



## Mathematics of Operations Research

Publication details, including instructions for authors and subscription information:  
<http://pubsonline.informs.org>

### Portfolio Selection with Transaction Costs

M. H. A. Davis, A. R. Norman,

To cite this article:

M. H. A. Davis, A. R. Norman, (1990) Portfolio Selection with Transaction Costs. *Mathematics of Operations Research* 15(4):676-713. <http://dx.doi.org/10.1287/moor.15.4.676>

Full terms and conditions of use: <http://pubsonline.informs.org/page/terms-and-conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact [permissions@informs.org](mailto:permissions@informs.org).

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

© 1990 INFORMS

Please scroll down for article—it is on subsequent pages



INFORMS is the largest professional society in the world for professionals in the fields of operations research, management science, and analytics.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

## PORTFOLIO SELECTION WITH TRANSACTION COSTS\*

M. H. A. DAVIS AND A. R. NORMAN†

In this paper, optimal consumption and investment decisions are studied for an investor who has available a bank account paying a fixed rate of interest and a stock whose price is a log-normal diffusion. This problem was solved by Merton and others when transactions between bank and stock are costless. Here we suppose that there are charges on all transactions equal to a fixed percentage of the amount transacted. It is shown that the optimal buying and selling policies are the local times of the two-dimensional process of bank and stock holdings at the boundaries of a wedge-shaped region which is determined by the solution of a nonlinear free boundary problem. An algorithm for solving the free boundary problem is given.

**1. Introduction.** This paper concerns the optimal investment and consumption decisions of an individual who has available just two investment instruments: a bank account paying a fixed interest rate  $r$ , and a risky asset ("stock") whose price is a geometric Brownian motion with expected rate of return  $\alpha$  and rate of return variation  $\sigma^2$ . Thus the stock grows at a mean rate  $\alpha$ , with white noise fluctuations. It is assumed that stock may be bought and sold in arbitrary amounts (not necessarily integral numbers of shares). The investor consumes at rate  $c(t)$  from the bank account; all income is derived from capital gains and consumption is subject to the constraint that the investor must be solvent, i.e. have nonnegative net worth, at all times. The investor's objective is to maximize the utility of consumption as measured by the quantity

$$(1.1) \quad \mathbb{E} \int_0^\infty e^{-\delta t} u(c(t)) dt.$$

Here  $\mathbb{E}$  denotes expectation and  $\delta > 0$  is the interest rate for discounting. In this paper the utility function  $u(c)$  will always be equal to  $c^\gamma/\gamma$  for some  $\gamma \in \Gamma := \{\gamma \in \mathbb{R} : \gamma < 1 \text{ and } \gamma \neq 0\}$  or  $u(c) = \log c$  ("log" denotes the natural logarithm). These functions form a subset of the so-called HARA (hyperbolic absolute risk aversion) class.

In the absence of any transactions between stock and bank, the investor's holding  $s_0(t)$  and  $s_1(t)$  in bank and stock respectively, expressed in monetary terms, evolve according to the following equations, the second of which is an Ito stochastic

\*Received July 27, 1987; revised April 24, 1989.

AMS 1980 subject classification. Primary: 93E20. Secondary: 90A16, 60H10.

LAOR 1973 subject classification. Main: Stochastic control. Cross references: Investment.

OR/MS Index 1978 subject classification. Primary: 200 Finance/portfolio. Secondary: 564 Probability/diffusion.

Key words. Portfolio selection, transaction costs, stochastic control, reflecting diffusions, free boundary problem, local time.

†Supported in part by an SERC Postgraduate Studentship.

differential equation driven by a standard Brownian motion  $z(t)$ .

$$(1.2) \quad ds_0(t) = (rs_0(t) - c(t)) dt,$$

$$(1.3) \quad ds_1(t) = \alpha s_1(t) dt + \sigma s_1(t) dz(t).$$

The investor starts off with an initial endowment  $s_0(0) = x$ ,  $s_1(0) = y$ . To complete the specification of the problem we need to state how funds are transferred from bank to stock and vice versa. In the original paper of R. C. Merton [18], and in almost all subsequent work in this area, it assumed that such transfers can be made instantly and costlessly. In this case we can re-parametrize the problem by introducing new variables  $w(t) = s_0(t) + s_1(t)$  (the total wealth) and  $\pi(t) = s_1(t)/w(t)$  (the fraction of total wealth held in stock). Adding (1.2) and (1.3) and using these variables gives us the basic wealth equation

$$(1.4) \quad dw(t) = [rw(t) + (\alpha - r)\pi(t)w(t) - c(t)] dt + \sigma\pi(t)w(t) dz(t),$$

$$w(0) = x + y.$$

Since transactions are free and instantaneous we can regard  $\pi(t)$ —as well as  $c(t)$ —as a *decision variable*. We now have a completely formulated stochastic control problem: choose nonanticipative processes  $\pi(t)$ ,  $c(t)$  so as to maximize (1.1) subject to (1.4) and the constraint  $w(t) \geq 0$  for all  $t$ .

It is a remarkable fact that this is one of the few nonlinear stochastic control problems that can be explicitly solved. It turns out, as we will show in §2 below, that for utility functions in the HARA class the optimal investment strategy is to keep a *constant fraction of total wealth in the risky asset*, and to *consume at a rate proportional to total wealth*, i.e. the optimal  $\pi(t)$ ,  $c(t)$  are  $\pi(t) = \pi^*$  and  $c(t) = Cw(t)$  for some constants  $\pi^*$ ,  $C$ . This means that, optimally, the investor acts in such a way that the portfolio holdings are always on the line  $s_1 = [\pi^*/(1 - \pi^*)]s_0$  in the  $(s_0, s_1)$  plane; we shall refer to this as the “Merton line” (Figure 1).

In a recent paper by Karatzas, Lehoczky, Sethi and Shreve [12], the constraint  $w(t) \geq 0$  is replaced by the stipulation that evolution of the wealth process terminates on bankruptcy (i.e. the first time at which  $w(t) = 0$ ), at which point the investor is retired on a lump-sum pension  $P$ . It turns out that if  $P$  is not large enough then it is optimal to avoid bankruptcy altogether (this can always be done) and then the policy described above is optimal. All of these results hold under a basic well-posedness condition, namely

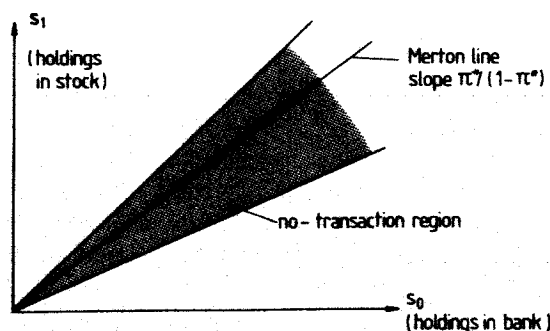


FIGURE 1. Space of Bank and Stock Holdings, Showing “Merton Line” and No-Transaction Wedge.

Condition A.  $\delta > \gamma[r + (\alpha - r)^2/\sigma^2(1 - \gamma)]$ .

If this condition is violated then growth of discounted utility is possible and arbitrarily large utility may be obtained by policies of prolonged investment followed by massive consumption. The well-posedness condition for  $u(c) = \log c$  is  $\delta > 0$ , which is just Condition A with  $\gamma = 0$ .

Any attempt to apply Merton's strategy in the face of transaction costs would result in immediate penury, since incessant trading is necessary to hold the portfolio on the Merton line. There must in such a case be some "no-transaction" region in  $(s_0, s_1)$  space inside which the portfolio is insufficiently far "out of line" to make trading worthwhile. In this paper we shall consider the case of proportional transaction costs:<sup>1</sup> the investor pays fractions  $\lambda$  and  $\mu$  of the amount transacted, on purchase and sale of stock respectively. All such charges are paid from the bank account. In this case the bank and stock holdings must retain their separate identities rather than being merged into a single wealth process. The equations describing their evolution are

$$(1.5) \quad \begin{aligned} ds_0(t) &= [rs_0(t) - c(t)] dt - (1 + \lambda) dL_t + (1 - \mu) dU_t, \\ ds_1(t) &= \alpha s_1(t) dt + \sigma s_1(t) dz(t) + dL_t - dU_t, \end{aligned}$$

where  $L_t, U_t$  represent cumulative purchase and sale of stock on the time interval  $[0, t]$  respectively. This allows for instantaneous purchase or sale of finite amounts of stock as well as purchase and sale at a given rate and various other sorts of behaviour. One notices from (1.5) that purchase of  $dL$  units of stock requires a payment  $(1 + \lambda)dL$  from the bank, while sale of  $dU$  units of stock realizes only  $(1 - \mu)dU$  in cash. Obviously, it will never be optimal to buy and sell at the same time. If Condition A does not hold then arbitrarily high utility can be achieved as described above (the "prolonged investment" just has to be a little more prolonged). Under Condition A, we will show that *the no-transaction region is a wedge containing the Merton line* (Fig. 1); equivalently, the proportion of total wealth held in stock should be maintained between fractions  $\pi_1^*$  and  $\pi_2^*$ , which of course depend on  $\lambda$  and  $\mu$  as well as the other constants in the problem. For example, when  $\lambda = \mu = 0.015$  and  $r = 0.07$ ,  $\alpha = 0.12$ ,  $\sigma = 0.4$ ,  $\delta = 0.1$ , and the utility function is  $u(c) = -1/c$  we find that  $\pi_1^* = 9.0\%$ ,  $\pi_2^* = 19.8\%$  whereas the Merton proportion is  $\pi^* = 15.6\%$  (see Figure 3 below). There is no closed-form expression for  $\pi_1^*, \pi_2^*$  but we state how to compute them. The optimal transaction policy is *minimal trading* to stay inside the wedge, preceded by an immediate transaction to the closest point in the wedge if the initial endowment is outside it. More technically, the optimally controlled process is a reflecting diffusion inside the wedge and the buying and selling policies  $(L_t, U_t)$  are the local times at the lower and upper boundaries respectively. Consumption takes place at a finite rate in the interior of the wedge (in which the process lies almost all of the time).

Our interest in this problem was aroused by the stimulating paper of Magill and Constantinides [17] on the same subject. This paper contains the fundamental insight that the no-transaction region is a wedge, but the argument is heuristic at best and no clear prescription as to how to compute the location of the boundaries, or what the controlled process should do when it reaches them, is given. The paper was in fact ahead of its time, in that an essential ingredient of any rigorous formulation, namely

<sup>1</sup>The case of fixed costs, where the investor pays a flat transaction fee regardless of the amount transacted, remains largely unexplored; this is a problem of *impulse control* [3]. There is some work in this direction by Duffie and Sun [7].

the theory of local time and reflecting diffusion, was unavailable to the authors, being at that time (1976) the exclusive property of a small band of pure mathematical votaries. Needless to say, Magill's and Constantinides' paper is far more valuable than many others of unimpeachable mathematical rectitude.

Stochastic control problems involving local time have received much attention in recent years. Early pioneering work of Bather and Chernoff was followed by the appearance of papers by Beneš, Shepp and Witsenhausen [2] and Harrison and Taylor [10] in which problems of "finite fuel" control and regulation of "Brownian storage systems" were solved rigorously, taking advantage of developments in stochastic calculus in the 1970s which made it possible to handle local times and reflecting diffusions in simple domains in a relatively straightforward way. The relevant theory can now be found in very compact form in Harrison's book [9]. All of these works, and the present paper, essentially concern free boundary problems, and indeed many of them are closely related to optimal stopping, as Karatzas and Shreve [14] have shown. This paper differs from all others we are aware of, however, in that our problem involves "continuous control" (i.e. consumption) as well as "singular control" (transactions). This leads to a free boundary problem for a *nonlinear* partial differential equation (PDE) as opposed to the linear PDEs which arise when singular control is the *only* control. The problem is for this reason substantially more delicate.

Three papers directly related to the present work are Constantinides [4], Duffie and Sun [7] and Taksar, Klass and Assaf [20]. Constantinides considers essentially the same problem as ours (or as the earlier paper of Magill and Constantinides [17]) and proposes an approximate solution based on making certain assumptions on the consumption process. Some further remarks on his results will be found in §7 below. Duffie and Sun [7] consider the case of fixed plus proportional transaction charges. Their results are quite different in character from ours. Taksar, Klass and Assaf [20], using the model (1.5) with  $c \equiv 0$  (no consumption), study the problem of maximizing the long-run growth rate

$$\mathbb{E} \left\{ \liminf_{t \rightarrow \infty} \frac{1}{t} \log(s_0(t) + s_1(t)) \right\}.$$

In an ingenious analysis they reduce the problem to a 1-dimensional one and show that a "two-sided regulator" is optimal, which means that the process  $(s_0(t), s_1(t))$  is, as in our case, optimally kept inside a wedge by reflection at the boundaries. The solution thus looks very similar to ours, but the details of the problem and the method of analysis are completely different. A study of the effects of transaction costs on *option pricing* is given by Leland [15].

The present paper is organized as follows. In §2 we give a self-contained treatment of the Merton (no transaction cost) problem. This is included because later on we need to use comparison arguments involving the Merton case, and also because no simple complete treatment seems to be readily available.<sup>2</sup> The transaction costs problem is formulated in §3, where we give informal arguments which indicate why the no-transaction region is wedge-shaped. We also show how the analytic problem may be reduced to a one-dimensional free boundary problem. In §4 we prove "verification theorems" which show that if the free boundary problem can be solved then a policy of minimal transaction to stay within the wedge defined by its solution is indeed optimal. Theorem 5.1 in §5 gives conditions under which the free boundary problem is solvable. In §6 we obtain a semimartingale representation of the evolution of the "value process". Apart from having some intrinsic interest, this is needed to

<sup>2</sup>The extra generality of the treatment in [12] necessitates more complicated arguments.



complete some technical argument in §4, and can be used to show that under the optimal policy the investor does not reach bankruptcy in finite time.

The main results of the paper are summarized in Theorem 7.1 in §7. This section also contains an algorithm for solving the free boundary problem together with numerical results, as well as concluding remarks. Finally, the Appendix contains a technical analysis of some differential equations arising in the solution of the free boundary problem.

A preliminary account of some of this work was given in Davis [5].

**2. No transaction costs: the Merton problem.** Throughout the paper  $(\Omega, \mathcal{F}, \mathbb{P})$  will denote a fixed complete probability space and  $(\mathcal{F}_t)_{t \geq 0}$  a given filtration, i.e. a family of sub- $\sigma$ -fields of  $\mathcal{F}$  such that (i)  $\mathcal{F}_s \subseteq \mathcal{F}_t$  for  $s \leq t$  and (ii) for each  $t \geq 0$ ,  $\mathcal{F}_t$  contains all null sets of  $\mathcal{F}$ . The stochastic process<sup>3</sup>  $(z_t)_{t \geq 0}$  will be a standard Brownian motion with respect to  $(\mathcal{F}_t)$ , i.e.  $(z_t)$  has almost all sample paths continuous, is adapted to  $(\mathcal{F}_t)$ , and for each  $s, t \geq 0$  the increment  $z_{t+s} - z_t$  is independent of  $\mathcal{F}_t$  and is normally distributed with mean 0 and variance  $s$ .

