criteria.

Lemma 1. *There is a unique non-negative solution a to (3.17) which satisfies the boundedness condition*

$$(3.20) \quad \sup_{S>0, 0 \leq t \leq T} \left| \frac{a(t, S)}{S^2} \right| < \infty.$$

REMARK. The point of the bound (4.8) is that any solution of (3.17) which satisfies this bound will be given by a Feynman-Kac representation:

$$(3.21) \quad a(t, S) = \frac{1}{2}E \left[\int_t^T \exp \left\{ - \int_t^u 2a(v, S_v) \frac{dv}{\varepsilon S_v} \right\} \sigma^2 S_u^2 du \mid S_t = S \right].$$

Proof . We shall construct an approximating sequence of functions which converge to a solution a , and then we shall establish uniqueness. Given a non-negative function $\alpha : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}^+$, we define the function $\Psi(\alpha) : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}^+$ by

$$(3.22) \quad \Psi(\alpha)(t, S) = \frac{1}{2}E \left[\int_t^T \exp \left\{ - \int_t^u 2\alpha(v, S_v) \frac{dv}{\varepsilon S_v} \right\} \sigma^2 S_u^2 du \mid S_t = S \right].$$

²Though convex and non-negative, this choice of l does not have derivative bounded below by -1. If ε is very small, whatever C^2 function l we may choose will look locally like this choice, so we expect that the conclusions we obtain here will apply widely.

Then it is clear from (3.21) that a solution to (3.17) satisfying the boundedness condition (3.20) is a function a such that $\Psi(a) = a$. Notice that Ψ is decreasing: $a \geq \tilde{a} \Rightarrow \Psi(a) \leq \Psi(\tilde{a})$.

We define a sequence of approximations $a^{(n)}$ to a solution by the recipe

$$a^{(0)} \equiv 0, \quad a^{(n+1)} = \Psi(a^{(n)}) \quad (n \geq 0).$$

Notice immediately that $\sup |a^{(1)}(t, S)/S^2|$ is bounded, and that

$$0 = a^{(0)} \leq a^{(2)} \leq a^{(1)}.$$

Hence

$$\begin{aligned} a^{(1)} &\geq a^{(3)} \geq a^{(2)}, \\ a^{(2)} &\leq a^{(4)} \leq a^{(3)}, \end{aligned}$$

and so on: $a^{(2n)} \leq a^{(2n+2)} \leq a^{(2n+1)}$, and $a^{(2n+1)} \geq a^{(2n+3)} \geq a^{(2n+2)}$ for all $n \geq 0$. Thus the sequence $a^{(2n)}$ increases to a limit \underline{a} , and the sequence $a^{(2n+1)}$ decreases to a limit \bar{a} , and we see that

$$\underline{a} \leq \bar{a}.$$

In the limit, we have $\Psi(\underline{a}) = \bar{a}$, and $\Psi(\bar{a}) = \underline{a}$. Thus \underline{a}, \bar{a} satisfy

$$\begin{aligned} \mathcal{L}\bar{a} + \frac{1}{2}\sigma^2 S^2 - \frac{2\underline{a}\bar{a}}{\varepsilon S} &= 0, \\ \mathcal{L}\underline{a} + \frac{1}{2}\sigma^2 S^2 - \frac{2\bar{a}\underline{a}}{\varepsilon S} &= 0. \end{aligned}$$

Hence the difference $f = \bar{a} - \underline{a} \geq 0$ satisfies the PDE $\mathcal{L}f = 0$; using the bound (4.8), which applies to \bar{a} , we deduce that $f \equiv 0$, and so $a = \bar{a} = \underline{a}$ is a solution.

As for uniqueness, if \tilde{a} is any other nonnegative solution to $\Psi(\tilde{a}) = \tilde{a}$, we have

$$\tilde{a} = \Psi(\tilde{a}) \leq \Psi(0) = \Psi(a^{(0)}) = a^{(1)},$$

from which we learn that $\tilde{a} = \Psi(\tilde{a}) \geq \Psi(a^{(1)}) = a^{(2)}$. Continuing, we find that $\tilde{a} \geq a^{(2n)}$ and $\tilde{a} \leq a^{(2n+1)}$ for any $n \geq 0$, so that $\tilde{a} = \bar{a} = \underline{a}$. ■

We summarise the preceding results as follows.

