SIAM J. CONTROL AND OPTIMIZATION
Vol. 31, No. 2, pp. 470-493, March 1993

006

EUROPEAN OPTION PRICING WITH TRANSACTION COSTS*

MARK H. A. DAVIS†, VASSILIOS G. PANAS†, AND THALEIA ZARIPHOPOULOU‡

This paper is dedicated to Wendell Fleming on the occasion of his 65th birthday.

Abstract. The authors consider the problem of pricing European options in a market model similar to the Black-Scholes one, except that proportional transaction charges are levied on all sales and purchases of stock. "Perfect replication" is no longer possible, and holding an option involves an essential element of risk. A definition of the option writing price is obtained by comparing the maximum utilities available to the writer by trading in the market with and without the obligation to fulfill the terms of an option contract at the exercise time. This definition reduces to the Black-Scholes value when the transaction costs are removed. Computing the price involves solving two stochastic optimal control problems. This paper shows that the value functions of these problems are the unique viscosity solutions, with different boundary conditions, of a fully nonlinear quasi-variational inequality. This fact implies convergence of discretisation schemes based on the "binomial" approximation of the stock price. Computational results are given. In particular, the authors show that, for a long dated option, the writer must charge a premium over the Black-Scholes price that is just equal to the transaction charge for buying one share.

Key words. option pricing, Black-Scholes formula, transaction costs, utility maximisation, stochastic control, free boundary problem, quasi-variational inequality, viscosity solution, Markov chain approximation

AMS(MOS) subject classifications. primary 35R35, 90A16, 93E20; secondary 35R45, 49B60, 90A09

1. Introduction. Option pricing has been a focus of mathematical research in finance since the publication of the Black-Scholes formula in 1973 [3]. Consult Cox and Rubinstein [5] for a full explanation of financial options and their uses. Black and Scholes gave a valuation for a *European call option*, a contract that confers on the holder the right to buy at the *exercise time* T one share of a specified stock at an agreed price E (known as the *strike price*). Let $\mathrm{S}(t)$ denote the stock price at time t. Clearly, the option is worthless if $\mathrm{S}(T) \leqq E$ but has positive value to the holder, and will be exercised, if $\mathrm{S}(T) > E$. The *writer* of the option thus has the obligation to deliver one share at time T for a cash payment of E if $\mathrm{S}(T) > E$. The pricing problem is to determine what a buyer should be prepared to pay at some earlier time t to acquire such an option and how much the writer should charge for issuing it.[1] Since holding an option is certainly a speculative position, it seems at first that the answer to this question must depend on the buyer's or writer's attitude to risk and therefore that there can be no "universal" pricing formula. However, Black and Scholes showed that, in certain circumstances, such a universal formula is indeed possible. Specifically, they assumed that the stock price process $\mathrm{S}(t)$ is a geometric Brownian motion (this is described in § 3, below), that a bank account, i.e., a riskless investment paying interest at a constant rate r, is available, and that funds may be transferred from bank to stock and vice versa without restrictions or costs. Then it turns out that *perfect hedging* is possible: we can form a (time-varying or *dynamic*), self-financing portfolio of holdings in bank and stock whose value at time T is, with probability one, equal to $(\mathrm{S}(T)-E)^+$.

* Received by the authors February 18, 1992; accepted for publication (in revised form) May 4, 1992.
† Department of Electrical and Electronic Engineering, Imperial College, London SW7 2BT, England.
‡ Department of Mathematics and Business School, University of Wisconsin, Madison, Wisconsin 53201.

[1] These are in general two separate calculations, although ultimately the writer and the buyer must of course agree on the same price. The circumstances required for such agreement to be reached are not discussed in this paper, but see § 8 for some further remarks.

(This is the value of the option at time T in the absence of any transaction costs, since if $\mathbf{S}(T) > E$ the option can be exercised and the stock immediately resold, yielding a profit of $\mathbf{S}(T) - E$.) It follows that the value of the option at time $t < T$ is the cash value $\mathbf{w}(t)$ of the hedging portfolio at that time. Indeed, suppose that the option is offered for $z < \mathbf{w}(t)$; then an investor can proceed as follows. He takes a short position in the hedging portfolio, thus acquiring $\mathbf{w}(t)$ at time t, of which z is used to purchase the option, the remainder $(\mathbf{w}(t) - z)$ being invested in the bank. At time T the investment in the bank is worth $x := (\mathbf{w}(t) - z) \exp(r(T-t))$ and by exercising the option and immediately reselling the stock if $\mathbf{S}(T) > E$ the investor acquires $(\mathbf{S}(T) - E)^+$. Since the latter is exactly the amount required to close the short position, a sure profit of x has been made. A similar *arbitrage opportunity* is available to the writer if he is able to command a higher price than $\mathbf{w}(t)$. It is axiomatic that arbitrage opportunities cannot exist (they contravene any concept of market equilibrium), and therefore $\mathbf{w}(t)$ is the unique fair price for the option, from either the buyer's or writer's point of view.

In fact, perfect hedging of very general European contingent claims (ECC) is possible in the Black–Scholes world: if ψ is any function of the stock price trajectory $\{\mathbf{S}(u): t \leqq u \leqq T\}$ whose expectation exists, then there is a dynamic portfolio whose value at time T is exactly ψ; this is the *replicating portfolio.* The ability to replicate arbitrary contingent claims is described as *completeness* of the market. In the Black–Scholes world, completeness ultimately hinges on the *martingale representation property* of Brownian motion; see Karatzas [13] for a full explanation. By the same argument as above, the fair price for an ECC with payoff ψ is the initial endowment of the replicating portfolio.

There is a paradoxical element to the Black–Scholes approach, which has been called the "Catch-22 of option pricing theory": the claims that can be priced are just those that are redundant in that the investor could, in principle, simply take a position in the replicating portfolio rather than actually buy the option. Thus, apparently such options have no reason to exist. The fallacy here is that we do not live in a Black–Scholes world. In particular, the replicating portfolio cannot be implemented exactly, since it involves incessant rebalancing, which is impractical in the face of any form of market friction such as transaction costs. In this paper, we develop a theory of option pricing in which transaction costs are explicitly taken into account. Perfect hedging is no longer possible, and therefore *buying or writing options involves an unavoidable element of risk.* For this reason, a preference-independent valuation is no longer possible, and the investor's or writer's attitude toward risk must be considered.

In this paper, a new definition is given for the writing price of a European option, based on utility maximisation theory. This is a modification of the definition introduced by Hodges and Neuberger [11], using a very similar approach. It is also shown that, if a replicating portfolio exists and the class of trading strategies forms a linear space, then the new definition of the writing price coincides with the Black–Scholes price for the contingent claim; in § 3 the Black–Scholes model is stated as an example of a market model where these conditions hold. In § 4 this model is modified to account for transaction costs. We assume for mathematical tractability that investors trade only in the underlying security, although in the presence of transaction costs they might well wish to invest in other securities also. The new definition involves the value functions of two different stochastic control problems and in § 4 the nonlinear partial differential equation (p.d.e.), satisfied by these value functions, is derived using informal arguments. Then we prove that these value functions are the unique viscosity solutions (with the appropriate boundary conditions) of this nonlinear p.d.e. in § 5, and a discretisation scheme is outlined in § 6, together with the proof of convergence to the unique solution. Finally, § 7 presents the results of this scheme for investors whose preferences are modeled by an exponential utility function (i.e., their index of risk aversion is independent of wealth).

An alternative approach to the pricing problem was outlined by Leland [16], where a hedging strategy is derived based on discrete time rebalancing under transaction costs, but this method is not optimal in any well-defined sense. Also, Bensaid et al. [2] and Edirisinghe, Naik, and Uppal [10] price options using the concept of super-replicating strategies in a discrete time (binomial) model. Our initial reaction is that this approach is unlikely to be viable in a continuous time setting, but this is a question that merits further investigation.

2. Option pricing via utility maximization. In this section, we give a general definition of option price based on utility maximization. This is a modification of work by Hodges and Neuberger [11], who introduced many of the key ideas. The definition can be stated in very general terms, not restricted to any particular market model. The main result of this section, which demonstrates that our approach is well founded, is Theorem 1, which shows that, if perfect replication is possible, then our price coincides with the Black–Scholes price. On the other hand, our definition is applicable in many situations where the Black–Scholes methodology fails.

We consider a time interval $[0, T]$ and a market, which consists of n stocks whose prices $\underset{\sim}{\mathbf{S}}(t)=(\mathbf{S}_1(t),\dots,\mathbf{S}_n(t))$ are assumed to be stochastic processes on a given probability space $(\Omega,\mathcal{F},P)$; their natural filtration is $\mathcal{F}_t=\sigma\{\mathbf{S}_1(u),\dots,\mathbf{S}_n(u)\colon 0\leqq u\leqq t\}$. Investors can also keep their funds in cash, i.e., a risk-free asset, denoted by B. We wish to give a price applicable at time zero for a European option with exercise time T on one of the stocks, say $\mathbf{S}_1(t)$.

Let $\mathcal{T}(B)$ denote the set of *admissible trading strategies* available to an investor who starts at time zero with an amount B in cash and no holdings in stock. We identify an element $\pi\in\mathcal{T}(B)$ with a vector stochastic process $(\mathbf{B}^\pi(t),\ \underset{\sim}{\mathbf{y}}^\pi(t))=(\mathbf{B}^\pi(t),\ \mathbf{y}_1^\pi(t),\dots,\mathbf{y}_n^\pi(t))$, $t\in[0,T]$, where $\mathbf{B}^\pi(t)$ denotes the amount held in cash and $\mathbf{y}_i^\pi(t)$ the *number of shares* of stock i held, $i=1,\dots,n$, over $[0,T]$ (this may or may not be constrained to be an integral number). There may be costs associated with transactions. In particular, $c(\underset{\sim}{y},\underset{\sim}{S})$ is the liquidated cash value of a portfolio vector $\underset{\sim}{y}$; i.e., the residual cash value when long positions $(y_i>0)$ are sold, and short positions $(y_i<0)$ closed. We assume that $c(\underset{\sim}{0},\underset{\sim}{S})=0$.

An option on stock $S_1(t)$ is the right to buy one share at time T at a price E, which may be a constant (in the case of a simple call option), or, more generally, may be an $\mathcal{F}_T$-measurable random variable (allowing for more exotic things such as *look-back* options). We suppose that the option writer forms a portfolio in order to hedge the option and liquidates the portfolio at time T. Suppose that $(B,\underset{\sim}{y},\underset{\sim}{S})$ denote the cash, stock holdings, and stock price vector at time T, respectively. If $S_1\leqq E$, the option is not exercised, and the cash value of the portfolio is $B+c(\underset{\sim}{y},\underset{\sim}{S})$. If $S_1>E$, then the buyer pays the writer E in cash, and the writer delivers one share to the buyer. The value of the writer's portfolio after this transaction is $B+E+c(\underset{\sim}{y}-\underset{\sim}{e}_1,\underset{\sim}{S})$, where $\underset{\sim}{e}_1$ denotes the vector $(1,0,\dots,0)$.

Let $\mathcal{U}:\mathbb{R}\to\mathbb{R}$, the writer's utility function, be a concave increasing function such that $\mathcal{U}(0)=0$. It is important that $\mathcal{U}(x)$ is defined for both positive and negative x. Now define the following value function:

$$V_w(B)=\sup_{\pi\in\mathcal{T}(B)}\mathbb{E}\{\mathcal{U}(\mathbf{B}^\pi(T)+I_{(S_1(T)\leqq E)}c(\underset{\sim}{\mathbf{y}}^\pi(T),\underset{\sim}{\mathbf{S}}(T))$$
$$+I_{(S_1(T)>E)}[c(\underset{\sim}{\mathbf{y}}^\pi(T)-\underset{\sim}{e}_1,\underset{\sim}{\mathbf{S}}(T))+E])\},\tag{2.1}$$

where $\mathbb{E}$ denotes expectation, and I_A is the indicator function of the event A. We assume that $V_w(B)<\infty$ for all $B\in\mathbb{R}$ and that $V_w(B)$ is a continuous and monotone-increasing function of B. Note that $V_w(B)$ is the maximum utility available to the writer at time T by liquidating his portfolio after his obligations to the buyer have been met, given an initial endowment B. Now define

$$B_w=\inf\{B\colon V_w(B)\geqq 0\}. \tag{2.2}$$

The writer is thus indifferent between (a) doing nothing and (b) accepting B_w and writing the option. This is, however, not the correct comparison for determining a fair option price. Let

$$V_1(B)=\sup_{\pi\in\mathcal{T}(B)}\mathbb{E}\{\mathcal{U}(\mathbf{B}^{\pi}(T)+c(\underline{\mathbf{y}}^{\pi}(T),\underline{\mathbf{S}}(T)))\} \tag{2.3}$$

and define the initial endowment B_1 by

$$B_1=\inf\{B\colon V_1(B)\geqq 0\}, \tag{2.4}$$

where $B_1\leqq 0$, since clearly $V_1(0)\geqq 0$. Think of $(-B_1)$ as the "entry fee" that the writer is prepared to pay to get into the market. Our definition of the option writing price p_w is now

$$p_w=B_w-B_1. \tag{2.5}$$

At this price, the writer is indifferent between going into the market to hedge the option and going into the market strictly on his own account. Note that in hedging the option in this way the writer may well hold stocks other than the one on which the option is written.