In this section we study the "Merton problem" of choosing investment and consumption policies  $(\pi_t, c_t)$  so as to maximize utility when wealth evolves according to equation (1.4). Let  $\mathcal{U}$  denote the set of policies. A policy is a pair  $(c, \pi)$  of  $\mathcal{F}_t$ -adapted processes such that

- (2.1) (i)  $c(t, \omega) \geq 0$  and  $\int_0^t c(s, \omega) ds < \infty$  for all  $(t, \omega)$ ,  
 (ii)  $|\pi(t, \omega)| \leq K$  for all  $(t, \omega)$ , where  $K$  is a constant which may vary from policy to policy, and  
 (iii)  $w(t, \omega) \geq 0$  for all  $(t, \omega)$ , where  $(w_t)$  is the unique strong solution of the wealth equation

$$(2.2) \quad dw_t = (rw_t + (\alpha - r)\pi_t w_t - c_t) dt + w_t \pi_t \sigma dz_t,$$

$$w_0 = w.$$

Here  $w \in \mathbb{R}_+$  is the initial endowment and  $\alpha, r, \sigma$  are positive constants as described in §1. The existence of such a solution follows from standard theorems, which show that the map  $t \rightarrow w_t(\omega)$  is continuous for almost all  $\omega \in \Omega$ . Let  $\tau := \inf t: w_t = 0$ . We note that  $(c, \pi) \in \mathcal{U}$  only if  $c_t = 0$  a.s. for all  $t \geq \tau$  and hence  $w_t = 0$  for all  $t \geq \tau$ . The investment problem is to choose  $(c, \pi) \in \mathcal{U}$  so as to maximize

$$(2.3) \quad J_w(c, \pi) := \mathbb{E}_w \int_0^\infty e^{-\delta t} u(c_t) dt.$$

**THEOREM 2.1.** (a) *Suppose Condition A holds and that  $u(c) = c^\gamma/\gamma$ ,  $\gamma \in \Gamma$ . Define*

$$(2.4) \quad C = \frac{1}{1-\gamma} \left[ \delta - \gamma r - \frac{\gamma \beta^2}{2(1-\gamma)} \right]$$

<sup>3</sup>Throughout the paper we take the usual probabilist's license of denoting a random process such as  $(z_t)$  interchangeably as  $z_t$ ,  $z(t)$ ,  $z_t(\omega)$  or  $z(t, \omega)$ . All exogenously defined processes are assumed to be measurable.

where  $\beta := (\alpha - r)/\sigma$ . Then  $\sup_{(c, \pi) \in \Pi} J_w(c, \pi) = v(w)$  where

$$(2.5) \quad v(w) = \frac{1}{\gamma} C^{\gamma-1} w^\gamma.$$

The optimal policy is

$$(2.6) \quad c_t^* = C w_t, \quad \pi_t^* = \frac{\beta}{(1 - \gamma)\sigma}.$$

(b) Suppose  $u(c) = \log c$  and that  $\delta > 0$ . Then

$$\sup_{(c, \pi) \in \Pi} J_w(c, \pi) = \frac{1}{\delta^2} \left[ r + \frac{1}{2} \beta^2 - \delta \right] + \frac{1}{\delta} \log \delta w$$

and the optimal policy is  $c_t^* = \delta w_t$ ,  $\pi_t^* = \beta/\sigma$ .

REMARK 2.2. As noted in §1,  $\pi^*$  is constant and  $c^*$  is proportional to current wealth. The condition  $\delta > 0$  and the optimal policies  $(c^*, \pi^*)$  of case (b) are formally obtained from case (a) by setting  $\gamma = 0$ . We see from (2.6) that  $\pi^* \in (0, 1)$  only when  $r < \alpha < r + (1 - \gamma)\sigma^2$ . This is *hedging*: assets are split between stock and bank to reduce volatility. If  $\alpha > r + (1 - \gamma)\sigma^2$  then *leverage* is optimal: funds are borrowed from bank to invest in stock,<sup>4</sup> while if  $\alpha < r$  then the converse—i.e. *shortselling*—is optimal. When  $\alpha = r$ ,  $\pi^* = 0$ ; in this case  $v(w) = ((\delta - \gamma r)/(1 - \gamma))^{\gamma-1} \gamma^{-1} w^\gamma$ , which is just the utility of optimally consuming an initial endowment  $w$  in the bank.

PROOF. We will only prove part (a). Part (b) is proved by exactly the same arguments. The proof is an application of dynamic programming, cf. Chapter VI of Fleming and Rishel [8]. When  $c, \pi$  are constants the wealth process  $w_t$  given by (2.2) is a diffusion process with generator

$$A^{c, \pi} \bar{v}(w) = [(t + (\alpha - r)\pi)w - c] \bar{v}'(w) + \frac{1}{2} w^2 \pi^2 \sigma^2 \bar{v}''(w)$$

acting on  $C^2$  functions  $\bar{v}$  ( $\bar{v}' = d\bar{v}/dw$ ). The so-called *Bellman equation* of dynamic programming for maximizing (2.2) over control policies  $(c_t, \pi_t)$ , to be solved for a function  $\bar{v}$ , is

$$(2.7) \quad \max_{c, \pi} \left\{ A^{c, \pi} \bar{v} + \frac{1}{\gamma} c^\gamma - \delta \bar{v} \right\} = 0.$$

The maxima are achieved at

$$c = (\bar{v}')^{-1/(1-\gamma)}, \quad \pi = \frac{-\beta \bar{v}'}{w \sigma \bar{v}''},$$

so that (2.7) is equivalent to

$$(2.8) \quad r w \bar{v}' - \frac{\beta^2 (\bar{v}')^2}{2 \bar{v}''} + \frac{1 - \gamma}{\gamma} (\bar{v}')^{-\gamma/(1-\gamma)} - \delta \bar{v} = 0.$$

The justification for introducing this equation will be seen when we obtain the semimartingale decomposition of the process  $M_t$  defined below. (2.8) is satisfied by

<sup>4</sup>We are assuming that interest rates are the same on lending and borrowing.

$\tilde{v} = v$  given by (2.5) and the maximizing  $\pi$  and  $c$  are equal to  $\pi^*$ ,  $c^*$  given by (2.6). Condition (2.3) ensures that  $c^*(t) \geq 0$ .

Let  $(c, \pi) \in \mathbb{U}$  be an arbitrary policy. The solution of (2.2) is given by

$$(2.9) \quad w_t = \exp \left\{ rt + \int_0^t \left( (\alpha - r)\pi_u - \frac{1}{2}\sigma^2\pi_u^2 \right) du + \sigma \int_0^t \pi_u Dz_u \right\} \\ \times \left\{ w - \int_0^t c_s \exp \left[ - \int_0^s \left( (\alpha - r)\pi_u - \frac{1}{2}\sigma^2\pi_u^2 \right) du - \sigma \int_0^s \pi_u dz_u \right] ds \right\}.$$

It follows from Hölder's inequality and the fact that  $\pi$  is bounded that  $w_t$  has finite moments of all orders, and it is also clear that  $w_t \leq w_t^0$  where  $w_t^0$  is the solution of (2.2) with the same  $\pi$  and  $c_t \equiv 0$ . Now define

$$(2.10) \quad M_t := \frac{1}{\gamma} \int_0^t e^{-\delta s} c_s^\gamma ds + e^{-\delta t} v(w_t),$$

where  $v$  is defined by (2.5). By the Ito formula,

$$M_t - M_0 = \int_0^t e^{-\delta s} \left[ A^{c, \pi} v + \frac{1}{\gamma} c_s^\gamma - \delta v \right] ds + \sigma C^{\gamma-1} \int_0^t e^{-\delta s} \pi_s w_s^\gamma dz_s.$$

It follows from the above argument that the second term on the right is a martingale, while the first term is, in view of (2.7), a decreasing process which is equal to zero when  $(c, \pi) = (c^*, \pi^*)$ . Thus  $M_t$  is a supermartingale, and is a martingale when  $(c, \pi) = (c^*, \pi^*)$ , so that

$$(2.11) \quad v(w) = M_0 \geq \mathbb{E}_w M_t = \mathbb{E}_w \frac{1}{\gamma} \int_0^t e^{-\delta s} c_s^\gamma ds + \mathbb{E}_w e^{-\delta t} v(w_t).$$

By using the Ito formula and the wealth equation (2.2) we find that

$$(2.12) \quad e^{-\delta t} w_t^\gamma = w_0^\gamma G_t \exp \left\{ \int_0^t a(s) ds \right\},$$

where  $G_t$  is the exponential martingale (it is a martingale since  $\pi$  is bounded)

$$G_t = \exp \left\{ \int_0^t \gamma \pi_s \sigma dz_s - \frac{1}{2} \int_0^t \gamma^2 \pi_s^2 \sigma^2 ds \right\} \quad \text{and}$$

$$a(s) = \gamma \left[ r + (\alpha - r)\pi_s - (c_s/w_s) - \frac{1}{2}(1 - \gamma)\pi_s^2 \sigma^2 \right] - \delta.$$

When  $(c, \pi) = (c^*, \pi^*)$  we find that  $a(s) = -C$  and hence from (2.12) that  $\mathbb{E}_w e^{-\delta t} v(w_t) \rightarrow 0$  as  $t \rightarrow \infty$ . It now follows from (2.11) that  $v(w) = J_w(c^*, \pi^*)$ .

To complete the proof we have to consider the cases  $0 < \gamma < 1$ ,  $\gamma < 0$  separately. In the former,

$$(2.13) \quad a(s) \leq -(1 - \gamma)C,$$

equality being achieved when  $c_s = 0$  and  $\pi_s = \pi^*$ . This implies as above that  $\mathbb{E}_w e^{-\delta t} v(w_t) \rightarrow 0$  as  $t \rightarrow \infty$  and hence from (2.11) that, for any policy  $(c, \pi)$ ,  $v(w) \geq J_w(c, \pi)$ . Thus  $(c^*, \pi^*)$  is optimal.



Now take  $\gamma < 0$ . The preceding argument fails because there is no longer an *a priori* upper bound for  $a(s)$ . Instead, we proceed as follows. For  $\epsilon > 0$  define

$$v_\epsilon(w) = \frac{1}{\gamma} C^{\gamma-1} (w + \epsilon)^\gamma.$$

Then  $v'_\epsilon(w) = v'(w + \epsilon)$  etc. and we see from (2.8) that  $v_\epsilon$  satisfies

$$r(w + \epsilon)v'_\epsilon(w) - \frac{\beta^2 (v'_\epsilon(w))^2}{2 v''_\epsilon(w)} + \frac{1 - \gamma}{\gamma} (v'_\epsilon(w))^{-\gamma/(1-\gamma)} - \delta v_\epsilon(w) = 0.$$

Since  $v'_\epsilon(w) > 0$  this shows that

$$rwv'_\epsilon - \frac{\beta^2 (v'_\epsilon)^2}{2 v''_\epsilon} + \frac{1 - \gamma}{\gamma} (v'_\epsilon)^{-\gamma/(1-\gamma)} - \delta v_\epsilon < 0,$$

i.e.

$$\max_{c, \pi} \left\{ A^{c, \pi} v_\epsilon + \frac{1}{\gamma} c^\gamma - \delta v_\epsilon \right\} < 0.$$

Now  $v_\epsilon$  is bounded on  $\mathbb{R}_+$  with bounded first and second derivatives. We easily conclude, by introducing the process  $M_t$  as in (2.10) but with  $v_\epsilon$  replacing  $v$ , that for any  $(c, \pi) \in \Pi$  and  $w > 0$ ,  $v_\epsilon(w) > J_w(c, \pi)$ . Since  $v_\epsilon(w) \downarrow v(w)$  as  $\epsilon \downarrow 0$  this shows that  $v(w) \geq \sup_{(c, \pi) \in \Pi} J_w(c, \pi)$ . But we know that  $v(w) = J_w(c^*, \pi^*)$ , so  $(c^*, \pi^*)$  is in fact optimal, as claimed.

**COROLLARY 2.3.** *Under the optimal policy  $(c^*, \pi^*)$  the wealth process  $w(t)$  is given by*

$$w(t) = w(0) \exp \left( \frac{1}{(1 - \gamma)} \left( r - \delta + \frac{\beta^2}{2} \left( \frac{2 - \gamma}{1 - \gamma} \right) \right) t \right) \exp \left( \frac{\beta}{(1 - \gamma)} z_t - \frac{1}{2} \frac{\beta^2}{(1 - \gamma)^2} t \right).$$

(This formula also applies when  $u(c) = \log c$ , setting  $\gamma = 0$ .) Using (2.5) we find that the evolution of *utility* is as follows:

$$(2.14) \quad v(w(t)) = v(w(0)) \exp \left( \frac{\gamma}{(1 - \gamma)} \left( r - \delta + \frac{\beta^2}{2(1 - \gamma)} \right) t \right) \\ \times \exp \left( \frac{\gamma\beta}{(1 - \gamma)} z_t - \frac{1}{2} \frac{\gamma^2 \beta^2}{(1 - \gamma)^2} t \right)$$

(cf. Remark 6.3 below).

**3. Transaction costs: preliminary discussion.** Let us now consider the situation, outlined in §1, in which transaction charges are imposed equal to a constant fraction of the amount transacted, the fractions being  $\lambda$  and  $\mu$  on purchase and sale respectively. The investor's holdings in bank and stock at time  $t$  are denoted  $s_0(t), s_1(t)$  and these are constrained to lie in the closed solvency region

$$\mathcal{S}_{\lambda, \mu} = \{(x, y) \in \mathbb{R}^2: x + (1 - \mu)y \geq 0 \text{ and } x + (1 + \lambda)y \geq 0\}.$$

We denote by  $\partial_\mu^+, \partial_\lambda^-$  the upper and lower boundaries of  $\mathcal{S}_{\lambda, \mu}$  respectively (see

Figure 2 below). It is clear that the investor's net worth is zero on  $\partial_\mu^+ \cup \partial_\lambda^-$ . A *policy* for investment and consumption is any triple  $(c_t, L_t, U_t)$  of adapted processes such that  $(c_t)$  satisfies (2.1) and  $(L_t)$  and  $(U_t)$  are right-continuous and nondecreasing with  $L_0 = U_0 = 0$ . ( $L$  and  $U$  are the cumulative purchases and sales of stock respectively.) The investor's holdings  $(s_0(t), s_1(t))$  starting with an endowment  $(x, y) \in \mathcal{S}_{\lambda, \mu}$  evolve in the following way in response to a given policy  $(c, L, U)$ :

$$(3.1) \quad \begin{aligned} ds_0(t) &= (rs_0(t) - c(t)) dt - (1 + \lambda) dL_t + (1 - \mu) dU_t, & s_0(0) &= x, \\ ds_1(t) &= \alpha s_1(t) dt + \sigma s_1(t) dz_t + dL_t - dU_t, & s_1(0) &= y. \end{aligned}$$

It follows from Doléans-Dade [6] that equations (3.1), have a unique strong solution at least up to the bankruptcy time  $\tau = \inf\{t \geq 0: (s_0(t), s_1(t)) \notin \mathcal{S}_{\lambda, \mu}\}$ .