Theorem 1. *Assuming (3.8), the value function V for the problem $\min \Phi$ is of the form (3.16), where a , b , and c are the unique solutions to (3.17), (3.18), (3.19) satisfying bounds of the form (3.8).*

Proof . Notice that

$$\begin{aligned} V(t, H_t, S_t) &\leq \sigma^2 E \left[\int_t^T (H_u^2 + \theta(u, S_u)^2) S_u^2 du \mid \mathcal{F}_t \right] \\ &\leq \sigma^2 E \left[\int_0^T (H_u^2 + \theta(u, S_u)^2) S_u^2 du \mid \mathcal{F}_t \right] \end{aligned}$$

and so for any H for which $E \int_0^T H_u^2 S_u^2 du < \infty$, the process $V(t, H_t, S_t)$ is dominated by a uniformly-integrable martingale.

In view of the bound (3.8) on θ , the Feynman-Kac representation of the solution to the PDE (3.18) is well defined (the integrals converge), and the solution b again satisfies a bound of the form (3.8). The same conclusion holds for c . Thus solving the PDEs (3.17), (3.18), (3.19), gives us a function V defined by (3.16) which solves the PDE (3.15). From this, whatever control h is used, the process

$$Y_t \equiv \int_0^t \frac{1}{2}(H_u - \theta(u, S_u))^2 \sigma^2 S_u^2 du + \int_0^t S_u l(h_u) du + V(t, H_t, S_t)$$

is expressed (by Itô's formula) as a local martingale plus a non-decreasing process. Using a stopping time $\tau \leq T$ which reduces the local martingale, we learn that

$$V(0, H_0, S_0) \leq E \left[\int_0^\tau \frac{1}{2}(H_u - \theta(u, S_u))^2 \sigma^2 S_u^2 du + \int_0^\tau S_u l(h_u) du + V(\tau, H_\tau, S_\tau) \right]$$

Now we let $\tau \uparrow T$; the integrals converge monotonically, and the final term $V(\tau, H_\tau, S_\tau)$ tends to 0 almost surely, and in view of the uniformly-integrable bound, also in L^1 . ■

REMARKS. (i) Notice that the PDE for a does not depend on the derivative to be hedged; the only effect of this is on b and c . This is perhaps not so surprising, because a controls the loss if H is very large; if H is very large, the most important thing is to get H back to somewhere near the Black-Scholes hedge, and it does not matter very much whether the Black-Scholes hedge is 4 or -65.

(ii) Our analysis tells us little about the magnitude of the value function, but we shall see in the next section that (roughly speaking) the magnitude of V is $O(\sqrt{\varepsilon})$.

4 Approximating the optimal solution.

The route we take here is to suppose that the illiquidity losses are small, and find an approximation to the optimal hedging policy. We still suppose that $l(h) = \frac{1}{2}\varepsilon h^2$.

For small ε , the terms in $\mathcal{L}V$ in the HJB equation (3.15) will be small, but the middle term in (3.15) will not be small, so the only way that the whole expression can be zero is if

$$\frac{V_H}{\sqrt{\varepsilon S}} \doteq \sigma S(H - \theta).$$

In view of (3.14) this leads us to consider the following candidate for a good control:

$$(4.1) \quad \bar{h} \equiv -\sigma \sqrt{\frac{S}{\varepsilon}} (H - \theta).$$

REMARKS. (i) Observe that this control \bar{h} , a choice for the rate of change of the holding H of the asset, has natural properties: it will pull H toward the Black-Scholes hedge θ , and will pull

more strongly as ε gets smaller. For large S , it is more important to get the hedging number of units of asset correct, since this corresponds to a larger sum invested in the asset. Thus we find that \bar{h} pulls H harder to θ for large S , but not proportional to S .

(ii) We are going to estimate the value of using $h = \bar{h}$, and show that this is small in a sense to be made precise in Theorem 2. This analysis gives an estimate of the liquidity costs (and the mishedging costs) incurred, so it gives a value for the ‘liquidity premium’ to be charged for selling this option written on a slightly illiquid asset.