A primary justification for this definition is that it reduces to the Black–Scholes valuation in cases where this is applicable. Given an option contract on S_1, a *replicating portfolio* for the writer is an element $\hat\pi\in\mathcal{T}(\hat B)$, for an initial endowment $\hat B$, such that $(\mathbf{B}^{\hat\pi}(T),\underline{\mathbf{y}}^{\hat\pi}(T))=I_{(S_1(T)>E)}(-E,\underline{e}_1)$.

THEOREM 1. *Suppose that the class $\mathcal{T}$ is a linear space; i.e., that if $\pi_i\in\mathcal{T}(B_i)$, $i=1,2$ then $a\pi_1+b\pi_2\in\mathcal{T}(aB_1+bB_2)$ for $a,b\in\mathbb{R}$, where $a\pi_1+b\pi_2=(a\mathbf{B}^{\pi_1}(t)+b\mathbf{B}^{\pi_2}(t),$ $a\underline{\mathbf{y}}^{\pi_1}(t)+b\underline{\mathbf{y}}^{\pi_2}(t))$. Suppose also that both $V_w(B)$ and $V_1(B)$ are continuous and strictly increasing functions of B. Then $p_w=\hat B$ if a replicating portfolio $\hat\pi\in\mathcal{T}(\hat B)$ exists.*

Proof. It follows from the linearity of $\mathcal{T}$ that an arbitrary trading strategy $\pi\in\mathcal{T}(B)$ can always be written in the form $\pi=\hat\pi+\tilde\pi$, where $\hat\pi\in\mathcal{T}(\hat B)$ and $\tilde\pi\in\mathcal{T}(\tilde B)$ with $B=\hat B+\tilde B$. Thus, by the continuity and monotonicity assumptions,

$$\begin{aligned}
0&=V_w(B_w)\\
&=\sup_{\pi\in\mathcal{T}(B_w)}\mathbb{E}\{\mathcal{U}(\mathbf{B}^{\pi}(T)+I_{(S_1(T)\leqq E)}c(\underline{\mathbf{y}}^{\pi}(T),\underline{\mathbf{S}}(T))\\
&\qquad\qquad+I_{(S_1(T)>E)}[E+c(\underline{\mathbf{y}}^{\pi}(T)-\underline{e}_1,\underline{\mathbf{S}}(T))])\}\\
&=\sup_{\tilde\pi\in\mathcal{T}(B_w-\hat B)}\mathbb{E}\{\mathcal{U}(\mathbf{B}^{\tilde\pi}(T)+\mathbf{B}^{\hat\pi}(T)+I_{(S_1(T)\leqq E)}c(\underline{\mathbf{y}}^{\tilde\pi}(T)+\underline{\mathbf{y}}^{\hat\pi}(T),\underline{\mathbf{S}}(T))\\
&\qquad\qquad+I_{(S_1(T)>E)}[E+c(\underline{\mathbf{y}}^{\tilde\pi}(T)+\underline{\mathbf{y}}^{\hat\pi}(T)-\underline{e}_1,\underline{\mathbf{S}}(T))])\}\\
&=\sup_{\tilde\pi\in\mathcal{T}(B_w-\hat B)}\mathbb{E}\{\mathcal{U}(\mathbf{B}^{\tilde\pi}(T)-EI_{(S_1(T)>E)}+I_{(S_1(T)\leqq E)}c(\underline{\mathbf{y}}^{\tilde\pi}(T),\underline{\mathbf{S}}(T))\\
&\qquad\qquad+I_{(S_1(T)>E)}[E+c(\underline{\mathbf{y}}^{\tilde\pi}(T),\underline{\mathbf{S}}(T))])\}\\
&=\sup_{\tilde\pi\in\mathcal{T}(B_w-\hat B)}\mathbb{E}\{\mathcal{U}(\mathbf{B}^{\tilde\pi}(T)+c(\underline{\mathbf{y}}^{\tilde\pi}(T),\underline{\mathbf{S}}(T)))\}\\
&=V_1(B_w-\hat B).
\end{aligned}$$

It follows that $B_1 = \inf\{B: V_1(B) \geqq 0\}$ is equal to $(B_w - \hat{B})$ and hence that $\hat{B} = B_w - B_1 = p_w$. $\square$

3. Option pricing without transaction costs. In this section, we show that the standard Black–Scholes model satisfies theconditions of Theorem 1. The stock price and the price of the amount in the bank are described by differential equations, and it is assumed that there are no transaction costs: $c(\underline{y}, \underline{S}) = \underline{y}^{\mathrm{T}}\underline{S}$. Several assumptions are made on the admissible trading strategies, which imply that $\mathcal{T}(B)$ is a linear space. For a more detailed presentation of a similar market model, refer to Karatzas [13].

The price of a stock $\mathbf{S}_i(t), i = 1, \ldots, n$, is modeled by the following Markov diffusion process:

$$\mathrm{d}\mathbf{S}_i(t) = \mathbf{S}_i(t)(\boldsymbol{\alpha}_i(t)\,\mathrm{d}t + \underline{\boldsymbol{\sigma}}_i^{\mathrm{T}}(t)\,\mathrm{d}\underline{\mathbf{R}}(t)), \tag{3.1}$$

and the price of the amount in the bank is modeled by the following ordinary differential equation (o.d.e.):

$$\mathrm{d}\mathbf{B}(t) = \mathbf{r}(t)\mathbf{B}(t)\,\mathrm{d}t, \tag{3.2}$$

where $\underline{\mathbf{R}}(t)$ is an n-dimensional P-Brownian motion, which generates the filtration, $\mathcal{F}_t$, to which $\mathbf{S}_i(t)$ is adapted, $\boldsymbol{\alpha}_i(t)$ is the mean growth rate of $\mathbf{S}_i(t)$, $\underline{\boldsymbol{\sigma}}_i^{\mathrm{T}}(t)$ is the ith row of the $n \times n$ volatility matrix $\boldsymbol{\sigma}$, and $\mathbf{r}(t)$ is the interest rate. All of these are stochastic processes adapted to $\mathcal{F}_t$. It is assumed that both $\boldsymbol{\alpha}_i(t)$ and $\mathbf{r}(t)$ are uniformly bounded and that

$$\exists \varepsilon > 0 \qquad \text{such that } \boldsymbol{\sigma}(t)\boldsymbol{\sigma}^{\mathrm{T}}(t) > \varepsilon I \quad \forall t \in [0, T], \tag{3.3}$$

where I is the $n \times n$ identity matrix. This last condition is known as the nondegeneracy condition, and it implies that at least one of the n sources of uncertainty in the model affects the price of $\mathbf{S}_i(t)$, for each $i = 1, \ldots, n$. The set of admissible trading strategies $\mathcal{T}(B)$ consists of the $(n+1)$-dimensional, right-continuous, measurable, adapted processes, $(\mathbf{B}^{\pi}(t), \underline{\mathbf{y}}^{\pi}(t))$, such that the investor's wealth is bounded below by a nonpositive integrable random variable. The wealth $\mathbf{W}(t)$ obeys the following stochastic differential equation:

$$\begin{aligned} d\mathbf{W}(t) &= d\mathbf{B}(t) + \underline{\mathbf{y}}^{\mathrm{T}}(t)\,d\underline{\mathbf{S}}(t) \\ &= \mathbf{r}(t)\mathbf{B}(t)\,dt + \sum_{i=1}^{n} \mathbf{y}_i(t)\mathbf{S}_i(t)\boldsymbol{\alpha}_i(t)\,dt + \sum_{i=1}^{n} \mathbf{y}_i(t)\mathbf{S}_i(t)\underline{\boldsymbol{\sigma}}_i^{\mathrm{T}}(t)\,d\underline{\mathbf{R}}(t), \end{aligned} \tag{3.4}$$

which can also be written as

$$d\mathbf{W}(t) = \mathbf{r}(t)\mathbf{W}(t)\,dt + \sum_{i=1}^{n} \mathbf{y}_i(t)\mathbf{S}_i(t)\underline{\boldsymbol{\sigma}}_i^{\mathrm{T}}(t)\,d\tilde{\underline{\mathbf{R}}}(t), \tag{3.5}$$

where $\tilde{\underline{\mathbf{R}}}(t)$ is a Brownian motion with drift.

That $\mathcal{T}(B)$ is a linear space can be verified directly from (3.5). Also, all the value functions are concave and increasing functions of B, as the utility function is a concave and increasing function of the investor's wealth. It can then be easily derived that both value functions are continuous functions of B, and Theorem 1 holds for all the contingent claims, $\psi = (S_1(T) - E)^+$, such that $\mathbb{E}\{\psi^{\nu}\} < \infty$ for some $\nu > 1$.

Remark. The validity of our price definition may alternatively be proved by deriving expressions for the value functions, following the steps in Karatzas [13].

4. Transaction costs: The Bellman equation for the value functions. It is now assumed that investors must pay transaction costs, which are proportional to the amount transferred from the stock to the bank. A market model, similar to that of Davis and Norman [8], is then developed, based on the model outlined in the previous section. The main purpose of this section is the derivation of the fully nonlinear p.d.e., actually a variational inequality, satisfied by all the value functions of the utility maximisation problems stated in § 2. Also, a special utility function is defined, the properties of which enable us to determine the dependence of the value functions on the initial endowment B, and thus reduce the dimensionality of the problem.

It is assumed that investors trade only in the underlying stock $\mathbf{S}(t)$, on which the contingent claim is written, to further reduce the dimensionality of the problem. Note that, in general, investors may wish to trade in all the risky securities available in the market to maximise their performance criteria. The cash value of a number of shares $\mathbf{y}(t)$ of the stock is

$$c(\mathbf{y}(t),\mathbf{S}(t))=\begin{cases}(1+\lambda)\mathbf{y}(t)\mathbf{S}(t), & \text{if } \mathbf{y}(t)<0,\\ (1-\mu)\mathbf{y}(t)\mathbf{S}(t), & \text{if } \mathbf{y}(t)\geqq 0.\end{cases} \tag{4.1}$$

where λ and μ are the fraction of the traded amount in stock, which the investor pays in transaction costs when buying or selling stock, respectively. The time interval considered is $[0, T]$, and the market model equations are

$$d\mathbf{B}(t)=r\mathbf{B}(t)\,dt-(1+\lambda)\mathbf{S}(t)\,d\mathbf{L}(t)+(1-\mu)\mathbf{S}(t)\,d\mathbf{M}(t), \tag{4.2}$$

$$d\mathbf{y}(t)=d\mathbf{L}(t)-d\mathbf{M}(t), \tag{4.3}$$

$$d\mathbf{S}(t)=\mathbf{S}(t)(\alpha\,dt+\sigma\,d\mathbf{R}(t)), \tag{4.4}$$

where $\mathbf{L}(t)$ and $\mathbf{M}(t)$ are the cumulative number of shares bought or sold, respectively, over $[0, T]$ by an investor, $\mathbf{R}(t)$ is a P-Brownian motion that represents the single source of uncertainty, and r, α, and σ are nonrandom constants. As before, $\mathscr{F}_t$ denotes the natural filtration of $\mathbf{R}(t)$. This system of equations describes a degenerate diffusion in $\mathbb{R}^3$.

DEFINITION 1. The set of trading strategies $\mathscr{T}(B)$ consists of all the two-dimensional, right-continuous, measurable processes $(\mathbf{B}^{\pi}(t), \mathbf{y}^{\pi}(t))$, which are the solution of equations (4.2)-(4.4), corresponding to some pair of right-continuous, measurable, $\mathscr{F}_t$-adapted, increasing processes $(\mathbf{L}(t), \mathbf{M}(t))$, such that

$$(\mathbf{B}^{\pi}(t),\mathbf{y}^{\pi}(t),\mathbf{S}(t))\in\mathscr{E}_K \quad \forall t\in[0, T], \tag{4.5}$$

where K is a constant, which may depend on the policy π, and

$$\mathscr{E}_K=\{(B, y, S)\in\mathbb{R}\times\mathbb{R}\times\mathbb{R}^+ \colon B+c(y, S)>-K\}. \tag{4.6}$$

By convention, $\mathbf{L}(0-)=\mathbf{M}(0-)=0$.

Remark. Investors may start with any combination of the two assets at time $s\in[0, T]$ in the general utility maximisaton problem, and in that case the class of admissible trading strategies depends on the time s and the initial portfolio, which is characterised by the initial amount in the bank B, an initial number of shares y, and the initial value of the stock S. The constraint (4.5) is required for technical reasons in § 5, and it only rules out strategies that are clearly nonoptimal, as the objective is the maximisation of the utility of final wealth. Also, either $\mathbf{L}(0)$ or $\mathbf{M}(0)$ may be positive; i.e., a jump at the initial time is possible. Finally, (4.2) and (4.3) imply that the trading strategies are self-financing.