An *admissible policy* is a policy  $(c, L, U)$  for which  $\tau = \infty$  a.s. or, equivalently, for which  $\mathbb{P}[(s_0(t), s_1(t)) \in \mathcal{S}_{\lambda, \mu} \text{ for all } t \geq 0] = 1$ . We denote by  $\mathcal{U}$  the set of admissible policies. This set is clearly nonempty; indeed let  $(c, L, U)$  be any policy such that  $(s_0(t), s_1(t))$  does not jump out of  $\mathcal{S}_{\lambda, \mu}$  (i.e. it is never the case that  $(s_0(t^-), s_1(t^-)) \in \mathcal{S}_{\lambda, \mu}$  but  $(s_0(t), s_1(t)) \notin \mathcal{S}_{\lambda, \mu}$ ). Then an admissible policy  $(\tilde{c}, \tilde{L}, \tilde{U})$  can be constructed by terminating  $(c, L, U)$  at bankruptcy, i.e. setting  $(\tilde{c}(t), \tilde{L}(t), \tilde{U}(t)) = (c(t), L(t), U(t))$  for  $t < \tau$ ,

$$\begin{aligned} \Delta \tilde{U}(\tau) &= s_1(\tau^-) \Bigg\} & \text{if } (s_0(\tau^-), s_1(\tau^-)) &\in \mu^+, \\ \Delta \tilde{L}(\tau) &= 0 \end{aligned}$$

$$\begin{aligned} \Delta \tilde{U}(\tau) &= 0 \Bigg\} & \text{if } (s_0(\tau^-), s_1(\tau^-)) &\in \partial_\lambda^- \\ \Delta \tilde{L}(\tau) &= s_1(\tau^-) \end{aligned}$$

and  $\tilde{L}(t) = \tilde{L}(\tau)$ ,  $\tilde{U}(t) = \tilde{U}(\tau)$ ,  $c(t) = 0$  for  $t \geq \tau$ . Then  $s_0(t) = s_1(t) = 0$  for  $t \geq \tau$  under  $(\tilde{c}, \tilde{L}, \tilde{U})$ .

The investor's objective is to maximize over  $\mathcal{U}$  the utility

$$J_{x,y}(c, L, U) = \mathbb{E}_{x,y} \int_0^\infty e^{-\delta t} u(c(t)) dt.$$

Here  $\mathbb{E}_{x,y}$  denotes the expectation given that the initial endowment is  $s_0(0) = x$ ,  $s_1(0) = y$ . Define the value function  $v$  as:

$$(3.2) \quad v(x, y) = \sup_{(c, L, U) \in \mathcal{U}} J_{x,y}(c, L, U).$$

The following properties of  $v$  are easily established directly from the definition.

**THEOREM 3.1.** Suppose  $u(c) = c^\gamma/\gamma$  for  $\gamma \in \Gamma$ , or  $u(c) = \log c$ . Then

- (a)  $v$  is concave.
- (b)  $v$  has the homothetic property: for  $\rho > 0$

$$v(\rho x, \rho y) = \rho^\gamma v(x, y), \quad [u(c) = c^\gamma/\gamma],$$

$$v(\rho x, \rho y) = \frac{1}{\delta} \log \rho + v(x, y), \quad [u(c) = \log c].$$

PROOF. (a) This is easily established by considering convex combinations of initial states and control process and using the linearity of equations (3.1) and concavity of the utility function. This idea appears in [13].

(b) Denote by  $\mathbb{U}(x, y)$  the class of admissible policies starting at  $(s_0(0), s_1(0)) = (x, y) \in \mathcal{S}_{\lambda, \mu}$ . Then it is easily checked from equations (3.1) that for any  $\rho > 0$ .

$$\mathbb{U}(\rho x, \rho y) = \{(\rho c, \rho L, \rho U) : (c, L, U) \in \mathbb{U}(x, y)\}.$$

Thus

$$\begin{aligned} v(\rho x, \rho y) &= \sup_{\mathbb{U}(\rho x, \rho y)} \mathbb{E}_{\rho x, \rho y} \int_0^\infty e^{-\delta t} u(c_t) dt \\ &= \sup_{\mathbb{U}(x, y)} \mathbb{E}_{x, y} \int_0^\infty e^{-\delta t} u(\rho c_t) dt =: \hat{v}. \end{aligned}$$

When  $u(c) = c^\gamma/\gamma$  we have  $u(\rho c) = \rho^\gamma u(c)$  so that  $\hat{v} = \rho^\gamma v(x, y)$ , whereas when  $u(c) = \log c$  then  $u(\rho c) = \log \rho + u(c)$  and  $\hat{v} = (\log \rho)/\delta + v(x, y)$ . This completes the proof.

In order to get some idea as to the nature of optimal policies, let us take  $u(c) = c^\gamma/\gamma$  and consider a restricted class of policies in which  $L$  and  $U$  are constrained to be absolutely continuous with bounded derivatives, i.e.

$$L_t = \int_0^t l_s ds, \quad U_t = \int_0^t u_s ds, \quad 0 \leq l_s, u_s \leq \kappa.$$

Equation (3.1) is then a vector SDE with controlled drift and the problem may be attacked in exactly the same way as in §2. The Bellman equation, to be solved for the value function  $\bar{v}$ , is

$$\max_{c, l, u} \left\{ A^{c, l, u} \bar{v}(x, y) + \frac{1}{\gamma} c^\gamma - \delta \bar{v}(x, y) \right\} = 0,$$

where  $A^{c, l, u}$  is the generator of (3.1) for fixed  $c, l, u$ . Written out in full this becomes

$$\begin{aligned} (3.3) \quad \max_{c, l, u} \left\{ \frac{1}{2} \sigma^2 y^2 \bar{v}_{yy} + r x \bar{v}_x + \alpha y \bar{v}_y + \frac{1}{\gamma} c^\gamma - c \bar{v}_x \right. \\ \left. + [-(1 + \lambda) \bar{v}_x + \bar{v}_y] l + [(1 - \mu) \bar{v}_x - \bar{v}_y] u - \delta \bar{v} \right\} = 0 \end{aligned}$$

where  $\bar{v}_x = \partial \bar{v} / \partial x$ ,  $\bar{v}_y = \partial \bar{v} / \partial y$ . Note that both of these derivatives must be positive since extra wealth will provide increased utility. The maxima are achieved as follows:

$$c = (\bar{v}_x)^{1/(\gamma-1)},$$

$$l = \begin{cases} \kappa & \text{if } \bar{v}_y \geq (1 + \lambda) \bar{v}_x, \\ 0 & \text{if } \bar{v}_y < (1 + \lambda) \bar{v}_x, \end{cases}$$

$$u = \begin{cases} 0 & \text{if } \bar{v}_y > (1 - \mu) \bar{v}_x, \\ \kappa & \text{if } \bar{v}_y \leq (1 - \mu) \bar{v}_x. \end{cases}$$

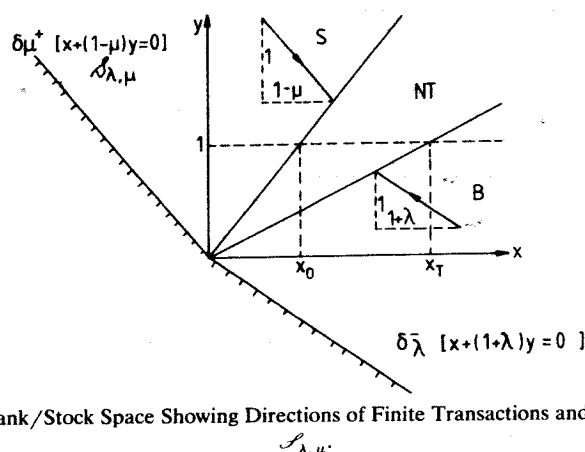


FIGURE 2. Bank/Stock Space Showing Directions of Finite Transactions and Solvency Region

This indicates that the optimal transaction policies are *bang-bang*: buying and selling either take place at maximum rate or not at all, and the solvency region  $\mathcal{S}_{\lambda, \mu}$  splits into three regions, “buy” ( $B$ ), “sell” ( $S$ ) and “no transactions” ( $NT$ ). At the boundary between the  $B$  and  $NT$  regions,  $\tilde{v}_y = (1 + \lambda)\tilde{v}_x$  whereas at the boundary between  $NT$  and  $S$ ,  $\tilde{v}_y = (1 - \mu)\tilde{v}_x$ . We now have to consider what shape these boundaries are, and here we use the homothetic property (b) of Theorem 3.1. This property does not hold for the restricted problem we are presently considering, but will hold in the limit as  $\kappa \rightarrow \infty$ . Assuming that  $\tilde{v}$  is  $C^1$  and homothetic, we find by direct calculation that for  $\rho > 0$

$$\tilde{v}_x(\rho x, \rho y) = \rho^{\gamma-1} \tilde{v}_x(x, y), \quad \tilde{v}_y(\rho x, \rho y) = \rho^{\gamma-1} \tilde{v}_y(x, y).$$

It follows that if  $\tilde{v}_y(x, y) = (1 + \lambda)\tilde{v}_x(x, y)$  or  $\tilde{v}_y(x, y) = (1 - \mu)\tilde{v}_x(x, y)$  for some  $(x, y)$  then the same is true at all points along the ray through  $(x, y)$ . This strongly suggests that the boundaries between the transaction and no-transaction ( $NT$ ) regions are straight lines through the origin. In the transaction regions, transactions take place at maximum, i.e. infinite, speed, which implies that the investor will make an instantaneous finite transaction to the boundary of  $NT$ . These considerations suggest the picture shown in Fig. 2: the no-transaction region  $NT$  is a wedge, the regions above and below it being the sell ( $S$ ) and buy ( $B$ ) regions respectively. Note that a finite transaction in the  $S$  [ $B$ ] region moves the portfolio down [up] a line of slope  $-1/(1 - \mu)$  [ $-1/(1 + \lambda)$ ].<sup>5</sup> After the initial transaction, all further transactions must take place at the boundaries, and this suggests a “local time” type of transaction policy. Meanwhile, consumption takes place at rate  $v_x^{1/(\gamma-1)}$ . In  $NT$  the value function  $v(x, y)$  satisfies the Bellman equation (3.3) with  $l = u = 0$ :

$$\max_c \left\{ \frac{1}{2} \sigma^2 y^2 v_{yy} + (rx - c)v_x + \alpha y v_y + \frac{1}{\gamma} c^\gamma - \delta v \right\} = 0$$

i.e.

$$(3.4) \quad \frac{1}{2} \sigma^2 y^2 v_{yy} + rx v_x + \alpha y v_y + \left( \frac{1 - \gamma}{\gamma} \right) v_x^{-\gamma/(-\gamma)} - \delta v = 0.$$

<sup>5</sup>The natural ranges of values for  $\mu$  and  $\lambda$  are  $[0, 1]$  and  $[0, \infty]$  respectively, these values corresponding to slopes of the finite transaction lines in  $S$  and  $B$  between  $-45^\circ$  and vertical (in  $S$ ) or horizontal (in  $B$ ).

To substantiate the conjectured solution just outlined we make essential use of the homothetic property of Theorem 3.1(b), by which the nonlinear partial differential equation (3.4) may be reduced to an equation in one variable. Indeed, define  $\psi(x) := v(x, 1)$ . Then by the homothetic property  $v(x, y) = y^\gamma \psi(x/y)$ . If our conjectured optimal policy is correct then  $v$  is constant along lines of slope  $(1 - \mu)^{-1}$  in  $S$  and along lines of slope  $(1 + \lambda)^{-1}$  in  $B$ , and this implies by the homothetic property that

$$(3.5) \quad \begin{aligned} \psi(x) &= \frac{1}{\gamma} A(x + 1 - \mu)^\gamma, & x \leq x_0, \\ \psi(x) &= \frac{1}{\gamma} B(x + 1 + \lambda)^\gamma, & x \geq x_T, \end{aligned}$$

for some constants  $A, B$ , where  $x_0$  and  $x_T$  are as shown in Figure 2. (The factor  $1/\gamma$  in these expressions turns out to be notationally convenient.) Using the homothetic property again we find that, with  $\psi' = d\psi/dx$ ,

$$(3.6) \quad \begin{aligned} v_y(x, 1) &= \gamma\psi(x) - x\psi'(x), & v_x(x, 1) &= \psi'(x), \\ v_{yy}(x, 1) &= -\gamma(1 - \gamma)\psi(x) + 2(1 - \gamma)x\psi'(x) + x^2\psi''(x), \end{aligned}$$

and hence equation (3.4) reduces to

$$(3.7) \quad \begin{aligned} \beta_3 x^2 \psi''(x) + \beta_2 x \psi'(x) + \beta_1 \psi(x) \\ + \left( \frac{1 - \gamma}{\gamma} \right) (\psi')^{-\gamma/(1-\gamma)} = 0, \quad x \in [x_0, x_T], \text{ where} \end{aligned}$$

$$(3.8) \quad \beta_1 = -\frac{1}{2}\sigma^2\gamma(1 - \gamma) + \alpha\gamma - \delta, \quad \beta_2 = \sigma^2(1 - \gamma) + r - \alpha, \quad \beta_3 = \frac{1}{2}\sigma^2.$$

The key to solving this problem is thus to find constants  $x_0, x_T, A, B$  and a globally  $C^2$  function  $\psi$  such that (3.5), (3.7) hold. The definitions of  $\beta_1, \beta_2, \beta_3$  in (3.8) will be maintained throughout the paper.

**4. A sufficiency theorem.** In this section we will show that the existence of a  $C^2$  function  $\psi$  satisfying (3.5) and (3.6) supplies a sufficient condition for optimality of a policy  $(c, L, U)$  such that the corresponding process  $(s_0(t), s_1(t))$  is a reflecting diffusion in the wedge  $NT$  and  $L_t, U_t$  are the local times at the lower and upper boundaries respectively. We first verify in Theorem 4.1 that such reflecting diffusions are well-defined in arbitrary wedges in the positive orthant of  $\mathbb{R}^2$ . We then give the sufficiency theorems, Theorems 4.2 and 4.3. The existence of such a function  $\psi$  is demonstrated in §5.

**THEOREM 4.1.** Take  $0 < x_0 < x_T$  and let  $NT$  be the closed wedge shown in Figure 2, with upper and lower boundaries  $\partial S, \partial B$  respectively. Let  $c: NT \rightarrow \mathbb{R}^+$  be any Lipschitz continuous function and let  $(x, y) \in NT$ . Then there exist unique processes  $s_0, s_1$  and continuous increasing processes  $L, U$  such that for  $t < \tau = \inf\{t: (s_0(t), s_1(t)) = 0\}$

$$(4.1) \quad \begin{aligned} ds_0(t) &= [rs_0(t) - c(s_0(t), s_1(t))] dt \\ &\quad - (1 + \lambda) dL_t + (1 - \mu) dU_t, & s_0(0) &= x, \\ ds_1(t) &= \alpha s_1(t) dt + \sigma s_1(t) dz_t - dU_t, & s_1(0) &= y, \\ L_t &= \int_0^t I_{\{(s_0(\xi), s_1(\xi)) \in \partial B\}} dL_\xi, & U_t &= \int_0^t I_{\{(s_0(\xi), s_1(\xi)) \in \partial S\}} dU_\xi. \end{aligned}$$

The process  $\tilde{c}_t := c(s_0(t), s_1(t))$  satisfies condition (2.1)(i).