Let v denote the value of using the policy \bar{h} : this solves the PDE

$$(4.2) \quad \mathcal{L}v - \frac{\sigma S(H - \theta)}{\sqrt{\varepsilon S}} v_H + \sigma^2 S^2 (H - \theta)^2 = 0, \quad v(T, \cdot, \cdot) = 0.$$

There is the probabilistic (Feynman-Kac) representation of the solution as

$$(4.3) \quad v(t, H, S) = E \left[\int_t^T \sigma^2 S_u^2 (H_u - \theta(u, S_u))^2 du \mid H_t = H, S_t = S \right],$$

where the expectation (4.3) is calculated under the assumed dynamics (3.2), (3.4), taking $h = \bar{h}$. The Ornstein-Uhlenbeck dynamics of H under control \bar{h} imply a solution of the form

$$(4.4) \quad v(t, H, S) = a(t, S)(H - \theta(t, S))^2 + b(t, S)(H - \theta(t, S)) + c(t, S)$$

for functions a , b , and c to be determined. Some routine calculations lead to the following equations for the three unknowns:

$$(4.5) \quad 0 = \mathcal{L}a - 2\sigma\sqrt{\frac{S}{\varepsilon}} a + \sigma^2 S^2,$$

$$(4.6) \quad 0 = \mathcal{L}b - \sigma\sqrt{\frac{S}{\varepsilon}} b + 2\sigma^2 S(a - S a_S) \theta_S,$$

$$(4.7) \quad 0 = \mathcal{L}c + \sigma^2 S \theta_S (b - S b_S) + \sigma^2 S^2 a \theta_S^2.$$

Theorem 2. *Assume that θ_S and θ_{SS} are uniformly bounded. Then the functions a , b and c determining v through (4.4) satisfy the bounds*

$$(4.8) \quad a(t, S) \leq k\varepsilon^2(z + z^{3/2});$$

$$(4.9) \quad |b(t, S)| \leq k\varepsilon^3(z + z^2);$$

$$(4.10) \quad c(t, S) \leq k\varepsilon^4 z^{5/2} (1 + z + \varepsilon(z^{3/2} + z^2))$$

for some positive constant k , where $z \equiv S/\varepsilon$.

REMARKS. (i) Notice a consequence of the bounds of Theorem 2: for fixed $S > 0$, and fixed H ,

$$V(t, H, S) \leq v(t, H, S) \leq \sqrt{\varepsilon} \kappa(t, S),$$

so we have roughly speaking that $V(t, H, S) = O(\sqrt{\varepsilon})$.

(ii) The assumption of uniform bounds on θ_S and θ_{SS} is quite likely far too strong, but it is sufficient for the required result.

The proof of Theorem 2 involves various estimations which are relegated to the Appendix.

5 Numerical solution

In this section, we discuss the numerical solution to the problem, and present some results. Our test problem is the hedging of a standard European call with strike 1; not all the boundedness hypotheses which we imposed to prove our results hold for this example, but this does not stop us computing the solution. The numerical results we obtain are consistent with the theory, suggesting that the hypotheses may be relaxed somewhat.

We firstly solve the original problem (3.15) using the form (3.16) of the solution. This requires us to solve the three coupled PDEs (3.17), (3.18), (3.19). For this, we worked with the variable $y = \log S$, and solved a finite-difference scheme by Crank-Nicolson. The computations were quick and accurate. The second form of the problem which we solved numerically was that studied in Section 4, where the good policy (4.1) is used, and the value function for that policy is found. While the analysis of Section 4 used the form

$$v(t, H, S) = a(t, S)(H - \theta(t, S))^2 + b(t, S)(H - \theta(t, S)) + c(t, S)$$

for the solution, which was necessary to establish bounds on the solution, what we did numerically was to take the solution in the equivalent form

$$v(t, H, S) = a(t, S)H^2 + b(t, S)H + c(t, S)$$

and solve for that instead. The equations we obtain are slightly different, but this time all are linear, and so may be solved easily:

$$(5.1) \quad 0 = \mathcal{L}a - 2\sigma\sqrt{\frac{S}{\varepsilon}}a + \sigma^2 S^2,$$

$$(5.2) \quad 0 = \mathcal{L}b - \sigma\sqrt{\frac{S}{\varepsilon}}(b - 2a\theta) - 2\sigma^2 s^2 \theta,$$

$$(5.3) \quad 0 = \mathcal{L}c + \sigma\sqrt{\frac{S}{\varepsilon}}\theta b + \sigma^2 S^2 \theta^2.$$