Now define the following two functions of wealth at the final time:

$$\Phi_1(T,\mathbf{B}(T),\mathbf{y}(T),\mathbf{S}(T))=\mathbf{B}(T)+c(\mathbf{y}(T),\mathbf{S}(T)) \tag{4.7}$$

and

$$\Phi_w(T, \mathbf{B}(T), \mathbf{y}(T), \mathbf{S}(T)) = \mathbf{B}(T) + I_{(S_1(T) \leqq E)} c(\mathbf{y}(T), \mathbf{S}(T)) + I_{(S_1(T) > E)}[c(\mathbf{y}(T) - 1, \mathbf{S}(T)) + E] \tag{4.8}$$

and the following value functions:

$$V_j(s, B, y, S) = \sup_{\pi \in \mathcal{T}} \mathbb{E}\{\mathcal{U}(\Phi_j(T, \mathbf{B}^\pi(T), \mathbf{y}^\pi(T), \mathbf{S}^\pi(T)))\} \tag{4.9}$$

where $(s, B, y, S) \in [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^+$, and the index j is 1 or w.

From these definitions, it is evident that the dynamic programming algorithm will yield the same p.d.e. for each value function, the terminal condition of which is determined by the utility of the two functions $\Phi_j(T, B, y, S)$, where j is 1 or w. In the following, we derive, at least formally, the Hamilton-Jacobi-Bellman equations, associated with the two stochastic control problems,[2] which prove to be variational inequalities with gradient constraints. Consider, temporarily, a smaller class of trading strategies $\mathcal{T}'$, such that $\mathbf{L}(t)$ and $\mathbf{M}(t)$ are absolutely continuous processes, given by

$$\mathbf{L}(t) = \int_s^t \mathbf{l}(\xi)\, d\xi \quad \text{and} \quad \mathbf{M}(t) = \int_s^t \mathbf{m}(\xi)\, d\xi, \tag{4.10}$$

where $\mathbf{l}(\xi)$ and $\mathbf{m}(\xi)$ are positive and uniformly bounded by $k < \infty$. Then (4.2)-(4.4) is a vector stochastic differential equation with controlled drift, and the Bellman equation for a value function denoted by V_j^k is

$$\max_{0 \leqq l, m \leqq k} \left\{ \left(\frac{\partial V_j^k}{\partial y} - (1+\lambda) S \frac{\partial V_j^k}{\partial B} \right) l - \left(\frac{\partial V_j^k}{\partial y} - (1-\mu) S \frac{\partial V_j^k}{\partial B} \right) m \right\} + \frac{\partial V_j^k}{\partial s} + rB \frac{\partial V_j^k}{\partial B} + \alpha S \frac{\partial V_j^k}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V_j^k}{\partial S^2} = 0 \tag{4.11}$$

for $(s, B, y, S) \in [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^+$. The optimal trading strategy is determined by considering the following three possible cases:

$$\frac{\partial V_j^k}{\partial y} - (1+\lambda) S \frac{\partial V_j^k}{\partial B} \geqq 0 \quad \text{and} \quad \frac{\partial V_j^k}{\partial y} - (1-\mu) S \frac{\partial V_j^k}{\partial B} > 0, \tag{4.12}$$

where the maximum is achieved by $m = 0$ and buying at the maximum possible rate $l = k$;

$$\frac{\partial V_j^k}{\partial y} - (1+\lambda) S \frac{\partial V_j^k}{\partial B} < 0 \quad \text{and} \quad \frac{\partial V_j^k}{\partial y} - (1-\mu) S \frac{\partial V_j^k}{\partial B} \leqq 0, \tag{4.13}$$

where the maximum is achieved by $l = 0$ and selling at the maximum possible rate $m = k$; and

$$\frac{\partial V_j^k}{\partial y} - (1+\lambda) S \frac{\partial V_j^k}{\partial B} \leqq 0 \quad \text{and} \quad \frac{\partial V_j^k}{\partial y} - (1-\mu) S \frac{\partial V_j^k}{\partial B} \geqq 0, \tag{4.14}$$

where the maximum is achieved by doing nothing; that is $m = 0$ and $l = 0$. (Note that in this case the process $(\mathbf{B}^\pi(t), \mathbf{y}^\pi(t), \mathbf{S}(t))$ becomes an uncontrolled diffusion, which drifts under the influence of the stock process only.) All the other permutations of inequalities are impossible, as all the value functions are increasing functions of B and y.

[2] The argument is very similar to that in Davis and Norman [8].

The above results suggest that the optimisation problem is a free boundary problem, where, if a value function is known in the four-dimensional space, defined by the state of the investor (s, B, y, S), the optimal trading strategy is determined by the above inequalities. Also, the state space is divded into three regions, called the *buy*, *sell* and *no-transaction* regions, which are characterised by (4.12), (4.13), and (4.14), respectively. Clearly, the buy and sell regions do not intersect, as it is not optimal to buy and sell at the same time. The boundaries between the no-transaction region and the buy and sell regions are denoted by ∂B and ∂S.

As $k \to \infty$, the class of admissible trading strategies becomes the class defined at the beginning of this section. It is conjectured that the state space remains divided into a buy region, a sell region, and a no-transaction region, and the optimal trading strategy mandates an immediate transaction to ∂B or ∂S if the state is in the buy region or sell region, followed by transactions of "local time" type at ∂B and ∂S. Therefore, each of the value functions satisfies the following set of equations:

(i) In the buy region, the value function remains constant along the path of the state, dictated by the optimal trading strategy, and therefore

$$V_j(s, B, y, S) = V_j(s, B-(1+\lambda)S\,\delta y_b, y+\delta y_b, S), \tag{4.15}$$

where δy_b (the number of shares bought by the investor) can take any positive value up to the number required to take the state to ∂B. Allowing $\delta y_b \to 0$, (4.15) becomes

$$\frac{\partial V_j}{\partial y} - (1+\lambda)S\frac{\partial V_j}{\partial B} = 0. \tag{4.16}$$

(ii) Similarly, in the sell region, the value function obeys the following equation:

$$V_j(s, B, y, S) = V_j(s, B+(1-\mu)S\,\delta y_s, y-\delta y_s, S), \tag{4.17}$$

where δy_s (the number of shares sold by the investor) can take any positive value up to the number required to take the state to ∂S. Allowing $\delta y_s \to 0$, (4.17) becomes

$$\frac{\partial V_j}{\partial y} - (1-\mu)S\frac{\partial V_j}{\partial B} = 0. \tag{4.18}$$

(iii) In the no-transaction region, the value function obeys the same set of equations obtained for the class of absolutely continuous trading strategies, and therefore the value function is given by

$$\frac{\partial V_j}{\partial s} + rB\frac{\partial V_j}{\partial B} + \alpha S\frac{\partial V_j}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V_j}{\partial S^2} = 0, \tag{4.19}$$

and the pair of inequalities, shown above in (4.14), also hold. Note that, due to the continuity of the value function, if it is known in the no-transaction region, it can be determined in both the buy and sell regions by (4.15) and (4.17), respectively.

In the buy region the left-hand side of (4.18) is negative, and, in the sell region, the left-hand side of (4.16) is positive. Also, from the two pairs of inequalities (4.12) and (4.13), it is conjectured that the left-hand side of (4.19) is negative in both the buy and sell regions. Therefore, the above set of equations is condensed into the following fully nonlinear p.d.e.:

$$\begin{aligned}\max\Bigg\{&\frac{\partial V_j}{\partial y} - (1+\lambda)S\frac{\partial V_j}{\partial B}, -\left(\frac{\partial V_j}{\partial y} - (1-\mu)S\frac{\partial V_j}{\partial B}\right),\\ &\frac{\partial V_j}{\partial s} + rB\frac{\partial V_j}{\partial B} + \alpha S\frac{\partial V_j}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V_j}{\partial S^2}\Bigg\} = 0\end{aligned} \tag{4.20}$$

for $(s, B, y, S) \in [0, T]\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}^+$.

Now consider the exponential utility function given by

$$\mathcal{U}(x) = 1 - \exp(-\gamma x), \tag{4.21}$$

where the index of risk aversion is $-\mathcal{U}''(x)/\mathcal{U}'(x) = \gamma$, which is independent of the investor's wealth. The definition of the value functions (4.9) can be written as

$$V_j(s, B, y, S) = 1 - \inf_{\pi \in \mathcal{T}} \mathbb{E}\{\exp(-\gamma \mathbf{B}^\pi(T)) \exp(-\gamma \Psi_j(T, \mathbf{y}^\pi(T), \mathbf{S}^\pi(T)))\}, \tag{4.22}$$

where $\Psi_j(T, \mathbf{y}^\pi(T), \mathbf{S}^\pi(T)) := \Phi_j(T, \mathbf{B}^\pi(T), \mathbf{y}^\pi(T), \mathbf{S}^\pi(T)) - \mathbf{B}^\pi(T)$, and the amount $\mathbf{B}^\pi(T)$ is given by the following integral version of the state equation (4.2):

$$\mathbf{B}(T) = \frac{B}{\delta(T, s)} - \int_s^T \frac{(1+\lambda)\mathbf{S}(t)}{\delta(T, t)} \, d\mathbf{L}(t) + \int_s^T \frac{(1-\mu)\mathbf{S}(t)}{\delta(T, t)} \, d\mathbf{M}(t), \tag{4.23}$$

where $\delta(T, t)$ is the discount factor, defined by

$$\delta(T, t) = \exp(-r(T-t)) \tag{4.24}$$

for the constant interest rate r. Therefore,

$$V_j(s, B, y, S) = 1 - \exp\left(-\gamma \frac{B}{\delta(T, s)}\right) Q_j(s, y, S), \tag{4.25}$$

where $Q_j(s, y, S)$ is a convex nonincreasing continuous function in y and S, defined by $Q_j(s, y, S) = 1 - V_j(s, 0, y, S)$. This representation means that, at time s, the monetary amount invested in the risky asset is independent of the total wealth. It also entails two very important simplifications. First, the writing price is given by the following explicit function of $Q_1(s, 0, S)$ and $Q_w(s, 0, S)$:

$$p_w(s, S) = \frac{\delta(T, s)}{\gamma} \ln\left(\frac{Q_w(s, 0, S)}{Q_1(s, 0, S)}\right). \tag{4.26}$$

Second, (4.20) is transformed into the following p.d.e. for $Q_j(s, y, S)$:

$$\min\left\{\frac{\partial Q_j}{\partial y} + \frac{\gamma(1+\lambda)S}{\delta(T, s)} Q_j, -\left(\frac{\partial Q_j}{\partial y} + \frac{\gamma(1-\mu)S}{\delta(T, s)} Q_j\right), \frac{\partial Q_j}{\partial s} + \alpha S \frac{\partial Q_j}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 Q_j}{\partial S^2}\right\} = 0 \tag{4.27}$$

with boundary conditions, for $Q_1(s, 0, S)$ and $Q_w(s, 0, S)$, given by

$$Q_1(T, y, S) = \exp(-\gamma c(y, S)) \tag{4.28}$$

and

$$Q_w(T, y, S) = \exp(-\gamma(I_{(S \leqq E)} c(y, S) + I_{(S > E)}[c(y-1, S) + E])). \tag{4.29}$$

Note that the function $Q_j(s, y, S)$ is evaluated in the three-dimensional space $[0, T] \times \mathbb{R} \times \mathbb{R}^+$.

If y defines the vertical axis, the optimal trading strategy when there are no transaction costs is defined by a surface denoted $\mathcal{I}(s, S)$; investors must trade in such a way that $y = \mathcal{I}(s, S)$ at all times. For V_1, this surface is given by

$$\mathcal{I}_1(s, S) = \frac{\delta(T, s)}{\gamma S} \frac{\alpha - r}{\sigma^2} \tag{4.30}$$

by using the results in Karatzas [13]. For V_w, this surface is given by

$$\mathcal{I}_w(s, S) = \frac{\partial \mathcal{C}(s, S)}{\partial S} + \frac{\delta(T, s)}{\gamma S} \frac{\alpha - r}{\sigma^2}, \tag{4.31}$$

where $\mathscr{C}(s, S)$ is the price of the European contingent claim in a market with no transaction costs and the partial derivative with respect to S is the replicating portfolio. In the case with transaction costs, it is conjectured that ∂B and ∂S are two surfaces, $\beta_b(s, S)$ and $\beta_s(s, S)$, which lie strictly below and above $\mathscr{I}(s, S)$, and the no-transaction region is between them (the numerical results, obtained in later sections, support this assertion). If the state of the investor (s, y, S) is in the no-transaction region, then it drifts, under the influence of the diffusion that decribes the stock price, on the plane defined by $y = \text{const}$. If the state is in the buy region or sell region, then the investor performs the minimum transaction required to take the state to $\beta_b(s, S)$ or $\beta_s(s, S)$ with an immediate purchase or sale of stock. As noted above, if the function $Q_j(s, y, S)$ is known in the no-transaction region, then (4.16), (4.18), and (4.22) are used to derive

$$(4.32) \qquad Q_j(s, y, S) = Q_j(s, y_b, S) \exp\left(-\frac{\gamma(1+\lambda)S}{\delta(T, s)}(y - y_b) \right) \quad \forall y \leqq y_b,$$

where $y_b \in \beta_b(s, S)$. A similar equation can be derived for $Q_j(s, y, S)$, for all $y \geqq \beta_s(s, S)$.