The proof will be omitted. Note that the directions of reflection are along the vectors  $((1 - \mu), -1)$  and  $(-(1 + \lambda), 1)$  at the upper and lower boundaries respectively. These coincide with the directions of finite transactions in  $S$  and  $B$ , as shown in Figure 2. The process  $(s_0, s_1)$  is a degenerate diffusion with coefficients which are not bounded away from zero, and with oblique reflection at a nonsmooth boundary. Because of this combination of factors, standard results on existence and uniqueness (e.g. Stroock and Varadhan [19]) do not apply. However, since one is only interested in the solution up to the first hitting time of the corner, the solution can be constructed piecewise, using a sequence of stopping times as in Varadhan and Williams [22], from diffusions reflecting off one or other of the two line boundaries. The technique also appears in Anderson and Orey [1]. The result may also be derived from Tanaka's theory of reflecting diffusions in convex regions [21].

For the sufficiency theorems which follow, we will only consider policies that do not involve shortselling (although borrowing is allowed). This class of policies is defined formally as

$$(4.2) \quad \mathcal{U}' = \{(c, L, U) \in \mathcal{U} : (s_0(t), s_1(t)) \in \mathcal{S}'_{\mu} \text{ for all } t \geq 0\}$$

where  $\mathcal{S}'_{\mu} = \{(x, y) \in \mathbb{R}^2 : y \geq 0 \text{ and } (x + (1 - \mu)y) \geq 0\}$ . The results can be extended to cover policies which allow shortselling, but because of the way we have parametrized the problem a separate argument has to be introduced to show that shortselling is not optimal; it does not seem worth presenting this argument here. For some of the results it is necessary to restrict the admissible policies further to the slightly smaller class  $\mathcal{U}''$  defined as follows. For  $\mu' > \mu$  let

$$\mathcal{U}''(\mu') = \{(c, L, U) \in \mathcal{U} : (s_0(t), s_1(t)) \in \mathcal{S}'_{\mu'} \text{ for all } t \geq 0\}.$$

Now define  $\mathcal{U}'' = \bigcup_{\mu' > \mu} \mathcal{U}''(\mu')$ . Using a policy in  $\mathcal{U}''$  means that the investor is always able to absorb a slight increase in the transaction costs. It seems certain that the results below remain true if  $\mathcal{U}''$  is replaced by  $\mathcal{U}'$ , but our proof technique requires the smaller class.

Theorem 4.2, which we present next, covers the case  $u(c) = c^{\gamma}/\gamma$  for  $\gamma \in \Gamma$ . The corresponding results for  $u(c) = \log c$  are stated separately as Theorem 4.3.

**THEOREM 4.2.** *Take  $\gamma \in \Gamma$  and assume that Condition A holds. Suppose there are constants  $A, B, x_0, x_T$  and a function  $\psi: [-(1 - \mu), \infty) \rightarrow \mathbb{R}$  such that*

$$(4.3) \quad 0 < x_0 < x_T < \infty,$$

$$(4.4) \quad \psi \text{ is } C^2 \text{ and } \psi'(x) > 0 \text{ for all } x,$$

$$(4.5) \quad \psi(x) = \frac{1}{\gamma} A(x + 1 - \mu)^{\gamma} \text{ for } x \leq x_0,$$

$$(4.6) \quad \beta_3 x^2 \psi''(x) + \beta_2 x \psi'(x) + \beta_1 \psi(x) + \left( \frac{1 - \gamma}{\gamma} \right) [\psi'(x)]^{-\gamma/(1-\gamma)} = 0 \text{ for } x \in [x_0, x_T],$$

$$(4.7) \quad \psi(x) = \frac{1}{\gamma} B(x + 1 + \lambda)^{\lambda} \text{ for } x \geq x_T.$$

Let  $NT$  denote the closed wedge  $\{(x, y) \in \mathbb{R}_+^2 : x_T^{-1} \leq yx^{-1} \leq x_0^{-1}\}$  and let  $B$  and  $S$



denote the regions below and above  $NT$  as shown in Figure 2. For  $(x, y) \in NT \setminus \{(0, 0)\}$  define

$$(4.8) \quad c^*(x, y) = y \left[ \psi' \left( \frac{x}{y} \right) \right]^{-1/(1-\gamma)}.$$

Let  $\tilde{c}_t^* = c^*(s_0(t), s_1(t))$  where  $(s_0, s_1, L^*, U^*)$  is the solution of (4.1) with  $c := c^*$ . Then the policy  $(\tilde{c}^*(t), L^*(t), U^*(t))$  is optimal in the class  $\mathcal{U}'$  for any initial endowment  $(x, y) \in NT$  when  $\gamma \in (0, 1)$ . When  $\gamma < 0$  this policy is optimal in the class  $\mathcal{U}''$ . In either case, if  $(x, y) \notin NT$  then an immediate transaction to the closest point in  $NT$  followed by application of this policy is optimal in  $\mathcal{U}', \mathcal{U}''$  respectively. The maximal expected utility is

$$(4.9) \quad v(x, y) = y^\gamma \psi \left( \frac{x}{y} \right).$$

For the proof, we require the following lemma.

LEMMA 4.3. Suppose  $v$  is defined as in (4.9). Then

(i)  $v$  is concave.

$$(ii) \quad v(x, y) = \begin{cases} \frac{1}{\gamma} A(x + (1 - \mu)y)^\gamma & \text{in } S, \\ \frac{1}{\gamma} B(x + (1 + \lambda)y)^\gamma & \text{in } B. \end{cases}$$

$$(iii) \quad (1 - \mu)v_x - v_y \leq 0 \quad \text{with equality in } S,$$

$$-(1 + \lambda)v_x + v_y \leq 0 \quad \text{with equality in } B.$$

(iv) Define

$$(4.10) \quad Gv = \frac{1}{2}\sigma^2 y^2 v_{yy} + rxv_x + \alpha yv_y - \delta v + \frac{1-\gamma}{\gamma} (v_x)^{-\gamma/(1-\gamma)}.$$

Then

$$(4.11) \quad Gv = \max_c \left\{ \frac{1}{2}\sigma^2 y^2 v_{yy} + (rx - c)v_x + \alpha yv_y + \frac{1}{\gamma} c^\gamma - \delta v \right\} \quad \text{and}$$

$$(4.12) \quad Gv = 0 \text{ in } NT \quad (\text{i.e. the Bellman equation is satisfied}),$$

$$Gv \leq 0 \quad \text{in } S \cup B.$$

PROOF. These properties can be verified directly from the construction of  $v$ .

PROOF OF THEOREM 4.2. Let  $(c, L, U)$  be a policy in  $\mathcal{U}'$  and let  $(s_0(t), s_1(t))$  be the corresponding solution of (3.1) with initial point  $(x, y) \in \mathcal{S}_\mu'$ . For an arbitrary  $C^2$  function  $\chi$  let  $M_T^\chi$  be the scalar process defined for  $T \geq 0$  by

$$M_T^\chi := \int_0^T e^{-\delta t} \frac{1}{\gamma} c^\gamma(t) dt + e^{-\delta T} \chi(s_0(T), s_1(T)).$$

An application of the generalized Ito formula (Harrison [9, §4.7]) gives

$$\begin{aligned}
 (4.13) \quad M_T^X - M_0^X &= \int_0^T e^{-\delta t} \left[ \frac{1}{\gamma} c^\gamma(t) + (rs_0(t) - c(t))\chi_x + \alpha s_1(t)\chi_y \right. \\
 &\quad \left. + \frac{1}{2} \sigma^2 s_1^2(t) \chi_{yy} - \delta \chi \right] dt \\
 &\quad + \int_0^T e^{-\delta t} [-(1+\lambda)\chi_x + \chi_y] dL_t \\
 &\quad + \int_0^T e^{-\delta t} [(1-\mu)\chi_x - \chi_y] dU_T \\
 &\quad + \sum_{0 \leq t \leq T} e^{-\delta t} [\chi(s_0(t), s_1(t)) - \chi(s_0(t^-), s_1(t^-)) \\
 &\quad - \chi_x \Delta s_0(t) - \chi_y \Delta s_1(t)] \\
 &\quad + \int_0^T e^{-\delta t} \sigma \chi_{y s_1}(t) dz_t \\
 &=: I_1 + I_2 + I_3 + I_4 + I_5.
 \end{aligned}$$

In this equation,  $\chi, \chi_x$  etc. are evaluated at  $(s_0(t), s_1(t))$  unless noted otherwise.

Suppose first that  $(c, L, U) = (c^*, L^*, U^*)$  as defined in the theorem statement and  $(x, y) \in NT$ , and let  $v$  be given by (4.9). The value of  $c^*$  is equal to  $[v_x]^{-1/(1-\gamma)}$ , the maximizing value obtained from the Bellman equation. Note from (4.8) that  $c^*(\rho x, \rho y) = \rho c^*(x, y)$  for  $\rho > 0$ , and hence that  $c_x^*(\rho x, \rho y) = c_x^*(x, y)$ ,  $c_y^*(\rho x, \rho y) = c_y^*(x, y)$ . It follows that  $c^*$  is Lipschitz continuous and that (4.1) has a unique solution  $(s_0(t), s_1(t), L_t^*, U_t^*)$  when  $c$  is equal to  $c^*$ . Now consider (4.13) with  $(c, L, U) = (c^*, L^*, U^*)$  and  $\chi = v$ . It follows from Lemma 4.3 that  $I_1 = I_2 = I_3 = 0$ , and  $I_4$  vanishes since  $(s_0(t), s_1(t))$  is continuous. Hence

$$\begin{aligned}
 (4.14) \quad v(x, y) &= M_0^v = \int_0^T e^{-\delta t} \frac{1}{\gamma} c^\gamma(t) dt + e^{-\delta T} v(s_0(T), s_1(T)) \\
 &\quad - \int_0^T e^{-\delta t} \sigma v_{y s_1}(t) dz_t.
 \end{aligned}$$

It is shown in Theorem 6.2 below that under  $(c^*, L^*, U^*)$  the value process  $v(s_0(t), s_1(t))$  can be represented in the form

$$(4.15) \quad v(s_0(t), s_1(t)) = v(x, y) \exp \left( \int_0^t \left( \delta - g \left( \frac{s_0(u)}{s_1(u)} \right) \right) du \right) \tilde{G}(t)$$

where  $\tilde{G}(t)$  is the Girsanov exponential

$$\tilde{G}(t) = \exp \left( \int_0^t \gamma \sigma (1-f) dz - \frac{1}{2} \int_0^t \gamma^2 \sigma^2 (1-f)^2 du \right)$$

and  $f$  and  $g$  are bounded functions in  $NT$  with  $g > 0$ . It is also shown in Lemma 6.1

that in  $NT$

$$(4.16) \quad yv_y(x, y) = \gamma \left( 1 - f\left(\frac{x}{y}\right) \right) v(x, y).$$

Since  $f$  is bounded,  $\mathbb{E}_{x,y} \tilde{G}^2(t) < \infty$  and it follows from (4.15) that the last term in (4.14) is a martingale. Thus

$$v(x, y) = \mathbb{E}_{x,y} \int_0^T e^{-\delta t} \frac{1}{\gamma} c^\gamma(t) dt + e^{-\delta T} \mathbb{E}_{x,y} [v(s_0(T), s_1(T))].$$

As  $T \rightarrow \infty$ , the last term converges to zero in view of (4.15) and this shows that, for  $(x, y) \in NT$ ,

$$(4.17) \quad v(x, y) = J_{x,y}(c^*, L^*, U^*).$$

For  $(x, y) \in \mathcal{S}'_\mu \setminus NT$  it is clear that  $J_{x,y}(c^*, L^*, U^*) = J_{x',y'}(c^*, L^*, U^*)$  where  $(x', y')$  is the point on the boundary of the wedge to which the initial transaction is made. Since  $v$  is by construction constant on the line joining  $(x, y)$  to  $(x', y')$ , it follows that (4.17) holds throughout  $\mathcal{S}'_\mu$ .

We now show that  $v(x, y) \geq J_{x,y}(c, L, U)$  for arbitrary  $(c, L, U)$ . We need separate arguments for the two cases  $0 < \gamma < 1$  and  $\gamma < 0$ .

Case (a):  $0 < \gamma < 1$ . Here we need the following lemma, whose proof is given later.

LEMMA 4.4. *Let  $v$  be defined by (4.9) with  $\gamma \in (0, 1)$ . Then there is a constant  $K$  and for each  $\epsilon > 0$  a constant  $K_\epsilon$  such that*

$$(4.18) \quad 0 \leq v(x, y) \leq K(x + y)^\gamma \quad \text{for all } (x, y) \in \mathcal{S}'_\mu,$$

$$(4.19) \quad 0 \leq yv(x, y) \leq K_\epsilon(1 + x + y)$$

$$\text{for all } (x, y) \in \mathbb{R}_+^2 \cup \{(x, y) \in \mathcal{S}'_\mu : x + (1 - \mu)y > \epsilon\}.$$

Let  $(c, L, U) \in \mathcal{U}'$  be an arbitrary policy and  $(s_0(t), s_1(t))$  the corresponding solution of (3.1). If we define  $\hat{w}_t := s_0(t) + s_1(t)$  and  $\pi_t = s_1(t)/\hat{w}_t$  then  $(c_t, \pi_t)$  is an admissible policy for the Merton (no transaction costs) problem. From (3.1) we see that  $\hat{w}_t$  satisfies

$$d\hat{w}_t = [(r + (\alpha - r)\pi_t)\hat{w}_t - c_t] dt + \sigma\hat{w}_t\pi_t dz_t - dA_t$$

where  $A = \lambda L + \mu U$ . By the Gronwall-Bellman lemma we conclude that  $\hat{w}_t \leq w_t$  for all  $t$ , where  $w_t$  is the solution of the Merton wealth equation (2.1) with  $w_0 = x + y$ . Fix  $\epsilon > 0$ , define  $v^\epsilon(x, y) = v(x + \epsilon, y)$ , and consider equation (4.13) with  $\chi := v^\epsilon$ . From (4.19) we see that for any  $T > 0$ ,

$$\begin{aligned} \mathbb{E}_{x,y} \int_0^T (v_y^\epsilon(s_0(t), s_1(t))s_1(t))^2 dt &\leq K_\epsilon^2 \int_0^T (1 + \epsilon + \hat{w}_t)^2 dt \\ &\leq K_\epsilon^2 \mathbb{E}_{x,y} \int_0^T (1 + \epsilon + w_t)^2 dt. \end{aligned}$$

The last expression is finite as was shown in the proof of Theorem 2.1. Thus, the last term  $I_5$  in (4.13) is a martingale. From part (iv) of Lemma 4.3 and since  $v_x^\epsilon(x, y) =$

$v_x(x + \epsilon, y)$  etc. we find that

$$Gv^\epsilon(x, y) = Gv(x + \epsilon, y) - r\epsilon v_x(x + \epsilon, y).$$

In view of (4.6), (4.11) and (4.12) we see that for any  $c \geq 0$  and  $(x, y) \in \mathcal{J}_\mu'$ ,

$$\frac{1}{\gamma}c^\gamma + (rx - c)v_x^\epsilon + \alpha y v_y^\epsilon + \frac{1}{2}\alpha^2 y^2 v_{yy}^\epsilon - \delta v \leq 0.$$

Hence  $I_1$  is a decreasing process. It follows from Lemma 4.3(iii) that  $I_2$  and  $I_3$  are decreasing while Lemma 4.3(i) implies that  $I_4$  is decreasing. Thus  $M_T^{v^\epsilon}$  is a supermartingale, i.e.

$$(4.20) \quad v^\epsilon(x, y) = M_0^{v^\epsilon} \geq \mathbb{E}_{x,y} \int_0^T e^{-\delta t} \frac{1}{\gamma} c^\gamma(t) dt + e^{-\delta T} \mathbb{E}_{x,y} v^\epsilon(s_0(T), s_1(T)).$$

Now using (4.18) we have

$$\begin{aligned} e^{-\delta T} \mathbb{E}_{x,y} v^\epsilon(s_0(T), s_1(T)) &\leq K e^{-\delta T} \mathbb{E}_{x,y} (\epsilon + \hat{w}_T)^\gamma \\ &\leq K e^{-\delta T} \mathbb{E}_{x,y} (\epsilon + w_T)^\gamma \\ &\leq K e^{-\delta T} \mathbb{E}_{x,y} (K' + w_T^{\gamma'}) \end{aligned}$$

for some  $K' > 0$ ,  $\gamma' \in (\gamma, 1)$ . It follows from (2.12), (2.13) that

$$\mathbb{E}_{x,y} [e^{-\delta T} w_T^{\gamma'}] \rightarrow 0 \quad \text{as } T \rightarrow \infty.$$

Thus taking the limit in (4.20) as  $T \rightarrow \infty$  we obtain  $v^\epsilon(x, y) \geq J_{x,y}(c, L, U)$ . Now  $v^\epsilon(x, y) \downarrow v(x, y)$  as  $\epsilon \downarrow 0$ , so that  $v(x, y) \geq J_{x,y}(c, L, U)$ . Thus  $(c^*, L^*, U^*)$  is optimal in the case  $\gamma \in (0, 1)$ .