The next question is what values we should take for the parameters of the problem. We suppose that $\sigma = 0.25$, but for ε the value we take should reflect realistic levels of liquidity. We asked a quant with experience of trading equities by how much he would expect the buying price to rise if one attempted to purchase 1% of the stock of a major company in one day, and his reply was “A ‘dawn raid’ of 10% or so of the shares will probably propel the market 15% higher. I would say 1.5% for a quiet 1% purchase.”³ From this, we deduce the value $\varepsilon = 0.006$. Figure 1 shows the optimal solution, and the approximate optimal solution, displaying for each the difference between the best hedge (defined to be the value of H which minimises the value function) and the Black-Scholes hedge; and also displaying the value function taking $H = \theta(t, S)$, scaled by $10/\sqrt{\varepsilon}$ to keep things $O(1)$. We see that qualitatively there is little difference between the optimal solution and the approximate optimal solution from Section 4. Notice that deep in or out of the money, the value function evaluated at the Black-Scholes hedge is very small, as one would expect, since at these levels the hedge is correctly set, at either 0 or 1 units of the stock, and there is little chance that it will need to be changed by very much. Near the money, we

³Interestingly, the scaling expressed in this reply matches the conclusions of our model.

expect more changes to be required, and we see that the minimised cost of liquidity is higher. The illiquid hedges are seen to be greater than the Black-Scholes hedge, probably because of convexity of the payoff.

6 Conclusions.

This paper has presented a model for the effects of illiquidity, and explored some of its consequences for the hedging of European-style options. We emphasised the distinction between price-impact effects and illiquidity effects, an important distinction that is sometimes blurred in the literature. Focussing on illiquidity allows us to suppose that hedging decisions do *not* affect price, and therefore allow us to analyse the effect of illiquidity on a small agent who is required to hedge an option. We pose an optimization problem for this agent, which can be solved in almost closed form, requiring the numerical solution of three parabolic PDEs, one of them non-linear. A near-optimal solution provides a good approximation, and useful bounds on the cost of illiquidity, which we deduce is $O(\sqrt{\varepsilon})$.

A Appendix.

Proof of Theorem 2. Throughout the proof, k will denote a positive constant whose (relatively unimportant) value changes from place to place.

The first step is to change to the variable $z \equiv S/\varepsilon$ in equations (4.5), (4.6), (4.7), transforming them to

$$(A.1) \quad 0 = \mathcal{L}\alpha - 2\sigma\sqrt{z}\alpha + \sigma^2\varepsilon^2z^2,$$

$$(A.2) \quad 0 = \mathcal{L}\beta - \sigma\sqrt{z}\beta + 2\sigma^2\varepsilon z(\alpha - z\alpha_z)\theta_S,$$

$$(A.3) \quad 0 = \mathcal{L}\gamma + \sigma^2\varepsilon z\theta_S(\beta - z\beta_z) + \sigma^2\varepsilon^2z^2\alpha\theta_S^2.$$

Here, we write $\alpha(t, z) = a(t, S)$, $\beta(t, z) = b(t, S)$, $\gamma(t, z) = c(t, S)$, and by slight abuse of notation we set

$$\mathcal{L} \equiv \frac{1}{2}\sigma^2z^2\frac{\partial^2}{\partial z^2} + \frac{\partial}{\partial t}.$$

Remarkably, there is an *explicit* solution to (A.1):

$$(A.4) \quad \alpha^{(0)}(t, z) = \alpha^{(0)}(z) = \varepsilon^2 \left(\frac{3\sigma^2}{32}z + \frac{\sigma}{2}z^{3/2} \right),$$

though this does not satisfy the boundary condition $\alpha(T, \cdot) = 0$. The solution to (4.5) which *does* satisfy the boundary condition is expressed as

$$\alpha(t, S) = \alpha^{(0)}(S) - \bar{\alpha}(t, S),$$

where $\bar{\alpha}$ solves

$$\mathcal{L}\bar{\alpha} - 2\sigma\sqrt{z}\bar{\alpha} = 0, \quad \bar{\alpha}(T, z) = \alpha^{(0)}(z).$$