5. Existence and uniqueness of the solutions of the p.d.e. In this section, we characterise the value functions V_j given by (4.9) as weak (viscosity) solutions of the variational inequality (4.20). Since the stochastic control problems, whose value functions are given by (4.9), are associated with the same Hamilton–Jacobi–Bellman equation, we only examine the problem with value function V_1. We next show that this value function is a constrained viscosity solution of (4.20) on $[0, T] \times \bar{\mathscr{E}}_K$, where $\mathscr{E}_K$ is defined by (4.6); the characterisation of V_1 as a constrained viscosity solution of (4.20) is natural due to the presence of the state constraints (4.5).

The notion of *viscosity solutions* was introduced by Crandall and Lions [6] for first-order equations, and by Lions [17] for second-order equations. For a general overview of the theory, we refer to the *user's guide* by Crandall, Ishii, and Lions [7]. Next, we recall the notion of constrained viscosity solutions, which was introduced by Soner [18] and Capuzzo-Dolcetta and Lions [4] for first-order equations (see also Ishii and Lions [12] and Katsoulakis [14]). To this end, we consider a nonlinear second-order p.d.e. of the form

$$(5.1) \qquad F(X, W, DW, D^2 W) = 0 \quad \text{in } [0, T] \times \mathscr{E},$$

where $\mathscr{E} \subseteq \mathbb{R}^3$, DW, and $D^2 W$ denote the gradient vector and the second derivative of W, and the function F is continuous in all its arguments and degenerate elliptic, meaning that

$$(5.2) \qquad F(X, p, q, A+N) \leqq F(X, p, q, A) \quad \text{if } N \geqq 0.$$

DEFINITION 2. A continuous function W: $[0, T] \times \bar{\mathscr{E}} \to \mathbb{R}$ is a constrained viscosity solution of (5.1) if (i) W is a viscosity subsolution of (5.1) on $[0, T] \times \bar{\mathscr{E}}$; that is, if, for any $\phi \in C^{1,2}([0, T] \times \bar{\mathscr{E}})$ and any local maximum point $X_0 \in [0, T] \times \bar{\mathscr{E}}$ of $W - \phi$,

$$(5.3) \qquad F(X_0, W(X_0), D\phi(X_0), D^2\phi(X_0)) \leqq 0,$$

and (ii) W is a viscosity supersolution of (5.1) on $[0, T] \times \mathscr{E}$; that is, if, for any $\phi \in C^{1,2}([0, T] \times \bar{\mathscr{E}})$ and any local minimum point $X_0 \in [0, T] \times \mathscr{E}$ of $W - \phi$,

$$(5.4) \qquad F(X_0, W(X_0), D\phi(X_0), D^2\phi(X_0)) \geqq 0.$$

THEOREM 2. *The value function $V_1(s, B, y, S)$ is a constrained viscosity solution of*

$$\min\left\{-\left(\frac{\partial W}{\partial y}-(1+\lambda)S\frac{\partial W}{\partial B}\right), \frac{\partial W}{\partial y}-(1-\mu)S\frac{\partial W}{\partial B}, -\left(\frac{\partial W}{\partial s}+rB\frac{\partial W}{\partial B}+\alpha S\frac{\partial W}{\partial S}+\frac{1}{2}\sigma^2S^2\frac{\partial^2 W}{\partial S^2}\right)\right\}=0 \tag{5.5}$$

on $[0, T]\times\bar{\mathscr{E}}_K$.

Proof. In our case, the state X is (s, x), where $x=(B, y, S)\in\bar{\mathscr{E}}_K$. Also, (4.20) has been given the above form to turn it into an elliptic p.d.e., to which the uniqueness theorems are applicable. Let $X_0=(s_0, B_0, y_0, S_0)\in[0, T]\times\bar{\mathscr{E}}_K$: it follows, from results of Zhu [19, Thm. iv.2.2], that there exists an optimal trading strategy, dictated by the pair of processes $(\mathbf{L}^*(t), \mathbf{M}^*(t))$, where $\mathbf{X}_0^*(t)=(t, \mathbf{B}_0^*(t), \mathbf{y}_0^*(t), \mathbf{S}_0^*(t))$ is the optimal trajectory, with $\mathbf{X}_0^*(s_0)=X_0$.

(i) First, we prove that V_1 is a viscosity subsolution of (5.5) on $[0, T]\times\bar{\mathscr{E}}_K$; for this, we must show that, for all smooth functions $\phi(X)$, such that $V_1(X)-\phi(X)$ has a local maximum at $X_0\in[0, T]\times\bar{\mathscr{E}}_K$, the following inequality holds:

$$\min\left\{-\left(\frac{\partial\phi(X_0)}{\partial y}-(1+\lambda)S_0\frac{\partial\phi(X_0)}{\partial B}\right), \frac{\partial\phi(X_0)}{\partial y}-(1-\mu)S_0\frac{\partial\phi(X_0)}{\partial B}, -\left(\frac{\partial\phi(X_0)}{\partial s}+rB_0\frac{\partial\phi(X_0)}{\partial B}+\alpha S_0\frac{\partial\phi(X_0)}{\partial S}+\frac{1}{2}\sigma^2S_0^2\frac{\partial^2\phi(X_0)}{\partial S^2}\right)\right\}\leqq 0. \tag{5.6}$$

Without loss of generality, we assume that $V_1(X_0)=\phi(X_0)$ and $V_1\leqq\phi$ on $[0, T]\times\bar{\mathscr{E}}_K$. We argue by contradiction: if the arguments inside the minimum operator of (5.6) satisfy

$$\frac{\partial\phi(X_0)}{\partial y}-(1+\lambda)S_0\frac{\partial\phi(X_0)}{\partial B}<0, \tag{5.7}$$

$$\frac{\partial\phi(X_0)}{\partial y}-(1-\mu)S_0\frac{\partial\phi(X_0)}{\partial B}>0, \tag{5.8}$$

then there exists $\theta>0$, such that

$$\frac{\partial\phi(X_0)}{\partial s}+rB_0\frac{\partial\phi(X_0)}{\partial B}+\alpha S_0\frac{\partial\phi(X_0)}{\partial S}+\frac{1}{2}\sigma^2S_0^2\frac{\partial^2\phi(X_0)}{\partial S^2}<-\theta. \tag{5.9}$$

From the fact that ϕ is smooth, the above inequalities become

$$\frac{\partial\phi(X)}{\partial y}-(1+\lambda)S\frac{\partial\phi(X)}{\partial B}<0, \tag{5.10}$$

$$\frac{\partial\phi(X)}{\partial y}-(1-\mu)S\frac{\partial\phi(X)}{\partial B}>0, \tag{5.11}$$

and

$$\frac{\partial\phi(X)}{\partial s}+rB\frac{\partial\phi(X)}{\partial B}+\alpha S\frac{\partial\phi(X)}{\partial S}+\frac{1}{2}\sigma^2S^2\frac{\partial^2\phi(X)}{\partial S^2}<-\theta, \tag{5.12}$$

where $X=(s, B, y, S)\in\mathscr{B}(X_0)$, a neighborhood of X_0. In Lemma 1 it is shown that $\mathbf{X}_0^*(t)$ has no jumps, P-a.s., at $X_0=\mathbf{X}_0^*(s_0)$. Hence, $\tau(\omega)$, defined by

$$\tau(\omega)=\inf\{t\in[s_0, T]: \mathbf{X}_0^*(t)\notin\mathscr{B}(X_0)\}, \tag{5.13}$$

is positive P-a.s., and therefore

$$
\begin{aligned}
-\theta\mathbb{E}\{\tau\} > \mathbb{E}\int_{s_0}^{\tau}&\left(\frac{\partial\phi(\mathbf{X}_0^*(t))}{\partial y} - (1+\lambda)\mathbf{S}_0^*(t)\frac{\partial\phi(\mathbf{X}_0^*(t))}{\partial B}\right)d\mathbf{L}^*(t)\\
-\mathbb{E}\int_{s_0}^{\tau}&\left(\frac{\partial\phi(\mathbf{X}_0^*(t))}{\partial y} - (1-\mu)\mathbf{S}_0^*(t)\frac{\partial\phi(\mathbf{X}_0^*(t))}{\partial B}\right)d\mathbf{M}^*(t)\\
+\mathbb{E}\int_{s_0}^{\tau}&\left(\frac{\partial\phi(\mathbf{X}_0^*(t))}{\partial s} + r\mathbf{B}_0^*(t)\frac{\partial\phi(\mathbf{X}_0^*(t))}{\partial B}\right.\\
&\left.+\alpha\mathbf{S}_0^*(t)\frac{\partial\phi(\mathbf{X}_0^*(t))}{\partial S} + \frac{1}{2}\sigma^2(\mathbf{S}_0^*(t))^2\frac{\partial^2\phi(\mathbf{X}_0^*(t))}{\partial S^2}\right)dt\\
= \mathbb{E}\{I_1\} - \mathbb{E}&\{I_2\} + \mathbb{E}\{I_3\},
\end{aligned}
\tag{5.14}
$$

where $(\mathbf{L}^*(t), \mathbf{M}^*(t))$ is the optimal trading strategy at X_0. Applying Itô's formula to $\phi(X)$, where the state dynamics are given by (4.2)–(4.4), we get

$$
\mathbb{E}\{\phi(\mathbf{X}_0^*(\tau))\} = \phi(X_0) + \mathbb{E}\{I_1\} - \mathbb{E}\{I_2\} + \mathbb{E}\{I_3\}. \tag{5.15}
$$

Since $V(X) \leqq \phi(X)$, for all $X \in \mathscr{B}(X_0)$, and $V_1(X_0) = \phi(X_0)$, (5.14) and (5.15) yield

$$
\mathbb{E}\{V_1(\mathbf{X}_0^*(\tau))\} \leqq V_1(X_0) + (\mathbb{E}\{I_1\} - \mathbb{E}\{I_2\} + \mathbb{E}\{I_3\}) < V_1(X_0) - \theta\mathbb{E}\{\tau\}, \tag{5.16}
$$

which violates the dynamic programming principle, together with the optimality of $(\mathbf{L}^*(t), \mathbf{M}^*(t))$. Therefore, at least one of the arguments inside the minimum operator of (5.6) is nonpositive, and hence the value function is a viscosity subsolution of (5.5).

(ii) In the second part of the proof, we show that V_1 is a viscosity supersolution of (5.5) in $[0, T] \times \mathscr{E}_K$; for this we must show that, for all smooth functions $\phi(X)$, such that $V_1(X) - \phi(X)$ has a local minimum at $X_0 \in [0, T] \times \mathscr{E}_K$, the following inequality holds:

$$
\begin{aligned}
\min\Bigg\{&-\left(\frac{\partial\phi(X_0)}{\partial y} - (1+\lambda)S_0\frac{\partial\phi(X_0)}{\partial B}\right), \frac{\partial\phi(X_0)}{\partial y} - (1-\mu)S_0\frac{\partial\phi(X_0)}{\partial B},\\
&-\left(\frac{\partial\phi(X_0)}{\partial s} + rB_0\frac{\partial\phi(X_0)}{\partial B} + \alpha S_0\frac{\partial\phi(X_0)}{\partial S} + \frac{1}{2}\sigma^2 S_0^2\frac{\partial^2\phi(X_0)}{\partial S^2}\right)\Bigg\} \geqq 0,
\end{aligned}
\tag{5.17}
$$

where, without loss of generality, $V_1(X_0) = \phi(X_0)$ and $V_1 \geqq \phi$ on $[0, T] \times \bar{\mathscr{E}}_K$. In this case, we prove that each argument of the minimum operator of (5.17) is nonnegative.

Consider the trading strategy $\mathbf{L}(t) = L_0 > 0$, $s_0 \leqq t \leqq T$, and $\mathbf{M}(t) = 0$, $s_0 \leqq t \leqq T$. By the dynamic programming principle,

$$
V_1(s_0, B_0, y_0, S_0) \geqq V_1(s_0, B_0 - (1+\lambda)S_0L_0, y_0 + L_0, S_0). \tag{5.18}
$$

This inequality holds for $\phi(s, B, y, S)$ as well, and, by taking the left-hand side to the right-hand side, dividing by L_0, and sending $L_0 \to 0$, we get

$$
\frac{\partial\phi(X_0)}{\partial y} - (1+\lambda)S_0\frac{\partial\phi(X_0)}{\partial B} \leqq 0. \tag{5.19}
$$

Similarly, by using the trading strategy $\mathbf{L}(t) = 0$, $s_0 \leqq t \leqq T$, and $\mathbf{M}(t) = M_0 > 0$, $s_0 \leqq t \leqq T$, the second argument inside the minimum operator is found to be nonnegative.