Case (b):  $\gamma < 0$ . The problem here is that the candidate value function  $v(x, y)$  is unboundedly negative at low levels of wealth. We use an argument similar to the proof of Theorem 2.1, based on the following lemma, whose proof is again given at the end of this section.

LEMMA 4.5. Let  $v$  be the function defined by (4.9) with  $\gamma < 0$ , and for  $\theta, \epsilon > 0$  define  $v^{\epsilon, \theta}(x, y) = v(x + \theta\epsilon, y + \epsilon)$ . Fix  $\mu' > \mu$ . Then there exists  $\theta > 0$  such that for all  $\epsilon > 0$

$$Gv^{\epsilon, \theta}(x, y) \leq 0 \quad \text{for all } (x, y) \in \mathcal{J}_{\mu'}',$$

where  $G$  is defined by (4.10).

We note from Lemma 4.3(iv) that this result implies that for all  $c \geq 0$ ,

$$(4.21) \quad \frac{1}{2}\sigma^2 y^2 v_{yy}^{\epsilon, \theta} + (rx - c)v_x^{\epsilon, \theta} + \alpha y v_y^\epsilon + \frac{1}{\gamma}c^\gamma - \delta v^{\epsilon, \theta} \leq 0, \quad (x, y) \in \mathcal{J}_{\mu'}'.$$

We also note that for any  $\epsilon, \theta > 0$ ,  $v^{\epsilon, \theta}$  is bounded on  $\mathcal{J}_{\mu'}'$  with bounded first and second partial derivatives. Now let  $(c, L, U)$  be an arbitrary policy in  $\Pi''$ ; then  $(c, L, U) \in \Pi''(\mu')$  for some  $\mu' > \mu$ . Consider the process  $M_T^\chi$  of (4.13) with  $\chi = v^{\epsilon, \theta}$ . We conclude from (4.21) and Lemma 4.3 as before that  $M_T^{v^{\epsilon, \theta}}$  is a supermartingale

and hence that

$$v^\epsilon(x, y) \geq \mathbb{E}_{x,y} \left[ \int_0^T e^{-\delta t} \frac{1}{\gamma} c^\gamma(t) dt \right] + \mathbb{E}_{x,y} \left[ e^{-\delta T} v^{\epsilon, \theta}(s_0(T), s_1(T)) \right].$$

The second term on the right converges to 0 as  $T \rightarrow \infty$ , while the bracket  $[\dots]$  in the first term is monotone decreasing. It follows that

$$v^{\epsilon, \theta}(x, y) \geq \mathbb{E}_{x,y} \left[ \int_0^\infty e^{-\delta t} \frac{1}{\gamma} c^\gamma(t) dt \right] = J_{x,y}(c, L, U).$$

Now  $v^{\epsilon, \theta}(x, y) \downarrow v(x, y)$  as  $\epsilon \downarrow 0$  and hence

$$v(x, y) \geq \sup_{(c, L, U) \in \mathcal{U}''} J_{x,y}(c, L, U).$$

On the other hand we know that  $v(x, y) = J_{x,y}(c^*, L^*, U^*)$  and it is clear that  $(c^*, L^*, U^*) \in \mathcal{U}''$  because the solution process is trapped inside the wedge  $NT$  after any initial transaction. Thus  $(c^*, L^*, U^*)$  is optimal in  $\mathcal{U}''$ , as claimed.

**PROOF OF LEMMA 4.4.** Denote  $S_+ = S \cap \mathbb{R}_+^2$  and  $S_- = S \setminus S_+$ , with similar definitions for  $B_+, B_-$  (note that  $B_+ = B \cap \mathcal{S}'_\mu$ ). Elementary geometry shows that

$$(4.22) \quad y \leq \frac{1}{\mu}(x + y) \quad \text{for all } (x, y) \in \mathcal{S}'_\mu.$$

Using this, (4.18) is readily verified in  $S \cup B_+$ , while in  $NT$ ,

$$v(x, y) = y^\gamma \psi\left(\frac{x}{y}\right) \leq M_0(x + y)^\gamma$$

where  $M_0 = \mu^{-\gamma} \max_{\xi \in [x_0, x_T]} \psi(\xi)$ . Thus (4.18) holds throughout  $S'_\mu$ . In view of (4.9), in  $NT$

$$(4.23) \quad yv_y(x, y) = \gamma y^\gamma \psi\left(\frac{x}{y}\right) - xy^{\gamma-1} \psi'\left(\frac{x}{y}\right) \leq \gamma M_0(x + y)^\gamma.$$

In  $B_+$ ,

$$(4.24) \quad yv_y(x, y) = By(x + (1 + \lambda)y)^{\gamma-1} \leq B(x + y)^\gamma,$$

and similarly in  $S_+$ ,

$$(4.25) \quad yv_y(x, y) \leq A(1 - \mu)^{\gamma-1}(x + y)^\gamma.$$

In  $S_-$  we note from (4.22) that when  $(x + (1 - \mu)y) > \epsilon$ ,

$$(4.26) \quad yv_y(x, y) = Ay(x + (1 - \mu)y)^{\gamma-1} \leq \frac{A}{\mu\epsilon^{1-\gamma}}(x + y)$$

We can now choose  $K_\epsilon$  such that  $K_\epsilon(1 + x + y)$  majorizes the right-hand sides of (4.23)–(4.26), thus establishing (4.19).

PROOF OF LEMMA 4.5. Take  $\epsilon, \theta > 0$ . We know from Lemma 4.3(iv) that

$$\begin{aligned} 0 &\geq Gv(x + \theta\epsilon, y + \epsilon) \\ &= \frac{1}{2}\sigma^2(y + \epsilon)^2 v_{yy}(x + \theta\epsilon, y + \epsilon) + r(x + \theta\epsilon)v_x(x + \theta\epsilon, y + \epsilon) \\ &\quad + \alpha(y + \epsilon)v_y(x + \theta\epsilon, y + \epsilon) + \frac{1-\gamma}{\gamma}(v_x)^{-\gamma/(1-\gamma)} - \delta v \\ &= Gv^{\epsilon, \theta}(x, y) + \epsilon\rho(\epsilon) \quad \text{where} \quad \rho(\epsilon) := \frac{1}{2}\sigma^2(2y + \epsilon)v_{yy}^{\epsilon, \theta} + r\theta v_x^{\epsilon, \theta} + \alpha v_y^{\epsilon, \theta}. \end{aligned}$$

We need to show that  $\rho(\epsilon) \geq 0$  so that then  $Gv^{\epsilon, \theta} \leq 0$ . First, consider points  $(x, y)$  such that  $(x + \theta\epsilon, y + \epsilon) \in NT$ . Then, with  $\xi := (x + \theta\epsilon)/(y + \epsilon)$ , we have from (4.10)

$$\begin{aligned} v^{\epsilon, \theta}(x, y) &= (y + \epsilon)^\gamma \psi(\xi), \\ v_x^{\epsilon, \theta}(x, y) &= (y + \epsilon)^{\gamma-1} \psi'(\xi), \\ v_y^{\epsilon, \theta}(x, y) &= (y + \epsilon)^{\gamma-1} (\gamma\psi(\xi) - \xi\psi'(\xi)), \\ v_{yy}^{\epsilon, \theta}(x, y) &= (y + \epsilon)^{\gamma-2} (\xi^2\psi'' + 2(1-\gamma)\xi\psi' - \gamma(1-\gamma)\psi) \end{aligned}$$

and hence

$$\begin{aligned} (4.27) \quad (y + \epsilon)^{2-\gamma} \rho(\epsilon) &= r\theta\psi' + \frac{1}{2}\sigma^2\eta\xi^2\psi'' + [\sigma^2\eta(1-\gamma) - \alpha]\xi\psi' \\ &\quad + [\alpha\gamma - \frac{1}{2}\sigma^2\eta\gamma(1-\gamma)]\psi \end{aligned}$$

where  $\eta := (2y + \epsilon)/(y + \epsilon)$ . Now  $\xi \in [x_0, x_T]$  and  $\eta \in [1, 2]$ , so that all the terms on the right of (4.27) are bounded. Since  $\psi' > 0$  it follows that  $\rho(\epsilon) > 0$  if  $\theta$  is chosen sufficiently large. Note also that the minimum value of  $\theta$  required does not depend on  $\epsilon$ .

For  $(x, y) \in S$  we have  $v(x, y) = A(x + (1 - \mu)y)^\gamma/\gamma$  and hence

$$\rho(\epsilon) = Ab^{\gamma-1} \left\{ \frac{1}{2}\sigma^2(\gamma-1)(1-\mu)^2 \frac{2y+\epsilon}{b} + r\theta + \alpha(1-\mu) \right\}$$

where  $b := x + (1 - \mu)y + (\theta + 1 - \mu)\epsilon$ . In  $S_+$  (i.e. when  $x \geq 0$ ) we have

$$\frac{2y+\epsilon}{b} < \frac{2y+\epsilon}{(1-\mu)(y+\epsilon)} \leq \frac{2}{1-\mu},$$

so  $\rho(\epsilon) \geq 0$  as long as

$$\theta > \frac{1-\mu}{\sigma} (\sigma^2(1-\gamma) - \alpha).$$

A similar calculation applies in  $B_+$ . The remaining case is  $S'_- = S'_\mu \cap \{x < 0\}$ . Here



we have  $x + (1 - \mu')y \geq 0$  for some  $\mu' > \mu$ , so that

$$\begin{aligned} \frac{2y + \epsilon}{b} &= \frac{2y + \epsilon}{x + (1 - \mu')y + (\mu' - \mu)y + (\theta + 1 - \mu)\epsilon} \\ &\leq \frac{2y + \epsilon}{(\mu' - \mu)y + (1 - \mu)\epsilon} \\ &= \frac{2y + \epsilon}{(\mu' - \mu)(y + \epsilon) + (1 - \mu')\epsilon} \leq \frac{2}{\mu' - \mu}. \end{aligned}$$

Thus  $\rho(\epsilon) > 0$  if

$$\theta > \frac{1 - \mu}{\sigma} \left( \sigma^2(1 - \gamma) \frac{(1 - \mu)}{\mu' - \mu} - \alpha \right),$$

the right-hand side being independent of  $\epsilon$ . Thus for  $\theta > \theta_0$ , where  $\theta_0$  is a constant not depending on  $\epsilon$ , we have  $\rho(\epsilon) > 0$  throughout  $S'_\mu$ . This completes the proof.

The following theorem covers the case  $u(c) = \log c$ . Its proof is similar to that of Theorem 4.2 and is omitted.

**THEOREM 4.3.** Assume  $\delta > 0$ , let  $\beta_1, \beta_2, \beta_3$  be defined as in (3.8) with  $\gamma = 0$ , and define  $\beta_4 := (-\frac{1}{2}\sigma^2 + \alpha - \delta)/\delta$ . Suppose that there are constants  $x_0, x_T, A, B$  and a function  $\psi: [-(1 - \mu), \infty) \rightarrow \mathbb{R}$  such that

$$0 < x_0 < x_T < \infty,$$

$$\psi \text{ is } C^2 \text{ and } \psi'(x) > 0 \text{ for all } x,$$

$$\psi(x) = \frac{1}{\delta} \log[A(x + (1 - \mu))] \text{ for } x \leq x_0,$$

$$\beta_1\psi(x) + \beta_2x\psi'(x) + \beta_3x^2\psi''(x) + \beta_4 - \log\psi'(x) = 0 \text{ for } x \in [x_0, x_T],$$

$$\psi(x) = \frac{1}{\delta} \log[B(x + 1 + \lambda)] \text{ for } x \geq x_T.$$

Let regions  $NT, B, S$  be defined as before and for  $(x, y) \in NT \setminus \{(0, 0)\}$  define  $c^*(x, y) = y[\psi'(x/y)]^{-1}$ . Then with the utility function  $u(c) = \log c$ , the policy  $(\bar{c}^*(t) := c^*(s_0(t), s_1(t)), L_t^*, U_t^*)$ , where  $(s_0, s_1, L^*, U^*)$  is the solution of (4.1) with  $c := c^*$ , is optimal in the class  $\mathcal{U}^n$  for any initial endowment  $(x, y) \in NT$ . If  $(x, y) \notin NT$  then an initial transaction to the nearest point in  $NT$  followed by application of this policy is optimal. The maximum expected utility is

$$(4.28) \quad v(x, y) = \frac{1}{\delta} \log y + \psi\left(\frac{x}{y}\right).$$

The existence of functions  $\psi$  satisfying the conditions of Theorems 4.2 and 4.3 is established in §5 below when  $\alpha > r$ . When  $\alpha = r$ , it is intuitive that we would invest only in the bank. Our final result in this section demonstrates that this is so.

**THEOREM 4.4.** Suppose  $\alpha = r$  and  $\delta > \gamma r$ . Then the optimal strategy in  $\mathcal{U}$  is to close out any position in stock and to consume optimally from bank.

PROOF. [sketch] For  $\gamma \in \Gamma$  and  $(x, y) \in \mathcal{S}_{\mu, \lambda}$  define  $w_1(x, y) = K(x + (1 - \mu)y)^\gamma$  and  $w_2(x, y) = K(x + (1 + \lambda)y)^\gamma$  where  $K := ((\delta - \gamma r)/(1 - \gamma))^{\gamma-1} \gamma^{-1}$ . In view of Remark 2.2,  $w(x, y) := w_1(x, y) \wedge w_2(x, y)$  is the utility of the policy described in the theorem statement. Taking an arbitrary policy  $(c, L, U) \in \mathcal{U}$  and applying a supermartingale argument as in the proof of Theorem 4.2 successively to the functions  $w_1, w_2$ , we find that  $J_{x,y}(c, L, U) \leq w_i(x, y)$ ,  $i = 1, 2$  and hence that  $J_{x,y}(c, L, U) \leq w(x, y)$ . This approach sidesteps the nonsmoothness of  $w$  at  $y = 0$ . A similar argument handles the case  $u(c) = \log c$ , using the functions  $w_1(x, y) = K' + \log[\delta(x + (1 - \mu)y)]/\delta$  and  $w_2 = K' + \log[\delta(x + (1 + \lambda)y)]/\delta$ , where  $K' = (r - \delta)/\delta^2$ .