The function $\bar{\alpha}$ therefore has the Feynman-Kac representation as

$$(A.5) \quad \bar{\alpha}(t, z) = E \left[\exp(-A_{t,T}) a^{(0)}(z_T) \mid H_t = H, z_t = z \right].$$

where we define

$$(A.6) \quad A_t \equiv \int_0^t 2\sigma\sqrt{z_u} du, \quad A_{t,s} \equiv A_s - A_t.$$

Elementary estimation of (A.5) using (A.4) gives us the uniform bound (4.8).

We therefore know that a is *small*, in the precise sense given by (4.8). The next step is to show that β is also comparably small, and for this we again use the Feynman-Kac representation of the solution to (4.6)

$$(A.7) \quad \beta(t, S) = E \left[\int_t^T \exp(-\frac{1}{2}A_{t,u}) q(u, z_u) du \mid H_t = H, z_t = z \right],$$

where

$$(A.8) \quad q(u, z) \equiv 2\sigma^2\varepsilon z(\alpha - z\alpha_z)(u, z)\theta_S(u, S).$$

From (A.5) we deduce that

$$z\bar{\alpha}_z = E_t \left[e^{-A_{t,T}} \left\{ z_T \alpha_z^{(0)}(z_T) - \frac{1}{2} A_{t,T} \alpha^{(0)}(z_T) \right\} \right],$$

and hence using the boundedness of θ we have a bound

$$|q(u, z)| \leq k\varepsilon^3 z(z + z^{3/2})$$

for some constant $k > 0$. We may therefore majorise β by the solution to the equation

$$(A.9) \quad \mathcal{L}g - \sigma\sqrt{z}g + k\varepsilon^3 z(z + z^{3/2}) = 0,$$

together with the boundary condition $g(T, \cdot) = 0$. Now one solution to (A.9) can be found explicitly:

$$g^{(0)}(t, z) = k\varepsilon^3 \sigma^{-1} \left(\frac{3\sigma}{8} (1 + \sigma)z + (1 + \sigma)z^{3/2} + z^2 \right).$$

The solution to (A.9) with the required boundary condition is related to $g^{(0)}$ as α is related to $\alpha^{(0)}$; we deduce similarly the bounds

$$|\beta(t, z)| \leq g(t, z) \leq k\varepsilon^3 (z + z^2).$$

Our last task is to show that γ is small in some suitable sense, and in view of Feynman-Kac representation of γ as

$$(A.10) \quad \gamma(t, z) = E \left[\int_t^T \exp(-\frac{1}{2}A_{t,u}) \psi(u, z_u) du \mid H_t = H, z_t = z \right],$$

where

$$(A.11) \quad \psi(u, z) = \sigma^2 \varepsilon z \theta_S (\beta - z\beta_z) + \sigma^2 \varepsilon^2 z^2 \alpha \theta_S^2,$$

the task is equivalently to show that ψ is suitably small. We already have bounds on α and β , but we need bounds on $z\beta_z$. For this, notice that

$$\mathcal{L}(z\beta_z) = z \frac{\partial}{\partial z} (\mathcal{L}\beta)$$

and so

$$\begin{aligned} \mathcal{L}(\beta - z\beta_z) &= z^2 \frac{\partial}{\partial z} \left(-\frac{1}{z} \mathcal{L}\beta \right) \\ &= z^2 \frac{\partial}{\partial S} \left(-\frac{\sigma}{\sqrt{z}} \beta + 2\sigma^2 \varepsilon (\alpha - z\alpha_z) \theta_S \right) \\ &= \sigma\sqrt{z} (\beta - z\beta_z) - \frac{\sigma}{2} \sqrt{z} \beta + 2\sigma^2 \varepsilon^2 z^2 (\alpha - z\alpha_z) \theta_{SS} - 2\sigma^2 \varepsilon z^3 \alpha_z \theta_S. \end{aligned}$$