Finally, consider the case where no trading is applied. By the dynamic programming principle

$$
\mathbb{E}\{V_1(\mathbf{X}_0^d(t))\} \leqq V_1(s_0, B_0, y_0, S_0), \tag{5.20}
$$

where $\mathbf{X}_0^d(t)$ is the state trajectory when $\mathbf{M}(t)=\mathbf{L}(t)=0$, $s_0 \leqq t \leqq T$, given by (4.2)-(4.4) as

$$\mathbf{X}_0^d(t)=(t, B_0 \exp(r(t-s_0)), y_0, S_0 \exp((\alpha-\tfrac{1}{2}\sigma^2)(t-s_0)+\sigma(\mathbf{R}(t)-\mathbf{R}(s_0)))) \tag{5.21}$$

and $\mathbf{X}_0^d(t) \in \mathscr{B}(X_0)$. Therefore, by applying Itô's rule on $\phi(s, B, y, S)$, inequality (5.20) yields

$$\begin{aligned} \mathbb{E}\bigg\{\int_{s_0}^{t} \bigg(&\frac{\partial \phi(\mathbf{X}_0^d(\xi))}{\partial s}+r\mathbf{B}_0^d(\xi)\frac{\partial \phi(\mathbf{X}_0^d(\xi))}{\partial B}+\alpha \mathbf{S}_0^d(\xi)\frac{\partial \phi(\mathbf{X}_0^d(\xi))}{\partial S} \\ &+\frac{1}{2}\sigma^2(\mathbf{S}_0^d(\xi))^2\frac{\partial^2 \phi(\mathbf{X}_0^d(\xi))}{\partial S^2}\bigg)\, d\xi\bigg\} \geqq 0, \end{aligned} \tag{5.22}$$

and, by sending $t \to s_0$, the third argument inside the minimum operator is found to be nonnegative (for detailed proof, see Lions [17]). This completes the proof. □

LEMMA 1. *Assume that inequality* (5.7) *holds and denote the event that the optimal trajectory* $\mathbf{X}_0^*(t)$ *has a jump of size* ε, *along the direction* $(0, -(1+\lambda)S_0, 1, 0)$, *by* $A(\omega)$. *Assume that the state* (*after the jump*) *is* $(s_0, B_0-(1+\lambda)S_0\varepsilon, y_0+\varepsilon, S_0) \in \mathscr{B}(X_0)$. *Then*

$$\left(\frac{\partial \phi(X_0)}{\partial y}-(1+\lambda)S\frac{\partial \phi(X_0)}{\partial B}\right)P(A) \geqq 0, \tag{5.23}$$

and therefore $P(A)=0$. *Similarly, if inequality* (5.8) *holds, then the optimal trajectory has no jumps along the direction* $(0, (1-\mu)S_0, -1, 0)$, *P-a.s. at* X_0.

Proof. By the principle of dynamic programming,

$$\begin{aligned} V_1(s_0, B_0, y_0, S_0) &= \mathbb{E}\{V_1(s_0, B_0-(1+\lambda)S_0\varepsilon, y_0+\varepsilon, S_0)\} \\ &= \int_{A(\omega)} V_1(s_0, B_0-(1+\lambda)S_0\varepsilon, y_0+\varepsilon, S_0)\, dP \\ &\quad +\int_{\Omega-A(\omega)} V_1(s_0, B_0, y_0, S_0)\, dP, \end{aligned} \tag{5.24}$$

and therefore

$$\int_{A(\omega)} (\phi(s_0, B_0-(1+\lambda)S_0\varepsilon, y_0+\varepsilon, S_0)-\phi(s_0, B_0, y_0))\, dP \geqq 0, \tag{5.25}$$

since $V_1(X) \leqq \phi(X)$ for all $X \in \mathscr{B}(X_0)$ and $V_1(X_0)=\phi(X_0)$. Therefore,

$$\limsup_{\varepsilon \to 0}\left\{\int_{A(\omega)} \frac{\phi(s_0, B_0-(1+\lambda)S_0\varepsilon, y_0+\varepsilon, S_0)-\phi(s_0, B_0, y_0, S_0)}{\varepsilon}\, dP\right\} \geqq 0, \tag{5.26}$$

and, by Fatou's lemma,

$$\int_{A(\omega)} \limsup_{\varepsilon \to 0}\left\{\frac{\phi(s_0, B_0-(1+\lambda)S_0\varepsilon, y_0+\varepsilon, S_0)-\phi(s_0, B_0, y_0, S_0)}{\varepsilon}\right\} dP \geqq 0, \tag{5.27}$$

which implies (5.23). □

This section is concluded by showing that the value function V_1 is the unique bounded constrained viscosity solution of (5.5). Since this uniqueness result will be used mainly for the convergence of the numerical scheme presented in the next section, we prove the theorem for the exponential utility function. For simplicity of exposition, we assume that the interest rate $r=0$. The argument is the same but notationally more cumbersome when $r>0$.

THEOREM 3. *Let u be a bounded upper semicontinuous viscosity subsolution of* (5.5) *on* $[0, T]\times\bar{\mathscr{E}}_K$, *and let v be a bounded from below lower semicontinuous viscosity supersolution of* (5.5) *in* $[0, T]\times\mathscr{E}_K$, *such that* $u(T, x)\leqq v(T, x)$, *for all* $x\in\bar{\mathscr{E}}_K$, *and* $u(t, B, y, 0)\leqq v(t, B, y, 0)$, *on* $[0, T]\times\bar{\mathscr{E}}_K$, *where* $u(T, x)=1-\exp(-\gamma(B+c(y, S)))$ *and* $u(t, B, y, 0)=1-\exp(-\gamma B)$. *Then* $u\leqq v$ *on* $[0, T]\times\bar{\mathscr{E}}_K$.

Note. The proof relies on arguments used in § v of Ishii and Lions [12]; only the main steps are presented.

Proof. Sketch. We first construct a positive strict supersolution of (5.5) in $[0, T]\times\mathscr{E}_K$. To this end, let $h: [0, T]\times\mathscr{E}_K\to\mathbb{R}^+$ be given by $h(t, B, y, S)=1-\exp(-\gamma(B+kyS))+C_1(T-t)+C_2$, where the constants k, C_1, and C_2 satisfy

$$1+\lambda>k>1-\mu,\quad C_1>\frac{\alpha^2}{2\sigma^2}\exp(\gamma K),\quad \text{and } C_2>\exp(K)-1. \tag{5.28}$$

Then

$$\begin{aligned} &H(X, h_t, Dh, D^2h)\\ &\quad=\min\left\{-\frac{\partial h}{\partial y}+(1+\lambda)S\frac{\partial h}{\partial B}, \frac{\partial h}{\partial y}-(1-\mu)S\frac{\partial h}{\partial B}, -\frac{\partial h}{\partial s}-\alpha S\frac{\partial h}{\partial S}-\frac{1}{2}\sigma^2S^2\frac{\partial^2 h}{\partial S^2}\right\}\\ &\quad=\exp(-\gamma(B+kyS))\\ &\qquad\cdot\min\Big\{\gamma S(1+\lambda-k), \gamma S(k-(1-\mu)),\\ &\qquad\qquad C_1\exp(\gamma(B+kyS))+\frac{1}{2}\gamma^2k^2\sigma^2(yS)^2-\alpha k\gamma(yS)\Big\}. \end{aligned} \tag{5.29}$$

Using (5.28) and the fact that the minimum value of the quadratic $\mathscr{Q}(\zeta)=\frac{1}{2}\gamma^2k^2\sigma^2\zeta^2-\alpha\gamma k\zeta$ is $-\alpha^2/2\sigma^2$, the above inequality yields

$$\begin{aligned} &H(X, h_t, Dh, D^2h)\\ &\quad>\exp(-\gamma(B+kyS))\min\{\gamma S(1+\lambda-k), \gamma S(k-(1-\mu)), K'\} \end{aligned} \tag{5.30}$$

in $[0, T]\times\bar{\mathscr{E}}_K$, where $0<K'<C_1-(\alpha^2/2\sigma^2)\exp(\gamma K)$.

Therefore, h is a strict supersolution (5.5). The fact that $h>0$ follows from the choice of the constant C_2.

To conclude the proof of the theorem, we will need the following key lemma. Its proof follows along the lines of Theorem vi.5 in Ishii and Lions [12], and therefore it is omitted.

LEMMA 2. *Let u be a bounded lower semicontinuous viscosity subsolution of* (5.5) *on* $[0, T)\times\bar{\mathscr{E}}_K$, *and let v be a bounded from below uniformly continuous viscosity supersolution of* (5.5) *in* $[0, T)\times\mathscr{E}_K$ *of the equation* $H(X, v_t, Dv, D^2v)-f(x)=0$, *where* $f>0$ *in* $\mathscr{E}_K$, $u(T, x)\leqq v(T, x)$, *for all* $x\in\bar{\mathscr{E}}_K$, *and* $u(t, B, y, 0)\leqq v(t, B, y, 0)$, *on* $[0, T]\times\bar{\mathscr{E}}_K$. *Then* $u\leqq v$ *on* $[0, T]\times\bar{\mathscr{E}}_K$.

We now conclude the proof of the theorem. To this end, we first observe that, because of the choice of k, C_1 and C_2,

$$h(T, B, y, S)>1-\exp(-\gamma(B+c(y, S)))\quad\text{and}\quad h(t, B, y, 0)>1-\exp(-\gamma B). \tag{5.31}$$

Next, we define the function $w^\theta=\theta v+(1-\theta)h$, where $0<\theta<1$, and, using (5.31), we get

$$w^\theta(T, B, y, S)\geqq u(T, B, y, S)\quad\text{and}\quad w^\theta(t, B, y, 0)\geqq u(t, B, y, 0). \tag{5.32}$$

We also observe that w^θ is a viscosity supersolution of $H-g=0$, in $[0, T]\times\bar{\mathscr{E}}_K$, where $g=(1-\theta)f$. In fact, let $\psi\in C^{1,2}([0,T]\times\mathscr{E}_K)$ and assume that $w^\theta-\psi$ has a minimum at X_0. Then $v-\phi$ also has a minimum at X_0, where $\phi=(1/\theta)\times(\psi-(1-\theta)h)$. Using the fact that v is a viscosity supersolution of $H=0$ and (5.30) we get

$$\begin{aligned}&\theta H(X_0, \phi_t(X_0), D\phi(X_0), D^2\phi(X_0))\\ &\quad+(1-\theta)H(X_0, h_t(X_0), Dh(X_0), D^2h(X_0))\geqq(1-\theta)f(X_0).\end{aligned} \tag{5.33}$$

As the Hamiltonian $H(x, p, q, A)$ is jointly concave with respect to (p, q, A), (5.33) yields

$$H(X_0, \psi_t(X_0), D\psi(X_0), D^2\psi(X_0))\geqq(1-\theta)f(X_0), \tag{5.34}$$

which in turn implies that w^θ is a viscosity supersolution of $H-g=0$. Finally, applying Lemma 2 to u and w^θ, we get

$$u\leqq w^\theta \quad \text{on } [0, T]\times\bar{\mathscr{E}}_K, \tag{5.35}$$

and sending $\theta\uparrow 1$ concludes the proof of the theorem. □

6. Discretisation and solution of the problem. The solution of the p.d.e. (4.20) is obtained by turning the stochastic differential equations (4.2)-(4.4) into Markov chains to apply the discrete time dynamic programming algorithm. The method here closely follows the one given in Martins and Kushner [15]. The discrete state is $\mathbb{X}=(\iota, \mathbb{B}, \eta, \mathbb{S})$, whose elements denote time, amount in the bank, number of shares, and stock price in a discrete space. The value functions, denoted by $\mathbb{V}_1$ and $\mathbb{V}_w$, are given a value at the final time by using the boundary conditions for the continuous value functions over the discrete subspace $(\mathbb{B}, \eta, \mathbb{S})$, and then they are estimated by proceeding backward in time by using the discrete time algorithm. As in the continuous time case, this algorithm is the same for both value functions and is derived below for a value function denoted by $\mathbb{V}_j^\rho(\iota, \mathbb{B}, \eta, \mathbb{S})$, where ρ is a discretisation parameter, which depends on the discrete time interval δt. If δt and the resolution of the η axis $\delta\eta$ are sent to zero, then the above discrete value function converges to a viscosity subsolution and a viscosity supersolution of the p.d.e. (4.20). Therefore, all the discrete value functions converge to their continuous counterparts; this is due to the uniqueness of the viscosity solution.