**5. Solution of the free boundary problem.** In this section we give our main result, Theorem 5.1, concerning the existence of a solution to the free boundary problem. To prove this, we have to establish certain properties of the system of ordinary differential equations (5.11) below; this is done in the Appendix. We find that a technical condition, introduced in the Appendix as Condition B, is required. We are virtually certain that this condition is nugatory, i.e. is always satisfied, but we can prove this only in some cases. Further comments are given in the Appendix.

**THEOREM 5.1.** *Suppose that Conditions A and B hold, that  $\mu \in [0, 1]$ ,  $\lambda \in [0, \infty]$ ,  $\mu \vee \lambda > 0$ , and that*

$$(5.1) \quad 0 < r < \alpha < r + (1 - \gamma)\sigma^2.$$

*Then there is a  $C^2$ -function  $\psi: (-(1 - \mu), \infty) \rightarrow \mathbb{R}$  and there are positive constants  $x_0, x_T, A, B$  with  $0 < x_0 < x_T < \infty$  satisfying:*

(a) *in case  $\gamma \in \Gamma$ :*

$$(5.2) \quad \psi(x) = \frac{1}{\gamma} A(x + 1 - \mu)^\gamma, \quad x \leq x_0,$$

$$(5.3) \quad \beta_1 \psi(x) + \beta_2 x \psi'(x) + \beta_3 x^2 \psi''(x) + \left( \frac{1 - \gamma}{\gamma} \right) [\psi'(x)]^{-\gamma/(1-\gamma)} = 0, \\ x \in [x_0, x_T],$$

$$(5.4) \quad \psi(x) = \frac{1}{\gamma} B(x + 1 + \lambda)^\gamma, \quad x \geq x_T,$$

$$(5.5) \quad \gamma \psi(x) \psi''(x) + (1 - \gamma)(\psi'(x))^2 \leq 0 \quad \text{with equality when } x \leq x_0 \text{ or } x \geq x_T.$$

(b) *in case  $\gamma = 0$ :*

$$(5.6) \quad \psi(x) = \frac{1}{\delta} \log[A(x + 1 - \mu)], \quad x \leq x_0,$$

$$(5.7) \quad \beta_1 \psi(x) + \beta_2 x \psi'(x) + \beta_3 x^2 \psi''(x) + \beta_4 - \log \psi'(x) = 0, \\ x \in [x_0, x_T],$$

$$(5.8) \quad \psi(x) = \frac{1}{\delta} \log[B(x + 1 + \lambda)], \quad x \geq x_T,$$

$$(5.9) \quad \frac{1}{\delta} \psi(x) + \psi'^2(x) \leq 0 \quad \text{with equality for } x \leq x_0, x \geq x_T.$$

**REMARK 5.2.** Properties (5.5), (5.9) will be needed to prove that the functions  $v$  constructed in (4.9), (4.28) are concave. (5.1) is the same as the condition which ensures, in the Merton problem, that hedging, rather than leverage or shortselling, is optimal. See Remark 2.2.

Before proceeding to the proof, let us point out conceptually what is involved in solving the free boundary problem. At a given point  $x_0$ , (5.2) and (5.3) provide two different expressions for  $\psi''$ . Equating these, we find that there is a unique value of  $A$ , say  $A(x_0)$ , such that  $\psi''$  is continuous at  $x_0$ . Similarly, continuity of  $\psi''$  at  $x_T$  fixes  $B(x_T)$ . Now write  $\zeta(x) = (\zeta^1(x), \zeta^2(x)) = (\psi(x), \psi'(x))$  and write the scalar second-order equation (5.3) as a 2-vector first-order equation in the form

$$(5.10) \quad \frac{d}{dx} \zeta(x) = f(x, \zeta(x)).$$

This can be solved, at least locally, from any initial vector  $\zeta_0$  at  $x = x_0$ ; denote this solution  $\zeta(x; \zeta_0, x_0)$ . Now (5.2) implies that the initial vector  $\zeta_0$  is fixed once  $x_0$  is given; indeed

$$\zeta_0 = \zeta_0(x_0) = \begin{bmatrix} \psi(x_0) \\ \psi'(x_0) \end{bmatrix} = \begin{bmatrix} \frac{1}{\gamma} A(x_0)(x_0 + 1 - \mu)^\gamma \\ A(x_0)(x_0 + 1 - \mu)^{\gamma-1} \end{bmatrix}.$$

Thus for given  $x_0$  the solution of (5.3) is  $\zeta(x; \zeta_0(x_0), x_0)$ . Similarly we see that (5.4) specifies the value  $\zeta_T(x_T)$  of the terminal vector  $\zeta_T$ . Thus the free boundary problem is solved if we can find  $x_0, x_T$  such that  $\zeta_T(x_T) = \zeta(x_T; \zeta_0(x_0), x_0)$ . Generically, there is no reason why this equation should be solvable, but the proof below shows that in fact such  $x_0, x_T$  exist under the specified conditions. As will be seen, the key to the proof is the introduction of nonlinear coordinate changes under which (5.10) takes a simpler form.

**PROOF OF THEOREM 5.1.** (a) For  $f \in [0, 1]$  define quadratic functions  $Q(f), R(f)$  by:

$$Q(f) := -\frac{1}{\gamma} \beta_1 - \beta_2 f + \beta_3(1 - \gamma)f^2,$$

$$R(f) := Q(f) + \beta_3(1 - f)f = -\frac{1}{\gamma} \beta_1 + (\beta_3 - \beta_2)f - \gamma\beta_3 f^2.$$

Notice that  $Q(0) = R(0)$ ,  $Q(1) = R(1)$ . Now consider the following one-parameter family of initial value problems  $\{(x_0, f(\cdot), h(\cdot)); 0 < x_0 < (1 - \mu)\beta_2/(\alpha - r)\}$ , where  $x_0$  is the starting value:

$$(5.11) \quad f'(x) = \frac{1}{\beta_3 x} [R(f(x)) - h(x)],$$

$$h'(x) = \frac{\gamma}{1 - \gamma} \frac{h(x)}{\beta_3 x f(x)} [h(x) - Q(f(x))], \quad x \geq x_0,$$

$$f(x_0) = f_0 = \frac{x_0}{x_0 + 1 - \mu}, \quad h(x_0) = h_0 = Q(f_0).$$

REMARK 5.3. The range of  $x_0$  implies that  $0 < f_0 < f_m = 1 - (\alpha - r)/(1 - \gamma)\sigma^2$ . Here  $f_m$  is the value of  $f$  at which  $Q(f)$  achieves its minimum. Condition (5.1) ensures that  $f_m \in (0, 1)$  and we find that

$$Q(f_m) = \frac{\delta}{\gamma} - r - \frac{1}{2(1-\gamma)} \left( \frac{\alpha - r}{\sigma} \right)^2$$

so that Condition A ensures that  $\gamma Q(f_m) > 0$ .

It is shown in the Appendix that there exist  $x_0$  and  $x_T$  and a solution  $(f(\cdot), h(\cdot))$  to (5.11) such that

- (5.12) (i)  $0 < x_0 < x_T$ ,
- (ii)  $f_0 = \frac{x_0}{x_0 + 1 - \mu}$ ,  $f(x_T) = \frac{x_T}{x_T + 1 + \lambda} =: f_T$ , say.
- (iii)  $h'(x_0) = h'(x_T) = 0$ ,  $h(x_0) = Q(f_0)$ ,  $h(x_T) = Q(f_T)$ ,
- (iv)  $h(\cdot), f(\cdot)$  are increasing function on  $(x_0, x_T)$  with  $\gamma h > 0$ .
- (v)  $f(x) \in (0, 1)$  for  $x \in (x_0, x_T)$ .

It is also shown that  $f_0$  decreases and  $f_T$  increases with respect to  $\lambda$  for fixed  $\mu$ . Next, define for  $x \in [x_0, x_T]$

$$p(x) := \left[ \frac{\gamma h(x)}{1 - \gamma} \right]^{-(1-\gamma)/\gamma}, \quad q(x) := \frac{x(1 - f(x))}{f(x)}.$$

In view of (5.12)(iv) above,  $p(x)$  is decreasing on  $[x_0, x_T]$ . Elementary computations show that

$$p'(x) = -\frac{1}{\beta_3 x f(x)} p(x) [h(x) - Q(f(x))],$$

$$q'(x) = \frac{1}{\beta_3 f^2(x)} \left[ \left( \frac{1-\gamma}{\gamma} \right) (p(x))^{-\gamma/(1-\gamma)} - Q(f(x)) \right].$$

From (ii) above, we have that

$$(5.13) \quad q(x_0) = 1 - \mu, \quad q(x_T) = 1 + \lambda.$$

Notice also that  $p'/p = -fq'/x = -q'/x + q$ ; this implies that  $q$  is increasing on  $[x_0, x_T]$ . Now define

$$p(x) [q(x)] := \begin{cases} p(x_0) [q(x_0)], & x < x_0, \\ p(x_T) [q(x_T)], & x > x_T, \end{cases} \quad \text{and}$$

$$(5.14) \quad \psi(x) := \frac{1}{\gamma} p^\gamma(x) [x + q(x)]^\gamma, \quad x \geq -(1 - \mu).$$

For  $x \in [x_0, x_T]$  we find, by using the above formulas for  $p'(x)$  and  $q'(x)$  and the

relationship  $x + q(x) = x/f(x)$ , that

$$(5.15) \quad \psi'(x) = \frac{1}{x} \gamma \psi(x) f(x)$$

(and hence that  $\psi'(x) > 0$  as (4.4) requires), and

$$\begin{aligned} \psi''(x) &= \frac{\gamma \psi' f}{x} + \frac{\gamma \psi f'}{x} - \frac{\gamma \psi f}{x^2} \\ &= \frac{\gamma^2 \psi f^2}{x^2} + \frac{\gamma \psi}{x} \frac{1}{\beta_3 x} [R - h] - \frac{\gamma \psi f}{x^2}. \end{aligned}$$

Also,

$$\begin{aligned} \left( \frac{1-\gamma}{\gamma} \right) [\psi']^{-\gamma/(1-\gamma)} &= p^{-\gamma^2/(1-\gamma)} [x + q]^\gamma \left( \frac{1-\gamma}{\gamma} \right) = \gamma \psi p^{-\gamma/(1-\gamma)} \left( \frac{1-\gamma}{\gamma} \right) \\ &= (1-\gamma) \psi p^{-\gamma/(1-\gamma)} \\ &= (1-\gamma) \frac{\psi \gamma}{1-\gamma} h = \gamma \psi h. \end{aligned}$$

Using these formulae it is easily verified that  $\psi$  satisfies (5.3). Putting  $A = p^\gamma(x_0)$ ,  $B = p^\gamma(x_T)$ , we note that  $\psi(\cdot)$  has the required form outside  $[x_0, x_T]$ . In view of (5.12)(iii), these continuations are in  $C^1$ , at least. Now, for  $x$  in  $[x_0, x_T]$

$$\begin{aligned} \gamma \psi \psi'' + (1-\gamma) \psi'^2 &= \gamma^3 \frac{\psi^2 f^2}{x^2} + \frac{\gamma^2 \psi^2}{\beta_3 x^2} [R - h] - \frac{\gamma^2 \psi^2 f}{x^2} + (1-\gamma) \frac{\gamma^2 \psi^2 f^2}{x^2} \\ &= \frac{\gamma^2 \psi^2}{\beta_3 x^2} [\beta_3 \gamma f^2 + R - h - \beta_3 f + \beta_3 (1-\gamma) f^2] \\ &= \frac{\gamma^2 \psi^2}{\beta_3 x^2} [-\beta_3 (1-f)f + R - h] \\ &= \frac{\gamma^2 \psi^2}{\beta_3 x^2} [Q - h] \leq 0 \end{aligned}$$

with equality when  $x = x_0$  or  $x = x_T$ , by (5.12)(iii) and the fact that  $h$  is increasing in  $(x_0, x_T)$ . When  $-(1-\mu) < x < x_0$ ,

$$\begin{aligned} \gamma \psi \psi'' + (1-\gamma) \psi'^2 &= \gamma A^2 (x + 1 - \mu)^{2\gamma-2} \frac{1}{\gamma} (\gamma - 1) + (1-\gamma) A^2 (x + 1 - \mu)^{2\gamma-2} \\ &= 0 \end{aligned}$$

and similarly when  $x > x_T$ . Then, since  $\psi(x) > 0$  for all  $x > -(1-\mu)$ , it follows that  $\psi'(\cdot)$  is continuous. Hence, the constructed function  $\psi$  is  $C^2$ , and satisfies (5.2)–(5.5).

(b) When  $\gamma = 0$ , the argument is similar. The corresponding initial value problems are:

$$(5.16) \quad f'(x) = \frac{1}{\beta_3 x} [R_1(f) - h],$$

$$h'(x) = \frac{\delta}{\beta_3 x f} [h - Q_1(f)], \quad x \geq x_0 > 0, \text{ where}$$

$$Q_1(f) = -\delta\beta_4 - \delta \log \delta - \beta_2 f + \beta_3 f^2,$$

$$R_1(f) = Q_1(f) + \beta_3(1-f)f = -\delta\beta_4 - \delta \log \delta + (\beta_3 - \beta_2)f,$$

$$f(x_0) = f_0 = \frac{x_0}{x_0 + 1 - \mu}, \quad h(x_0) = h_0 = Q_1(f_0),$$

and  $0 < x_0 < (1 - \mu)\beta_2/(\alpha - r)$ . In this case we define

$$q(x) := \frac{x(1-f(x))}{f(x)}, \quad p(x) := \exp\left(\frac{-h(x)}{\delta}\right).$$

Remarkably, the relation  $p' = -pq'/(x + q)$  still holds; indeed, we find that

$$p' = \frac{-p}{\beta_3 x f} (h - Q_1),$$

$$\begin{aligned} q' &= \frac{1-f}{f} - \frac{1}{\beta_3 f^2} (R_1 - h) \\ &= \frac{1}{\beta_3 f^2} (\beta_3(1-f)f - R_1 + h) \\ &= \frac{1}{\beta_3 f^2} (h - Q_1). \end{aligned}$$

As before, we have  $q(x_0) = 1 - \mu$ ,  $q(x_T) = 1 + \lambda$ , and we define  $q(x) = q(x_0)$  for  $x < x_0$ ,  $q(x) = q(x_T)$  for  $x \geq x_T$ . Similarly, we define  $p(x) = A = p(x_0)$  for  $x \leq x_0$ ,  $p(x) = p(x_T) = B$  for  $x \geq x_T$ . We can now verify in a similar manner to case (a) that

$$\psi(x) := \frac{1}{\delta} \log(p(x)(x + q(x)))$$

satisfies (5.6)–(5.11). This completes the proof.

**6. A representation theorem for the value process.** From the function introduced in the proof of Theorem 5.1, we can obtain useful representations of  $y v_y(x, y)$  and  $v(s_0^*(t), s_1^*(t))$ , evolving under the optimal policy  $(c^*, L^*, U^*)$ . These are needed to complete the proof of Theorem 4.2. They also, show that under the optimal policy, bankruptcy does not occur in finite time.