Rearranging this gives a PDE for $(\beta - z\beta_z)$:

$$\mathcal{L}(\beta - z\beta_z) - \sigma\sqrt{z} (\beta - z\beta_z) = -\frac{\sigma}{2} \sqrt{z} \beta + 2\sigma^2 \varepsilon^2 z^2 (\alpha - z\alpha_z) \theta_{SS} - 2\sigma^2 \varepsilon z^3 \alpha_z \theta_S.$$

We need a bound on α_{zz} , which is derived from the Feynman-Kac representation of α ; we deduce that

$$|z^2\alpha_{zz}| \leq \varepsilon^2(z + z^{3/2}),$$

and hence

$$\begin{aligned} |\beta - z\beta_z| &\leq k\{\varepsilon^3 z^{3/2}(1+z) + \varepsilon^4(z^{7/2} + z^3)\} \\ &\leq k\varepsilon^3 z^{3/2}[1+z + \varepsilon(z^{3/2} + z^2)]. \end{aligned}$$

Using again the Feynman-Kac representation of the solution to (A.3), we deduce that

$$\begin{aligned} |\gamma| &\leq k\varepsilon^4 z^{5/2}[1+z + \varepsilon(z^{3/2} + z^2)] + k\varepsilon^4 z^3(1 + \sqrt{z}) \\ &\leq k\varepsilon^4 z^{5/2}(1+z + \varepsilon(z^{3/2} + z^2)) \end{aligned}$$

as required.

References

- [1] R. Almgren and N. Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, 3:5–39, 2001.
- [2] D. Bakstein and S. Howison. A non-arbitrage liquidity model with observable parameters for derivatives. *Oxford University Preprint*, 2003.
- [3] P. Bank and D. Baum. Hedging and portfolio optimization in financial markets with a large trader. *Mathematical Finance*, 14:1–18, 2004.
- [4] U. Cetin, R. A. Jarrow, and P. Protter. Liquidity risk and arbitrage pricing theory. *Finance and Stochastics*, 8:311–341, 2004.
- [5] U. Cetin and L. C. G. Rogers. Modelling liquidity effects in discrete time. *Mathematical Finance*, page ??, 2006.
- [6] M. H. A. Davis and A. R. Norman. Portfolio selection with transaction costs. *Mathematics of Operations Research*, 15:676–713, 1990.
- [7] R. Frey. Perfect option hedging for a large trader. *Finance & Stochastics*, 2:115–141, 1988.
- [8] R. Frey and A. Stremme. Market volatility and feedback effects from dynamic hedging. *Mathematical Finance*, 7:351–374, 1997.
- [9] J. P. Heritage and L. C. G. Rogers. Large investors, takeovers, and the rule of law. *Monte Carlo Methods and Applications*, 8:357–370, 2002.
- [10] J. C. Hull. *Options, Futures and Other Derivative Securities, 2nd Edition*. Prentice-Hall, Englewood Cliffs, NJ, 1993.
- [11] S. Isaenko. Portfolio choice under convex transaction costs. Technical report, Concordia University, 2005.
- [12] R. A. Jarrow. Market manipulation, bubbles, corners, and short squeezes. *Journal of Financial and Quantitative Analysis*, 27:311–336, 1992.
- [13] F. A. Longstaff. Optimal portfolio choice and the valuation of illiquid securities. *Review of Financial Studies*, 14:407–431, 2001.
- [14] M. J. P. Magill and G. M. Constantinides. Portfolio selection with transaction costs. *Journal of Economic Theory*, 13:245–263, 1976.
- [15] G. Papanicolaou and K. R. Sircar. General black-scholes models accounting for increased market volatility from hedging strategies. *Applied Mathematical Finance*, 5:45–82, 1998.
- [16] E. Platen and M. Schweizer. On feedback effects from hedging derivatives. *Mathematical Finance*, 8:67–84, 1998.
- [17] L. C. G. Rogers and S. Singh. Modelling liquidity and its effects on price. Technical report, Cambridge University, 2004.

- [18] P. J. Schönbucher and P. Wilmott. The feedback effect of hedging in illiquid markets. *SIAM Journal on Applied Mathematics*, 61:232–272, 2000.
- [19] S. Singh. *Modelling Liquidity and the Valuation of American Options using the Dual Method*. PhD thesis, University of Bath, 2005.

Figure 1: Numerical results, $\sigma = 0.25$, $\varepsilon = 0.006\%$.