The discrete time variable ι takes values in $\{0, \delta t, 2\delta t, \dots, N\delta t\}$, where δt is the discrete time interval and $T-s=N\delta t$. The Markov chain for the discrete stock price process $\mathbf{S}(\iota)$ is modeled by

$$\mathbf{S}(\iota+1)=\begin{cases}\mathbf{S}(\iota)\times k_u & \text{with probability } \frac{1}{2},\\ \mathbf{S}(\iota)\times k_d & \text{with probability } \frac{1}{2},\end{cases} \tag{6.1}$$

where the values of k_u and k_d are determined by equating the first and second moments of the chain with those of the diffusion, which describes S, and therefore

$$k_u=\exp(\alpha\,\delta t+\sigma\sqrt{\delta t}) \quad \text{and} \quad k_d=\exp(\alpha\,\delta t-\sigma\sqrt{\delta t}). \tag{6.2}$$

The discretisation scheme and its convergence properties are more thoroughly explained in Chapter 5 of Cox and Rubinstein [5]. The discrete time equation for the amount in the bank $\mathbb{B}(\iota)$ is

$$\mathbb{B}(\iota+1)=\mathbb{B}(\iota)\exp(r\,\delta t), \tag{6.3}$$

which is a deterministic difference equation.

The discrete time dynamic programming principle is invoked, and the following discretisation scheme is proposed for the p.d.e. (4.20):

$$\mathscr{S}(\rho)\mathbb{V}_j^\rho-\mathbb{V}_j^\rho=0, \tag{6.4}$$

where $\mathcal{S}(\rho)$ is an operator given by

$$\begin{aligned}\mathcal{S}(\rho)\mathbb{V}_j^\rho = \max\{&\mathbb{V}_j^\rho(\iota, \mathbb{B}-(1+\lambda)\mathbb{S}\kappa\rho, \eta+\kappa\rho, \mathbb{S}), \mathbb{V}_j^\rho(\iota, \mathbb{B}+(1-\mu)\mathbb{S}\kappa\rho, \eta-\kappa\rho, \mathbb{S}),\\ &\mathbb{E}\{\mathbb{V}_j^\rho(\iota+\rho, \mathbb{B}\exp(r\rho), \eta, \mathbb{S}\exp(\alpha\rho+\theta\sigma\sqrt{\rho}))\}\},\end{aligned} \tag{6.5}$$

where $\rho=\delta t$, κ is a real constant and θ is a random variable taking values ± 1 with probability $\frac{1}{2}$ each. This scheme is based on the principle that the investor's policy is the choice of the optimum transaction, that is, to buy or sell or do nothing for a particular state given the value function for all the states in the next time instant. We next show that, as the discretisation parameter $\rho\to 0$, the solution $\mathbb{V}_j^\rho$ of (6.4) converges to the value function V, or, equivalently, to the unique constrained viscosity solution of (4.20). Although the proof of the theorem follows along the lines of Barles and Souganidis [Thm. 2.1, p. 1], it is presented here for completeness.

THEOREM 4. *The solution $\mathbb{V}_j^\rho$ of* (6.4) *converges locally uniformly as $\rho\to 0$ to the unique continuous constrained viscosity solution of* (4.20).

Proof. Let

$$V_j^\rho(t, B, y, S) = \begin{cases}\mathbb{V}_j^\rho(\iota, \mathbb{S}, \eta, \mathbb{B}), & \text{if } t\in[\iota, \iota+\rho), y\in[\eta, \eta+\kappa\rho),\\ \Phi_j(\mathbb{B}, \eta, \mathbb{S}), & \text{if } t=T\end{cases}, \tag{6.6}$$

and

$$\underline{V}_j(X) = \liminf_{Y\to X, \rho\to 0}\{\mathbb{V}_j^\rho(Y)\} \quad\text{and}\quad \bar{V}_j(X) = \limsup_{Y\to X, \rho\to 0}\{\mathbb{V}_j^\rho(Y)\}, \tag{6.7}$$

where $X=(t, B, y, S)$. We will show that $\underline{V}_j$ and $\bar{V}_j$ are a viscosity supersolution and a viscosity subsolution of (4.20), respectively. Combining this with the uniqueness result of Theorem 3 yields $\underline{V}_j \geqq \bar{V}_j$ on $[0, T]\times\mathscr{E}_K$. The opposite inequality is true by the definition of $\underline{V}_j$ and $\bar{V}_j$, and therefore,

$$\underline{V}_j(X) = \bar{V}_j(X) = V_j(X), \tag{6.8}$$

which, together with (6.7), also implies the local uniform convergence of $\mathbb{V}_j^\rho$ to V_j (see [1]).

We will only prove that $\underline{V}_j$ is a viscosity supersolution of (4.20), since the arguments for $\bar{V}_j$ are identical. Let X_0 be a local minimum of $\underline{V}_j-\phi$ on $[0, T]\times\mathscr{E}_K$, for $\phi\in C^{1,2}([0, T]\times\mathscr{E}_K)$.Without loss of generality, we may assume that X_0 is a strict local minimum, that $\underline{V}_j(X_0)=\phi(X_0)$, and that $\phi\leqq -2\times\sup_\rho\{\|\mathbb{V}_j^\rho\|_\infty\}$ outside the ball $\mathscr{B}(X_0, R)$, $R>0$, where $\underline{V}_j(X)-\phi(X)\geqq 0$.

Then there exist sequences $\rho_n\in\mathbb{R}^+$ and $Y_n\in[0, T]\times\mathscr{E}_K$, such that

$$\rho_n\to 0,\ Y_n\to X_0, \mathbb{V}_j^{\rho_n}(Y_n)\to \underline{V}_j(X_0),\ Y_n \quad\text{is a global minimum point of } \mathbb{V}_j^{\rho_n}-\phi. \tag{6.9}$$

Let $h_n=\mathbb{V}_j^{\rho_n}-\phi$; then

$$h_n\to 0 \quad\text{and}\quad \mathbb{V}_j^{\rho_n}(X)\geqq \phi(X)+h_n(X) \qquad \forall X\in[0, T]\times\mathscr{E}_K. \tag{6.10}$$

To show that $\underline{V}_j$ is a viscosity supersolution of (4.20), it suffices to show that

$$\begin{aligned}\min\Bigg\{&-\left(\frac{\partial\phi(X_0)}{\partial y}-(1+\lambda)S_0\frac{\partial\phi(X_0)}{\partial B}\right), \left(\frac{\partial\phi(X_0)}{\partial y}-(1-\mu)S_0\frac{\partial\phi(X_0)}{\partial B}\right),\\ &-\left(\frac{\partial\phi(X_0)}{\partial s}+rB_0\frac{\partial\phi(X_0)}{\partial B}+\alpha S_0\frac{\partial\phi(X_0)}{\partial S}+\frac{1}{2}\sigma^2S_0^2\frac{\partial^2\phi(X_0)}{\partial S^2}\right)\Bigg\}\geqq 0.\end{aligned} \tag{6.11}$$

Let $Y_n = (t_{\rho_n}, \mathbb{B}_{\rho_n}, y_{\rho_n}, \mathbb{S}_{\rho_n})$, where $t_{\rho_n} \in [\iota_{\rho_n}, \iota_{\rho_n} + \rho_n)$ and $y_{\rho_n} \in [\eta_{\rho_n}, \eta_{\rho_n} + \kappa\rho_n)$. Then

$$\begin{aligned} &\mathbb{V}_j^{\rho_n}(\iota_{\rho_n}, \mathbb{B}_{\rho_n}, y_{\rho_n}, \mathbb{S}_{\rho_n}) \\ &\quad = \max \{\mathbb{V}_j^{\rho_n}(\iota_{\rho_n}, \mathbb{B}_{\rho_n} - (1+\lambda)\mathbb{S}_{\rho_n}\kappa\rho_n, \eta_{\rho_n} + \kappa\rho_n, \mathbb{S}_{\rho_n}), \\ &\qquad \mathbb{V}_j^{\rho_n}(\iota_{\rho_n}, \mathbb{B}_{\rho_n} + (1-\mu)\mathbb{S}_{\rho_n}\kappa\rho_n, \eta_{\rho_n} - \kappa\rho_n, \mathbb{S}_{\rho_n}), \\ &\qquad \mathbb{E}\{\mathbb{V}_j^{\rho_n}(\iota_{\rho_n} + \rho_n, \mathbb{B}_{\rho_n} \exp(r\rho_n), \eta_{\rho_n}, \mathbb{S}_{\rho_n} \exp(\alpha\rho_n + \theta\sigma\sqrt{\rho_n}))\}\}. \end{aligned} \tag{6.12}$$

Now we look at the following three cases.

Case 1. It holds that

$$\mathbb{V}_j^{\rho_n}(\iota_{\rho_n}, \mathbb{B}_{\rho_n}, \eta_{\rho_n}, \mathbb{S}_{\rho_n}) = \mathbb{V}_j^{\rho_n}(\iota_{\rho_n}, \mathbb{B}_{\rho_n} - (1+\lambda)\mathbb{S}_{\rho_n}\kappa\rho_n, \eta_{\rho_n} + \kappa\rho_n, \mathbb{S}_{\rho_n}).$$

Then (6.10) implies that

$$\begin{aligned} \mathbb{V}_j^{\rho_n}(\iota_{\rho_n}, \mathbb{B}_{\rho_n}, \eta_{\rho_n}, \mathbb{S}_{\rho_n}) &\geqq \phi(\iota_{\rho_n}, \mathbb{B}_{\rho_n} - (1+\lambda)\mathbb{S}_{\rho_n}\kappa\rho_n, \eta_{\rho_n} + \kappa\rho_n, \mathbb{S}_{\rho_n}) \\ &\quad + \mathbb{V}_j^{\rho_n}(\iota_{\rho_n}, \mathbb{B}_{\rho_n}, \eta_{\rho_n}, \mathbb{S}_{\rho_n}) - \phi(\iota_{\rho_n}, \mathbb{B}_{\rho_n}, \eta_{\rho_n}, \mathbb{S}_{\rho_n}), \end{aligned} \tag{6.13}$$

and therefore

$$\begin{aligned} 0 &\geqq \liminf_n \left\{ \frac{\phi(\iota_{\rho_n}, \mathbb{B}_{\rho_n} - (1+\lambda)\mathbb{S}_{\rho_n}\kappa\rho_n, \eta_{\rho_n} + \kappa\rho_n, \mathbb{S}_{\rho_n}) - \phi(\iota_{\rho_n}, \mathbb{B}_{\rho_n}, \eta_{\rho_n}, \mathbb{S}_{\rho_n})}{\rho_n} \right\} \\ &\geqq \liminf_{\rho \to 0} \left\{ \frac{\phi(t_0, B_0 - (1+\lambda)S_0\kappa\rho, y_0 + \kappa\rho, S_0) - \phi(t_0, B_0, y_0, S_0)}{\rho} \right\} \\ &= \frac{\partial\phi(X_0)}{\partial y} - (1+\lambda)S_0 \frac{\partial\phi(X_0)}{\partial B}. \end{aligned} \tag{6.14}$$

Case 2. It holds that

$$\mathbb{V}_j^{\rho_n}(\iota_{\rho_n}, \mathbb{B}_{\rho_n}, \eta_{\rho_n}, \mathbb{S}_{\rho_n}) = \mathbb{V}_j^{\rho_n}(\iota_{\rho_n}, \mathbb{B}_{\rho_n} + (1-\mu)\mathbb{S}_{\rho_n}\kappa\rho_n, \eta_{\rho_n} - \kappa\rho_n, \mathbb{S}_{\rho_n}).$$

Working similarly to the above case, we get

$$0 \geqq -\left(\frac{\partial\phi(X_0)}{\partial y} - (1-\mu)S_0 \frac{\partial\phi(X_0)}{\partial B} \right). \tag{6.15}$$

Case 3. It holds that

$$\mathbb{V}_j^{\rho_n}(\iota_{\rho_n}, \mathbb{B}_{\rho_n}, \eta_{\rho_n}, \mathbb{S}_{\rho_n}) = \mathbb{E}\{\mathbb{V}_j^{\rho_n}(\iota_{\rho_n} + \rho_n, \mathbb{B}_{\rho_n}, \exp(r\rho_n), \eta_{\rho_n}, \mathbb{S}_{\rho_n} \exp(\alpha\rho_n + \theta\sigma\sqrt{\rho_n}))\}.$$

Then (6.10) implies that

$$\begin{aligned} &\mathbb{V}_j^{\rho_n}(\iota_{\rho_n}, \mathbb{B}_{\rho_n}, \eta_{\rho_n}, \mathbb{S}_{\rho_n}) \\ &\quad \geqq \mathbb{E}\{\phi(\iota_{\rho_n} + \rho_n, \mathbb{B}_{\rho_n} \exp(r\rho_n), \eta_{\rho_n}, \mathbb{S}_{\rho_n} \exp(\alpha\rho_n + \theta\sigma\sqrt{\rho_n}))\} \\ &\qquad + \mathbb{V}_j^{\rho_n}(\iota_{\rho_n}, \mathbb{B}_{\rho_n}, \eta_{\rho_n}, \mathbb{S}_{\rho_n}) - \phi(\iota_{\rho_n}, \mathbb{B}_{\rho_n}, \eta_{\rho_n}, \mathbb{S}_{\rho_n}), \end{aligned} \tag{6.16}$$

and therefore

$$\begin{aligned} 0 &\geqq \liminf_n \left\{ \frac{\phi(\iota_{\rho_n} + \rho_n, \mathbb{B}_{\rho_n} \exp(r\rho_n), \eta_{\rho_n}, \mathbb{S}_{\rho_n} \exp(\alpha\rho_n + \theta\sigma\sqrt{\rho_n})) - \phi(\iota_{\rho_n}, \mathbb{B}_{\rho_n}, \eta_{\rho_n}, \mathbb{S}_{\rho_n})}{\rho_n} \right\} \\ &\geqq \liminf_{\rho \to 0} \left\{ \frac{\phi(t_0 + \rho, B_0 \exp(r\rho), y_0, S_0 \exp(\alpha\rho + \theta\sigma\sqrt{\rho})) - \phi(t_0, B_0, y_0, S_0)}{\rho} \right\} \\ &= \frac{\partial\phi(X_0)}{\partial s} + rB_0 \frac{\partial\phi(X_0)}{\partial B} + \alpha S_0 \frac{\partial\phi(X_0)}{\partial S} + \frac{1}{2}\sigma^2 S_0^2 \frac{\partial^2\phi(X_0)}{\partial S^2}. \end{aligned} \tag{6.17}$$