LEMMA 6.1.

$$(6.1) \quad (i) \quad yv_y(x, y) = \gamma \left( 1 - f\left(\frac{x}{y}\right) \right) v(x, y) \quad \text{for } (x, y) \in NT \text{ when } \gamma \in \Gamma$$

where  $f$  is given by (5.12)

$$(6.1') \quad (ii) \quad yv_y(x, y) = \frac{1}{\delta} \left( 1 - f\left(\frac{x}{y}\right) \right) \quad \text{for } (x, y) \in NT \text{ when } \gamma = 0$$

where  $f$  is given by (5.16).

PROOF. For case (i), express  $v_y$  in terms of  $\psi$  using (3.6) and obtain  $f$  from (5.14), together with the relations (5.15) and (5.16). An analogous procedure gives (6.1').

THEOREM 6.2. (a) Let the conditions of Theorems 4.2 and 5.1 hold, let  $(c^*, L^*, U^*)$  be the policy described by Theorem 4.2, let  $s_0^*(t), s_1^*(t)$  be the solution of equation (4.1) with starting point  $(x, y)$ , and define  $\xi_t^* = s_0^*(t)/s_1^*(t)$ . Then, for  $u(c) = c^\gamma/\gamma$ ,  $\gamma \in \Gamma$ , we have

$$(6.2) \quad v(s_0^*(T), s_1^*(T)) = v(x, y) \exp \left[ \int_0^T \left( \delta - \frac{\gamma}{1-\gamma} h(\xi_t^*) \right) dt \right] \\ \times \exp \left[ \int_0^T \gamma \sigma (1 - f(\xi_t^*)) dz(t) - \frac{1}{2} \int_0^T \gamma^2 \sigma^2 (-f(\xi_t^*))^2 dt \right].$$

(b) When  $u(c) = \log c$ , the corresponding result is:

$$(6.2') \quad v(s_0^*(t), s_1^*(T)) = v(x, y) + \int_0^T \left( -\log \delta - \frac{1}{\delta} h(\xi_t^*) \right) dt \\ + \frac{1}{2} \int_0^T \frac{\sigma}{\delta} (1 - f(\xi_t^*)) dz(t).$$

REMARK 6.3. (6.2) and (6.2') are the counterparts in the transactions costs case of the expression for the wealth process in the Merton problem obtained in Corollary 2.3 above. Indeed, (6.2) reduces to (2.14) when  $\lambda = \mu = 0$ , because then  $f(\xi_t^*) = f_m$ ,  $h(\xi_t^*) = h_m$ , where  $f_m$  and  $h_m = Q(f_m)$  are as given in Remark 5.3 above. Substituting these values in (6.2), we obtain the Merton expression (2.14).

PROOF. First, observe that for any  $(x, y) \in NT$ , we have:

(a.i) When  $u(c) = c^\gamma/\gamma$ ,  $\gamma \in (0, 1)$ , then  $v(x, y) \geq 0$  and  $v(x, y) \neq 0 \Leftrightarrow (x, y)^T \neq (0, 0)^T$ .

(a.ii) When  $u(c) = c^\gamma/\gamma$ ,  $\gamma < 0$ , then  $v(x, y) < 0$  and  $v(x, y) > -\infty \Leftrightarrow (x, y)^T \neq (0, 0)^T$ .

(b)  $u(c) = \log c$ , then  $v(x, y)$  takes both positive and negative values, and  $v(x, y) > -\infty \Leftrightarrow (x, y)^T \neq (0, 0)^T$ .

Suppose  $\underline{s}^*(0) := (s_0^*(0), s_1^*(0)) = (x, y) \in NT \setminus \{0\}$  (without loss of generality). Let  $\tau := \inf\{t: \underline{s}^*(t) = 0\}$ . Then  $\tau > 0$  a.s. by continuity of the wealth process.

(a) Applying the Ito differentiation formula to  $\log[e^{-\delta(T \wedge \tau)} \gamma v(\underline{s}^*(T \wedge \tau))]$  yields:

$$\begin{aligned}
 (6.3) \quad & \log[e^{-\delta(T \wedge \tau)} \gamma v(\underline{s}^*(T \wedge \tau))] - \log \gamma v(\underline{s}^*(0)) \\
 &= \int_0^{T \wedge \tau} \frac{1}{v} \left[ -\delta v + r s_0^* v_x + \alpha s_1^* v_y + \frac{1}{2} \sigma^2 s_1^{*2} v_{yy} - c^* v_x \right] dt \\
 &\quad - \int_0^{T \wedge \tau} \frac{1}{2v^2} \sigma^2 s_1^{*2} v_y^2 dt \\
 &\quad + \int_0^{T \wedge \tau} \frac{1}{v} [(-(1+\lambda)v_x + v_y) dL^*(t) + ((1-\mu)v_x - v_y) dU^*(t)] \\
 &\quad + \int_0^{T \wedge \tau} \frac{1}{v} v_y \sigma s_1^* dz(t) \\
 &= \int_0^{T \wedge \tau} \frac{1}{v} \left[ Gv - \frac{1}{\gamma} (v_x)^{-\gamma/(1-\gamma)} \right] dt + \int_0^{T \wedge \tau} \frac{1}{v} v_y \sigma s_1^* dz(t) \\
 &\quad - \int_0^{T \wedge \tau} \frac{1}{2v^2} \sigma^2 s_1^{*2} v_y^2 dt,
 \end{aligned}$$

since the boundary terms vanish, where

$$Gv = \frac{1}{2} \sigma^2 y^2 v_{yy} + r x v_x + \alpha y v_y - \delta v + \left( \frac{1-\gamma}{\gamma} \right) (v_x)^{-\gamma/(1-\gamma)}.$$

First,  $Gv(\underline{s}^*) \equiv 0$ . Next, recall from (4.9) that  $v(x, y) = y^\gamma \psi(x/y)$ . From (5.14), (5.15) we find that  $(\gamma v)^{-1} (v_x)^{-\gamma/(1-\gamma)} = p^{-\gamma/(1-\gamma)}$ , and from (6.1) that  $y v_y / v = \gamma(1 - f(x/y))$ . Thus (6.3) simplifies to

$$\begin{aligned}
 \log \left( \frac{e^{-\delta(T \wedge \tau)} \gamma v(\underline{s}^*(T \wedge \tau))}{\gamma v(\underline{s}^*(0))} \right) &= \int_0^{T \wedge \tau} \left( -p^{-\gamma/(1-\gamma)} - \frac{1}{2} \gamma^2 \sigma^2 (1-f)^2 \right) dt \\
 &\quad + \int_0^{T \wedge \tau} \gamma \sigma (1-f) dz(t).
 \end{aligned}$$

Put  $h \equiv ((1-\gamma)/\gamma) p^{-\gamma/(1-\gamma)}$  and obtain the following, by exponentiation:

$$\begin{aligned}
 (6.4) \quad & \gamma v(s_0^*(T \wedge \tau), s_1^*(T \wedge \tau)) \\
 &= \gamma v(x, y) \exp \left( \int_0^{T \wedge \tau} \left( \delta - \frac{\gamma}{1-\gamma} h(\xi_t^*) \right) dt \right) \\
 &\quad \times \exp \left( \int_0^{T \wedge \tau} \gamma \sigma (1-f(\xi_t^*)) dz(t) - \frac{1}{2} \int_0^{T \wedge \tau} \gamma^2 \sigma^2 (1-f(\xi_t^*))^2 dt \right).
 \end{aligned}$$

Since  $f \in (0, 1)$ ,  $\int_0^t \gamma \sigma (1-f) dz$  is a time-changed Brownian motion, finite for all  $t$  almost surely. Also  $\gamma h$  is bounded (away from zero). Write the right-hand side of (6.4) as  $\gamma V(T \wedge \tau)$ . Suppose  $\tau < \infty$ , on some set  $\Delta$  of positive measure and choose a sample path in  $\Delta$ .

If  $0 < yv(x, y) < \infty$ , then  $0 < \lim_{t \rightarrow \tau} \gamma V(t \wedge \tau) < \infty$  since  $V(t \wedge \tau)$  is a continuous process. On the other hand by definition of  $\tau$ ,

$$\lim_{t \rightarrow \tau} \gamma v(s_0^*(t \wedge \tau), s_1^*(t \wedge \tau)) = 0 \quad \text{or} \quad \infty$$

contradicting (6.4). Thus  $\tau = \infty$ , and  $0 < yv(s_0^*(t), s_1^*(t)) < \infty$  for all  $t \geq 0$  almost surely. So (6.4) is equivalent to (6.2).

(b) In the log case,

$$(6.3)' \quad v(s^*(T \wedge \tau)) - v(s^*(0)) = \int_0^{T \wedge \tau} [\log v_x + \delta v] dt + \int_0^{T \wedge \tau} v_y \sigma s_1^* dz(t).$$

Since  $v$  satisfies the Bellman equation, we find by introducing the functions  $p, q, f$  of §5 that the right-hand side of (6.3') is equal to

$$\begin{aligned} & \int_0^{T \wedge \tau} [-\log(\delta(s_0^* + q(\xi_t^*)s_1^*))] dt \\ & + \int_0^{T \wedge \tau} \log[p(\xi_t^*)(s_0^* + q(\xi_t^*)s_1^*)] dt + \int_0^{T \wedge \tau} \frac{\sigma}{\delta} (1 - f) dz(t) \\ & = \int_0^{T \wedge \tau} \log\left[\frac{1}{\delta} p(\xi_t^*)\right] dt + \int_0^{T \wedge \tau} \frac{1}{\delta} \sigma [1 - f(\xi_t^*)] dz(t) \\ & = \int_0^{T \wedge \tau} \left[-\log \delta - \frac{1}{\delta} h(\xi_t^*)\right] dt + \int_0^{T \wedge \tau} \frac{1}{\delta} \sigma [1 - f(\xi_t^*)] dz(t) \end{aligned}$$

where  $h = -\delta \log p$ . So

$$\begin{aligned} v(s_0^*(T \wedge \tau), s_1^*(T \wedge \tau)) &= v(x, y) + \int_0^{T \wedge \tau} \left[-\log \delta - \frac{1}{\delta} h(\xi_t^*)\right] dt \\ &+ \int_0^{T \wedge \tau} \frac{1}{\delta} \sigma [1 - f(\xi_t^*)] dz(t). \end{aligned}$$

Again, the integrands are bounded. So the assumption that  $\tau < \infty$  leads to a contradiction, as before. Thus  $\tau = \infty$  and  $v(s_0^*(t), s_1^*(t)) > -\infty$  for all  $t$  almost surely if  $v(x, y) > -\infty$ .

**COROLLARY 6.4.** *The optimal strategies do not lead to bankruptcy in finite time.*

**7. Summary, numerical results and conclusions.** Combining Theorems 4.2, 4.3 and 5.1 we can now state the main result of this paper.

**THEOREM 7.1.** *Suppose that  $\mu \in [0, 1]$ ,  $\lambda \in [0, \infty]$  and  $\mu \vee \lambda > 0$ .*

*Case (a): utility function  $u(c) = c^\gamma/\gamma$ ,  $\gamma \in \Gamma$ .*

*Suppose that Conditions A and B hold, and that*

$$(7.1) \quad r < \alpha < r + (1 - \gamma)\sigma^2.$$

*Then there is a  $C^2$  solution  $\psi$  to the free boundary problem (5.2)–(5.4) with  $0 < x_0 < x_T < \infty$ , and the policy  $(c^*, L^*, U^*)$  described in Theorem 4.2 is optimal in the class  $\mathcal{U}'$  ( $\gamma \in (0, 1)$ ) or  $\mathcal{U}''$  ( $\gamma < 0$ ). In particular,  $L^*, U^*$  are the local times of the wealth*

equation (4.1) at the boundaries of the no-transaction region  $NT$ . Thus the optimal strategy is minimal trading to keep the proportion of wealth held in stock between  $\pi_1^* := (1 + x_T)^{-1}$  and  $\pi_2^* := (1 + x_0)^{-1}$ .

Case (b): utility function  $u(c) = \log c$ .

Suppose that  $\delta > 0$ , that Condition B holds, and that  $r < \alpha < r + \sigma^2$  (i.e. (7.1) holds with  $\gamma = 0$ ). Then there is a  $C^2$  solution to the free boundary problem (5.7)–(5.9) with  $0 < x_0 < x_T < \infty$  and the policy  $(c^*, L^*, U^*)$  described in Theorem 4.3 is optimal in  $\mathcal{U}$ .

If condition (7.1) is not met then the situation is similar to that described for the Merton problem in Remark 2.2: If  $\alpha > r + (1 - \gamma)\sigma^2$  then leverage is optimal and the no-transaction region is a wedge in  $\mathcal{S}'_\mu \cap \{x < 0\}$ , whereas if  $\alpha < r$  shortselling is optimal. If  $\alpha = r$  then cashing out all stock holdings is optimal, as shown in Theorem 4.7. The case  $\alpha = r + (1 - \gamma)\sigma^2$  is unsolved, but we conjecture that it involves a wedge with vertical upper barrier, with a discontinuity in the second derivative of the value function at this barrier. If Condition A does not hold then arbitrarily high utility can be attained.

We now turn to computation of the optimal policy. This is best done in terms of the transformed coordinates introduced in the proof of Theorem 5.1. Recall from (5.11), (5.12) that, for  $u(c) = c^\gamma/\gamma$ , if  $\psi(x), x_0, x_T$  is the solution of the free boundary problem (5.3)–(5.5) then there are functions  $f(x), h(x)$  which satisfy the system of differential equations

$$(7.2) \quad f' = \frac{1}{\beta_3 x} (R(f) - h) \quad h' = \frac{\gamma}{1 - \gamma} \frac{h}{\beta_3 x f} (h - Q(f)),$$

$$(7.3) \quad f(x_0) = \frac{x_0}{x_0 + 1 - \mu} =: f_0, \quad h(x_0) = Q(f_0),$$

$$(7.4) \quad f(x_T) = \frac{x_T}{x_T + 1 + \lambda} =: f_T, \quad h(x_T) = Q(f_T).$$

Now from (5.14)  $\psi$  is given by

$$\psi(x) = \frac{1}{\gamma} p^\gamma(x) (x + q(x))^\gamma,$$

where  $p, q$  are defined in terms of  $x, f, h$  by

$$p(x) = \left( \frac{\gamma}{1 - \gamma} h(x) \right)^{-(1-\gamma)/\gamma}, \quad q(x) = x \left( \frac{1}{f(x)} - 1 \right).$$

We then see that

$$(7.5) \quad \psi(x) = \frac{1}{\gamma} \left( \frac{\gamma}{1 - \gamma} h(x) \right)^{-(1-\gamma)} x^\gamma f^{-\gamma}.$$

From (5.15) we have

$$(7.6) \quad \psi'(x) = x^{-1} \gamma \psi(x) f(x),$$

while from (4.8) we know that the optimal consumption policy  $c^*(x, y)$  is given by

$$c^*(x, y) = y \left[ \psi' \left( \frac{x}{y} \right) \right]^{-1/(1-\gamma)}.$$

In view of (7.5), (7.6) we can calculate  $c^*$  directly in terms of the function  $f(x), h(x)$ :

$$c^*(x, y) = \frac{\gamma x h(x/y)}{(1-\gamma)f(x/y)}.$$

The problem therefore reduces to computing the solution to (7.2) with boundary conditions (7.3), (7.4), which say that both at  $x_0$  and at  $x_T$  we have  $(f, h) \in \Omega = \{(f, h): h = Q(f)\}$ . Figure 4 below shows some typical trajectories (one should appreciate that the “phase space” is actually 3-dimensional since (7.2) is nonautonomous). The minimum of  $Q$  occurs at  $f = f_m = 1 - (\alpha - r)/(\sigma^2(1 - \gamma)) \in (0, 1)$ . Thus  $f_0 \leq f_m, f_T \geq f_m$ .