Combining (6.14), (6.15), and (6.17) yields (6.11), and the proof is complete. □

In the discrete time framework, the exponential utility function, given by (4.21), has the same effect on the value function as before, and therefore (4.25) can be written as follows:

$$\mathbb{V}_j(\iota, \mathbb{B}, \eta, \mathbb{S}) = 1 - \exp\left(-\gamma \frac{\mathbb{B}}{\Delta(N, \iota)}\right)\mathbb{Q}_j(\iota, \eta, \mathbb{S}), \tag{6.18}$$

where $\Delta(\nu, \nu')$ is the discrete time discount factor, given by

$$\Delta(\nu, \nu') = \exp\left(-r(\nu - \nu')\right), \tag{6.19}$$

where ν and ν' take values in the same set with ι and $\nu \geqq \nu'$. The discretisation scheme for the new functions $\mathbb{Q}_j(\iota, \eta, \mathbb{S}) := 1 - \mathbb{V}_j(\iota, 0, \eta, \mathbb{S})$ is derived from (6.4) and (6.18) to be

$$\begin{aligned}\mathbb{Q}_j(\iota, \eta, \mathbb{S}) = \min\{&\mathbb{F}_b(\iota, \zeta, \mathbb{S}) \times \mathbb{Q}_j(\iota, \eta + \zeta, \mathbb{S}),\\ &\mathbb{F}_s(\iota, \zeta, \mathbb{S}) \times \mathbb{Q}_j(\iota, \eta - \zeta, \mathbb{S}), \mathbb{E}\{\mathbb{Q}_j(i + \zeta, \eta, \Theta \times \mathbb{S})\}\},\end{aligned} \tag{6.20}$$

where

$$\mathbb{F}_b(\iota, \zeta, \mathbb{S}) = \exp\left(\gamma \frac{(1+\lambda)\mathbb{S}\zeta}{\Delta(N, \iota)}\right) \tag{6.21}$$

and

$$\mathbb{F}_s(\iota, \zeta, \mathbb{S}) = \exp\left(-\gamma \frac{(1-\mu)\mathbb{S}\zeta}{\Delta(N, \iota)}\right), \tag{6.22}$$

and the boundary conditions at $\iota = N$ are given by the discrete space versions of (4.28) and (4.29). As in the continuous time case, if the value functions are known in the no-transaction region, then they can be calculated in the buy and sell regions by using the discrete space versions of (4.15) and (4.17): Suppose that η_b^* is the value of η, at which it is optimal to buy ζ shares, whereas, at $\eta = \eta_b^* + \zeta$ it is optimal to perform no transaction at all; then the function $\mathbb{Q}_j(\iota, \eta, \mathbb{S})$ is determined by

$$\mathbb{Q}_j(\iota, \eta, \mathbb{S}) = \mathbb{F}_b(\iota, \eta_b^* - \eta, \mathbb{S}) \times \mathbb{Q}_j(\iota, \eta_b^*, \mathbb{S}) \quad \forall \eta < \eta_b^*, \tag{6.23}$$

and a similar equation can be derived for $\eta > \eta_s^*$, the value of η at which it is optimal to sell ζ shares by using $\mathbb{F}_s(\iota, \eta - \eta_s^*, \mathbb{S})$. Finally, the price of the European contingent claim is given by

$$\mathbb{P}_w(\iota, \mathbb{S}) = \frac{\Delta(N, t)}{\gamma} \ln\left(\frac{\mathbb{Q}_w(\iota, 0, \mathbb{S})}{\mathbb{Q}_1(\iota, 0, \mathbb{S})}\right), \tag{6.24}$$

which is the discrete time version of (4.26).

7. Numerical results. The algorithm developed in the previous section was implemented, and the writing price of a European call option was calculated for a writer with exponential utility function given by (4.21) and for constant model coefficients. For comparison, the Black–Scholes value was also calculated from

$$p_{bs}(s, S) = SN(\chi) - Ee^{-r(T-s)}N(\chi - \sigma\sqrt{T-s}), \tag{7.1}$$

where

$$\chi = \frac{\ln(S/Ee^{-r(T-s)})}{\sigma\sqrt{T-s}} + \frac{1}{2}\sigma\sqrt{T-s}, \tag{7.2}$$

and $N(\cdot)$ denotes the normal distribution function with mean zero and variance 1. The number of stock units held in the hedging portfolio is simply $N(\chi)$. The boundary conditions for the value functions were set according to the analysis presented in § 1, and several values of proportional transaction costs were tried. The effects on p_w of the model parameters, the time to expiration, and the stock price at the time the option is written were in line with expectations and are outlined below. The parameters λ and μ were set equal to each other and are denoted by T.C. in the figures.

For both value functions, the boundaries ∂B and ∂S were found to lie below and above the optimal trading strategy without transaction costs, and the no-transaction region was observed to widen as the expiration date approached. (Note that we have not proved that the optimal transaction policies consist of reflection off these boundaries, although, of course, we believe this to be the case.) This shows that the investor is reluctant to transact toward the end of the trading interval, as he thinks that the stock price may not vary too much until the final time. (The cost of transactions is likely to reduce the utility of the final wealth more than the cost of providing one share at the final time.) Also, the boundary ∂B, for the value function $V_w(s, B, y, S)$, was virtually equal to the Black-Scholes trading strategy $N(\chi)$, indicating that the writer considers the cost of the obligation to provide one share of stock at the final time (if he does not already own it) as the most significant factor, affecting the trading strategy for this value function.

The most important result appears in Fig. 1, where the price difference $p_w - p_{b_s}$ is plotted against time over a three-year period, with the parameter values for a one-year period shown in its title (p_{b_s} is the Black-Scholes price, given by (7.1)). The expiration date is at the end of the third year, where both prices vanish, as the option is worthless. The price difference a long time before the expiration of the claim is equal to λS, which is the amount required to buy one share of the stock. The reason for this is revealed by observing $N(\chi)$ over the three time periods, plotted in Fig. 2 for the same parameter values. Although $S < E$, $N(\chi)$ increases as the time to expiration increases and dictates that the investor must own almost one share of the stock in a market with no transaction costs if the time to expiration is long enough. By that time, the extra price that the writer charges is the amount required to buy one share of the stock,

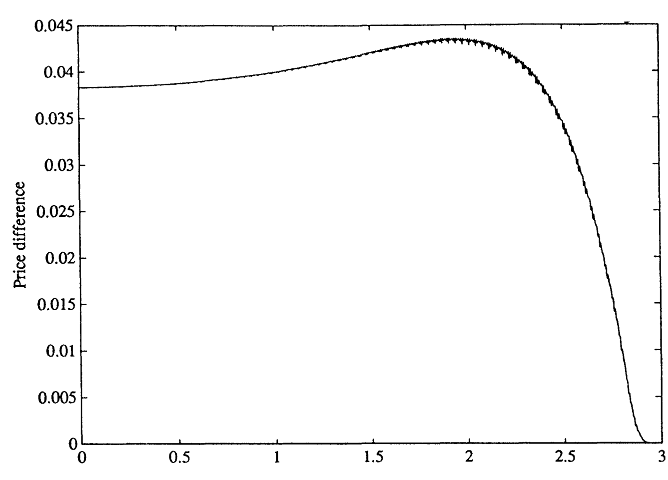

FIG 1. T.C. = 0.2 *percent*; $S = 19$; $E = 20$; $\gamma = 1.0$; $\sigma = 0.05$; $r = 8.5$ *percent*; $\alpha = 10$ *percent*.

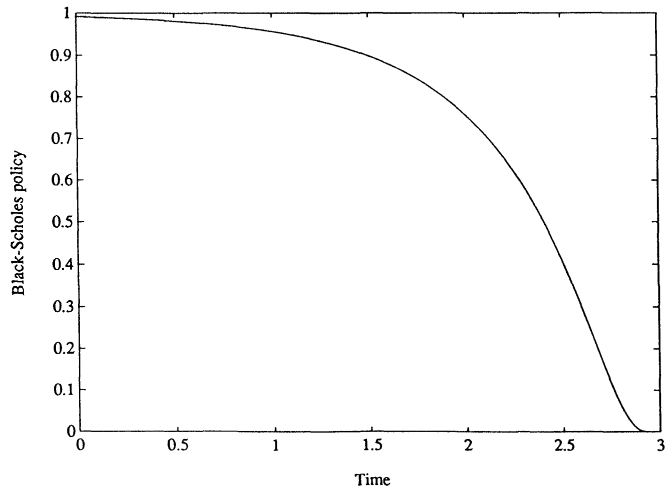

FIG 2. $S=19$; $E=20$; $\sigma=0.05$; $r=8.5$ *percent.*

which is the "hedging" strategy of the option writer, who performs few transactions long before expiration because of the wide range of possible paths of the stock price until the final time. Also, as the final time approaches, the price difference shows a "hump," which represents the period of active trading by the option writer. Finally, if $S > E$, the price difference is λS at the final time.

The variation of the peak of the price difference with he model parameters was investigated, and the following results were obtained. The "overshoot" ratio $((p_w - p_{b_s}) - \lambda S)/\lambda S$ was calculated, and it was observed that it is (i) a linear increasing function of the logarithm of the index of risk aversion γ (ii) a linear increasing function of the volatility σ, (iii) a decreasing function of the stock price S, (iv) a convex decreasing function of the proportional transaction charge $\lambda = \mu$, (v) a convex function of the interest rate r, and (vi) a linear decreasing function of the stock's mean growth rate α. These results are illustrated in Figs. 3-8, where the relevant parameter values

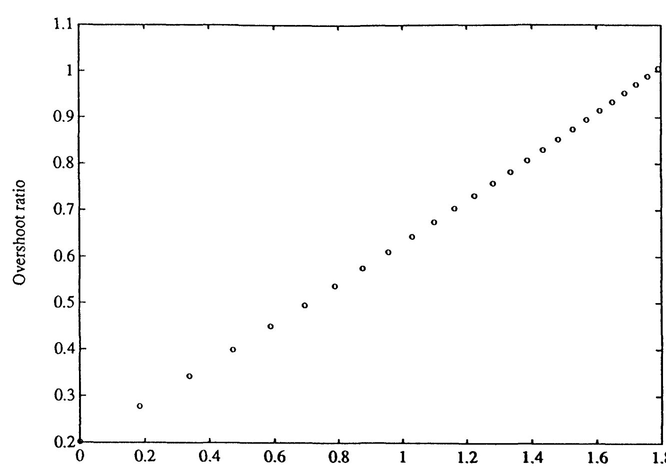

FIG. 3. T.C. $=0.2$ *percent*; $S=9$; $E=10$; $\sigma=0.075$; $r=7$ *percent*; $\alpha=10$ *percent.*

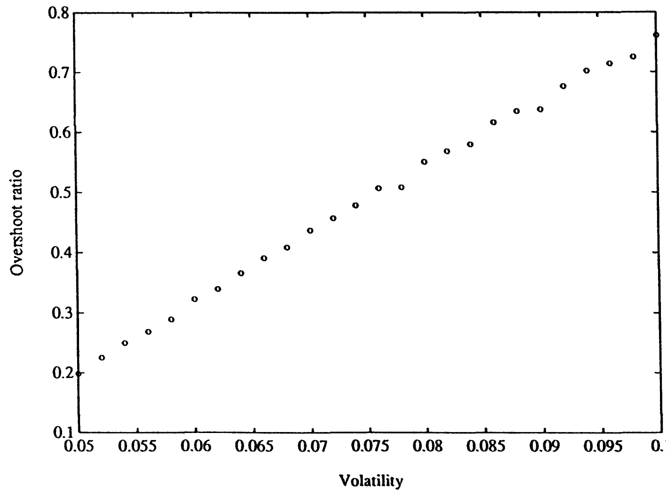

FIG. 4. T.C. = 0.2 *percent*; $S = 29$; $E = 30$; $\gamma = 1.0$; $r = 7$ *percent*; $\alpha = 10$ *percent*.

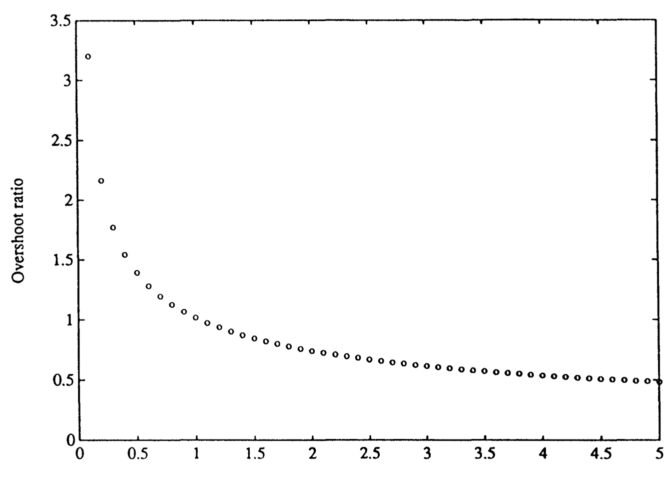

FIG. 5. $S = 29$; $E = 30$; $\gamma = 2.0$; $\sigma = 0.075$; $r = 7$ *percent*; $\alpha = 10$ *percent*.

are shown in their titles. The results are interpreted as follows. As the writer becomes more risk averse, the boundaries ∂B and ∂S come closer to the optimal trading strategy without transaction costs, thus mandating more transactions and increasing the option price. The linearity with $\ln(\gamma)$ is probably due to the form of the utility function. As the volatility of the underlying risky security increases, the uncertainty facing the option writer is greater, and the option price increases, as is in the case without transaction costs. As the stock price increases over the exercise price, the price difference is λS; at expiration, the Black-Scholes strategy $N(\chi)$ dictates holding one share of the stock until expiration, and the "hump" is small. As the transaction costs increase, λS increases; but the above ratio decreases as the writer tries to perform less transactions (the boundaries ∂B and ∂S move away from the optimal trading strategy without transaction costs). Finally, the mean growth rate α and the interest rate r have little effect on the above ratio as compared to other parameters.

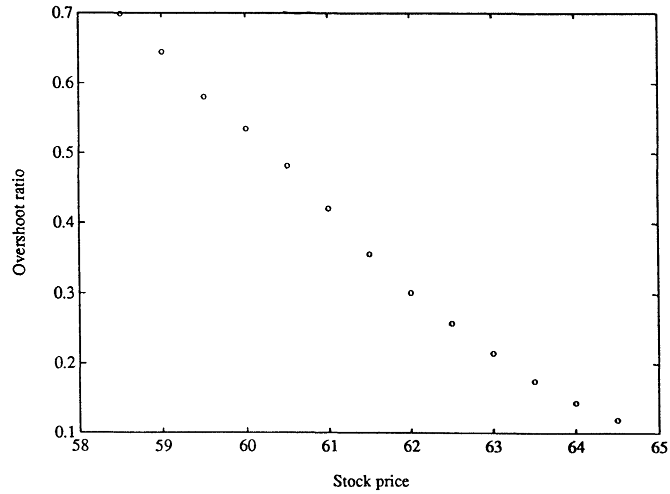

FIG. 6. T.C. $=0.2$ *percent*; $E=60$; $\gamma=1.0$; $\sigma=0.075$; $r=7$ *percent*; $\alpha=10$ *percent*.

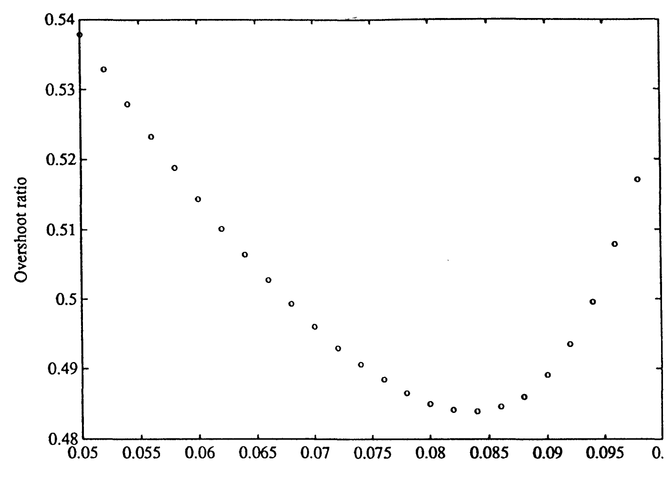

FIG. 7. T.C. $=0.2$ *percent*; $S=29$; $E=30$; $\gamma=1.0$; $\sigma=0.075$; $\alpha=10$ *percent*.

8. Concluding remarks. There are several directions in which our approach needs further investigation.

1. *Nonexponential utilities.* There is no issue of principle here, but only of a further increased computational load, since the reduction from four dimensions to three is no longer available. Since the risk averse writer's strategy is basically a hedging strategy, we believe that the form of the utility function is unimportant and that only its curvature at the origin plays any real role. If true, this would provide a justification for using the computationally simpler exponential function.

2. *Diversified portfolios.* As pointed out earlier, in our framework the writer may well wish to include other stocks (not just the one on which the option is written) in the hedging portfolio. Again, this simply increases the dimensionality of the problem. By allowing investment in other stocks we are enlarging the class of possible hedging strategies, and hence the option writing price will be reduced; we do not know, however, by how much.

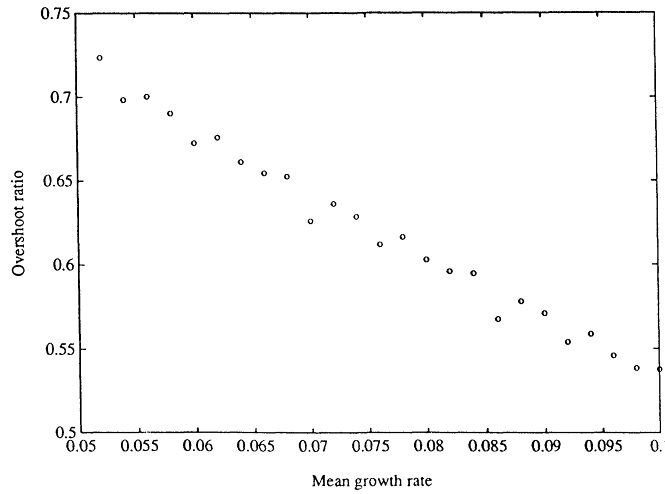

FIG. 8. T.C. $= 0.2$ *percent*; $S = 29$; $E = 30$; $\gamma = 1.0$; $\sigma = 0.075$; $r = 5$ *percent*.

3. *American options.* Clearly, a similar approach could be taken to the pricing of American options. There is, however, a conceptual problem in that the buyer, not the writer, controls the exercise strategy, and the pricing problem must involve the solution of one more utility maximisation problem over all the exercising strategies available to the buyer. In particular, there seems no reason why the buyer should use the "frictionless exercise strategy" as described in § 6 of Karatzas [13]. The precise definition of the problem is currently under investigation.

4. *Equilibrium.* Under what circumstances will a writer and a buyer agree on a "deal," i.e., a common price for an option contract in the framework we have described? This is a very important question; one that we do not claim to understand fully. It is possible to define a buying price in a way that mirrors our definition of the writing price, and this is what Hodges and Neuberger [11] do. In the notation of § 2, if the buyer forms a hedging portfolio whose composition at the exercise time T is $(B, \underline{y})$ then its cash value after exercise of the option is $B - E + c(\underline{y} + \underline{e}_1, \underline{S})$. Analogous to the definition (2.1) for V_w, we can therefore define

$$V_b(B) = \sup_{\pi \in \mathcal{T}(B)} \mathbb{E}\{\mathcal{U}(\mathbf{B}^\pi(T) + I_{(S_1(T) \leqq E)} c(\underline{\mathbf{y}}^\pi(T), \underline{\mathbf{S}}(T)) + I_{(S_1(T) > E)}[c(\underline{\mathbf{y}}^\pi(T) + \underline{e}_1, \underline{\mathbf{S}}(T)) - E])\} \tag{8.1}$$

and

$$B_b = \inf\{B \colon V_b(B) \geqq 0\}, \tag{8.2}$$

and the buying price p_b as

$$p_b = B_b - B_1, \tag{8.3}$$

where B_1 is given by (2.4). However, we do not believe this definition to be an appropriate one. The most obvious objection is that it is very hard to see how writer and buyer could ever agree on a price. Certainly $p_w \neq p_b$ if all parameters are the same for all calculations. One may hypothesise that writer and buyer agree on market parameters, but have different utility functions. This fails, however, because it is always the case that $p_w > p_{bs}$ and $p_b < p_{bs}$, where p_{bs} is the (preference-independent) Black-Scholes price. At a more fundamental level the above buying price seems inappropriate because it fails to respect the essential asymmetry in an option contract, namely that buying an option is a form of insurance, whereas writing one is a gamble. This distinction disappears in the Black–Scholes world, as there is no essential element of risk on either side, but no general theory of option pricing can be satisfactory if this distinction is not taken into account. This is an interesting area for further research.

Acknowledgments. The starting point for this work was a preliminary version of the paper [11], presented by Stewart Hodges and Anthony Neuberger at a workshop organised by the Financial Options Research Centre of Warwick University in 1989. Readers of [11] will appreciate that several of the key ideas were introduced in that paper. Collaborative work on the present paper was initiated while the first and third authors were visiting the Graduate School of Business, Stanford University, December 1990. We would like to thank Darrell Duffie, Avi Mandelbaum and the Finance and Decision Sciences groups at the Graduate School of Business at Stanford University for this opportunity as well as for stimulating comments on the work in progress. We also appreciate some incisive comments by Lucien Foldes at the London School of Economics and Political Science and discussions with Helyette Geman, which clarified our thinking.

REFERENCES

[1] G. BARLES AND P. E. SOUGANIDIS, *Convergence of approximation schemes for fully nonlinear second order equations*, Asymptotic Anal., 4 (1991), pp. 271–283.

[2] B. BENSAID, J. LESNE, H. PAGES, AND J. SCHEINKMAN, *Derivative asset pricing with transaction costs*, working paper, Bank of France, Centre de Recherche (1991).

[3] F. BLACK AND M. SCHOLES, *The pricing of options and corporate liabilities*, J. Political Economy, 81 (1973), pp. 637–659.

[4] CAPUZZO-DOLCETTA, AND P.-L. LIONS, *Hamilton–Jacobi equations and state constraints problems*, IMA preprint ser. 342, University of Minnesota, MN, 1987.

[5] J. C. COX AND M. RUBINSTEIN, *Options Markets*, Prentice-Hall, Englewood Cliffs, NJ, 1985.

[6] M. G. CRANDALL AND P.-L. LIONS, *Viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc., 277 (1983), pp. 1–42.

[7] M. G. CRANDALL, H. ISHII, AND P.-L. LIONS, *User's guide to viscosity solutions of second order partial differential equations*, Department of Mathematics, University of California, Santa Barbara, 1990, preprint.

[8] M. H. A. DAVIS AND A. R. NORMAN, *Portfolio selection with transaction costs*, Math. Oper. Res., 15 (1990), pp. 676–713.

[9] M. H. A. DAVIS AND V. G. PANAS, *European option pricing with transaction costs*, Proc. 30th I.E.E.E. Conference on Decision and Control (1991), pp. 1299–1304.

[10] C. EDIRISINGHE, V. NAIK, AND R. UPPAL, *Optimal replication of option with transaction costs*, University of British Columbia, working paper, 1991.

[11] S. D. HODGES AND A. NEUBERGER, *Optimal replication of contingent claims under transaction costs*, Rev. Futures Markets, 8 (1989), pp. 222–239.

[12] H. ISHII AND P.-L. LIONS, *Viscosity solutions of fully non-linear second-order elliptic partial differential equations*, J. Differential equations, 83, 1 (1990), pp. 26–78.

[13] I. KARATZAS, *Optimisation problems in the theory of continuous trading*, SIAM J. Control Optim., 27 (1989), pp. 1221–1259.

[14] M. KATSOULAKIS, *State-constraint problems for second order fully nonlinear degenerate partial differential equations*, Ph.D. thesis, Brown University, Providence, RI, 1991.

[15] L. H. MARTINS AND H. J. KUSHNER, *Numerical methods for stochastic singularly controlled problems*, SIAM J. Control Optim., 29 (1991), pp. 1443–1475.

[16] H. E. LELAND, *Option pricing and replication with transactions costs*, J. Finance, 40 (1985), pp. 1283–1301.

[17] P.-L. LIONS, *Optimal control of diffusion processes and Hamilton–Jacobi–Bellman equations*, parts 1 & 2, Comm. Partial Differential Equations, 8 (1983), pp. 1101–1174, 1229–1276.

[18] H. M. SONER, *Optimal control with state-space constraints*, SIAM J. Control Optim., 24 (1986), pp. 552–561.

[19] H. ZHU, *Characterisation of variational inequalities in singular control*, Ph.D. thesis, Brown University, Providence, RI, 1991.