The algorithm for solving (7.2)–(7.4) is as follows.

1. Choose any  $x_T$  such that  $f_T \in (f_m, 1)$ , where  $f_T := x_T/(x_T + 1 + \lambda)$ . Define  $h_T = Q(f_T)$ .
2. Using numerical integration, solve (7.2) backwards (i.e. in the direction of decreasing  $x$ ) until the trajectory re-crosses  $\Omega$ . Let  $x_0$  be the value of  $x$  at which this happens ( $x_0 = \sup\{x < x_T: (f(x), h(x)) \in \Omega\}$ ) and let  $f_0 := f(x_0)$ .
3. Define

$$(7.7) \quad m := x_0 + 1 - \frac{x_0}{f_0}.$$

We see from (7.7) that  $f_0 = x_0/(x_0 + 1 - m)$ , and hence, referring to (7.3), that  $(x_0, x_T)$  determined in this way solves (7.2)–(7.4) for the given value of  $\lambda$  and with  $\mu = m$ . We can regard Steps 1–3 above as a function which maps  $x_T$  to  $m = m(x_T)$ . The solution of (7.2)–(7.4) is completed by embedding 1–3 in a *one-dimensional search procedure* to find a value of  $x_T$  such that  $m(x_T) = \mu$  (the prescribed proportional cost for sales). The argument given in the Appendix shows that this search will always be successful. The reason for integrating backwards rather than forwards in Step 2 is that this is the “stable” direction of (7.2): the solution integrated forwards is very sensitive to the initial condition  $f_0$ , whereas the backwards solution is much more robust. An exactly similar algorithm, based on equations (5.16), solves the problem for the utility function  $u(c) = \log c$ .

We plan to report more fully on the numerical results in a later publication, but Figure 3 shows some typical results. Our result says that the proportion of wealth held in stock should be kept between  $\pi_1^* = 100(1 + x_T)^{-1}\%$  and  $\pi_2^* = 100(1 + x_0)^{-1}\%$ . The Merton proportion (no transaction costs) is  $\pi^* = 100 \times 2(\alpha - r)/\sigma^2$ . In the present case this is equal to 15.63%. We have taken  $\mu = \lambda$  (equal transaction costs on sale and purchase) and plotted  $\pi_1^*, \pi_2^*$  against  $\lambda$ . The most noticeable feature of these curves is that the upper (sell) barrier is very insensitive to  $\lambda$  while the lower (buy) barrier decreases quite rapidly as  $\lambda$  increases. This is probably due to the asymmetry in the model: all consumption takes place from the bank, so stock must be sold (and transaction charges paid) before it can be realized for consumption. If the selling charge is high then this is unfortunate but unavoidable. On the other hand if the bulk of the investor's holdings are in cash then the potential gains from investing

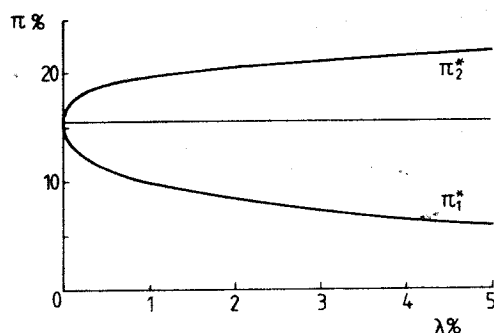


FIGURE 3. Dependence of  $\pi_1^*, \pi_2^*$  on  $\lambda$  (with  $\mu = \lambda$ ). The Parameter Values are  $\alpha = 0.12$ ,  $r = 0.07$ ,  $\sigma = 0.4$ ,  $\delta = 0.1$ ,  $\gamma = -1$ .

in stock and then reselling at some later date may not be worthwhile if the associated costs are too high. Hence the decreasing value for  $\pi_1^*$ .

In [4], Constantinides considers exactly the same problem as in this paper and obtains approximate results by restricting the class of consumption policies  $c(x, y)$  to those satisfying

$$(7.8) \quad \frac{c(x, y)}{x} = \text{constant}.$$

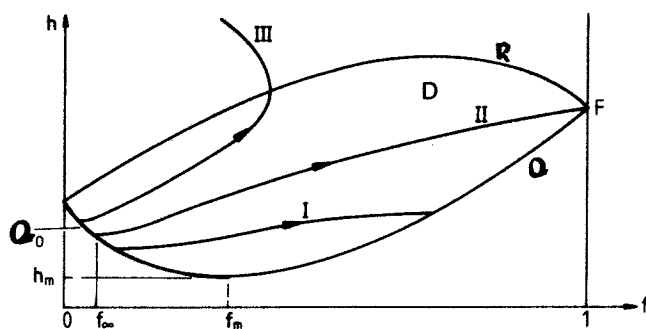
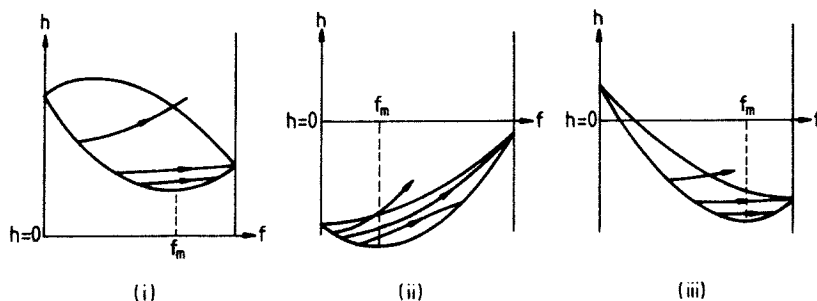
His results are qualitatively similar to those displayed in Figure 3. We can use our numerical technique to check whether (7.8) is exactly or approximately satisfied for truly optimal policies. For  $\gamma \neq 0$  the optimal consumption  $c^*$  is given by (4.9) and we see that for  $(x, y) \in NT$

$$\frac{c^*(x, y)}{x} = \xi^{-1} [\psi'(\xi)]^{-1/(1-\gamma)} =: \bar{c}(\xi), \quad \xi := \frac{x}{y}.$$

In  $NT$ ,  $\xi$  ranges over the interval  $[x_0, x_T]$ . Taking parameter values as above and  $\lambda = \mu = 0.01$ , we find that  $x_0 = 0.406$ ,  $x_T = 1.112$ ,  $\bar{c}(x_0) = 0.188$  and  $\bar{c}(x_T) = 0.105$ . Thus the ratio in (7.8) actually varies over a range of nearly two to one. We have not, however, investigated how much *utility* is lost by imposing the restriction (7.8).

Finally, let us consider possible extensions of this work. Of course, the most interesting extension would be to the case of  $m > 1$  risky assets, but unfortunately this is essentially impossible except perhaps for  $m = 2$  or  $3$ . As Magill and Constantinides [17] point out,  $m$  risky assets imply  $3^m$  possible transaction regions and, for example,  $3^{10} \approx 60000$ . Although the coordinate transformation introduced in §5 generalizes to higher dimensions, it is not clear how to locate the boundaries even when  $m = 2$ . The only readily solvable case is that in which transactions *between risky assets* are costless. The risky assets can then be combined via a mutual fund theorem [17] and the problem reduces to the single risky asset case considered here. See Magill [16]. Another important question is *nonconstant model parameters* (interest rates, return on assets, volatility). It is unlikely that our problem could ever be solved at the level of generality of, say, Karatzas *et al.* [11], where these parameters are taken as general stochastic processes. However, it might be solvable if the randomness of the parameters were modelled in specific ways, for example as finite state Markov processes. This is an interesting area for further research.



FIGURE 4.  $(f, h)$ -Plane for  $\gamma \in (0, 1)$ ,  $f_m < \frac{1}{2}$ .FIGURE 5.  $(f, h)$  plane for (i)  $\gamma \in (0, 1)$ ,  $f_m > \frac{1}{2}$ , (ii)  $\gamma < 0$ ,  $f_m < \frac{1}{2}$ , (iii)  $\gamma < 0$ ,  $f_m > \frac{1}{2}$ . When  $\gamma = 0$  the pictures are similar, but parabola  $\mathfrak{R}$  becomes a straight line.

**Acknowledgements.** We would like to thank V. E. Beneš, I. Karatzas, P. Varaiya and R. J. Williams for helpful discussions and advice, and the two referees for their detailed attention and many valuable remarks. We would also like to thank D. Abeysekera for her unstinting help in the preparation of the manuscript.

**Appendix.** In this Appendix we examine the solution trajectories  $(x, f(x), h(x)) \in \mathbb{R}^3$  of the nonautonomous first-order system defined by equation (5.11):

$$(5.11) \quad f'(x) = \frac{1}{\beta_3 x} [R(f(x)) - h(x)],$$

$$h'(x) = \frac{\gamma}{1 - \gamma} \frac{h}{\beta_3 x f(x)} [h(x) - Q(f(x))].$$

We show that for each  $\mu < 1$  and  $\lambda > -\mu$ , these equations have a solution satisfying (5.12)(i)–(v).<sup>6</sup> This will follow from the sequence of lemmas presented below. The proofs of these lemmas are collected together at the end of the Appendix.

Denote by  $\mathcal{Q}$  and  $\mathfrak{R}$  respectively the graphs in  $\mathbb{R}^2$  of the functions  $Q$  and  $R$  for  $f \in [0, 1]$ , and define  $\mathcal{Q}_0 := \{(f, h): 0 < f < f_m\}$  where  $f_m := 1 - (\alpha - r)/(1 - \gamma)\sigma^2$  is the value of  $f$  at which  $Q(f)$  attains its minimum. We also define  $\tilde{\mathcal{Q}} = \mathbb{R} \times \mathcal{Q}$ ,  $\tilde{\mathcal{Q}}_0 = \mathbb{R} \times \mathcal{Q}_0$  and  $\mathfrak{R} = \mathbb{R} \times \mathfrak{R}$ .  $\tilde{\mathcal{Q}}$  and  $\mathfrak{R}$  enclose a region  $\tilde{D}$  in  $(x, f, h)$  space and we denote by  $D$  the projection of  $\tilde{D}$  onto the  $(f, h)$ -plane (see Figures 4 and 5). The

<sup>6</sup>We only need to prove existence because the uniqueness of the optimal payoff functions (3.2) will follow immediately.

only "stationary point"  $F := (1, Q(1)) = (1, R(1)) = (1, (\delta/\gamma) - r)$  corresponds to a straight line trajectory  $\bar{F} = (x, 1, Q(1)) \subset \tilde{\mathcal{Q}} \cap \tilde{\mathcal{R}}$ .<sup>7</sup>

Fix  $\mu < 1$  and consider the solution of (5.11) starting at  $(x_0, f(x_0), h(x_0)) = (x_0, f_0, h_0) \in \tilde{\mathcal{Q}}_0$  with  $x_0 := X_\mu(f_0) := (1 - \mu)f_0/(1 - f_0)$ . We note that  $f'(x_0) > 0$  and  $h'(x_0) = 0$ , and that  $f', h' > 0$  inside  $\bar{D}$ . Thus, the trajectory moves monotonically in each variable across  $\bar{D}$  and can only (I) exit through  $\tilde{\mathcal{Q}}$ , (II) exit through  $\tilde{\mathcal{R}}$  or (III) tend to  $\bar{F}$ . We call such trajectories types I, II, III respectively.

Suppose for a moment that (I) holds. Denote by  $\bar{x}_T, \bar{f}_T, \bar{h}_T$  the exit values of  $x, f, h$ , and define  $l = l(f_0) := (\bar{x}_T/\bar{f}_T) - x_T - 1$ . Since  $Q(\bar{f}_m) = (\delta/\gamma) - r - (\alpha - r)^2/2(1 - \gamma)\sigma^2$ , Condition A implies that  $h > 0$  for all  $(f, h) \in D$  when  $\gamma > 0$ . When  $\gamma < 0$ , at each point  $(f, h)$  on our trajectory we have  $h < \bar{h}_T < Q(1) = (\delta/\gamma) - r < 0$ . Thus  $\gamma h > 0$  in either case, and properties (5.13)(i)–(v) are all satisfied with  $\lambda = l$ . We therefore need to show that for each  $\lambda > -\mu$  there is a type I trajectory with  $l = \lambda$ .

Consider first the extreme case  $f_0 = f_m$ . We have the following result.

**LEMMA A.1.** Fix  $\mu < 1$ , as before. Set  $f_0 = f_m$ ,  $h_0 = Q(f_0)$  and  $x_0 = X_\mu(f_0) > 0$ . Then  $h''(x_0) = 0$ ,  $h'''(x_0) < 0$ .

This shows that the trajectory of (5.11) does not enter the interior of  $D$ , and hence that  $x_T = x_0$ ,  $f_T = f_m$ ,  $h_T = Q(f_m)$  and  $l = -\mu$ .<sup>8</sup>

It remains to show that  $l(f_0)$  ranges continuously up to infinity as we decrease the starting value  $f_0$  of the type I trajectory keeping  $(x_0, f_0, h_0)$  on  $\tilde{\mathcal{Q}}_0$  and maintaining the relationship  $x_0 = X_\mu(f_0)$ . We shall discover the behaviour of  $l(f_0)$  for  $f_0 < f_m$  by investigating the function  $q(x) = x(1 - f(x))/f(x)$  introduced in the proof of Theorem 5.1, and noting that  $q(x_0) = 1 - \mu$ , and  $q(x_T) = 1 + l$ . We shall also use (without proof) the fact that solutions to (5.11) are line integrals which are continuous functions of the initial values. The following lemma is of crucial importance. It shows that the projections onto  $D$  of trajectories with distinct starting points in  $\tilde{\mathcal{Q}}_0$  do not intersect.

**LEMMA A.2.** Let  $(x, f_1(x), h_1(x))$  and  $(x, f_2(x), h_2(x))$  be two solution trajectories of (5.11) starting at  $(f_{i0}, h_{i0}) \in \tilde{\mathcal{Q}}_0$ ,  $x_{i0} = X_\mu(f_{i0})$ ,  $i = 1, 2$ . Define  $\mathcal{G}'_i = \{(f_i(x), h_i(x)) : x \geq x_{i0}\}$  and  $\mathcal{G}_i = \mathcal{G}'_i \cap D$ ,  $i = 1, 2$ . Suppose that  $f_{10} > 0$  and  $f_{20} > 0$ ,  $Q(f_{10}) \leq Q(1)$  and  $Q(f_{20}) \leq Q(1)$ , and  $h_{20} > h_{10}$  (so  $f_{20} < f_{10}$ ). Suppose  $(f, h_1) \in \mathcal{G}_1$  and  $(f, h_2) \in \mathcal{G}_2$  for  $f \in [0, 1]$ . Then  $h_2 > h_1$ .

**COROLLARY.** Let  $(x, f_1(x), h_1(x))$  and  $(x, f_2(x), h_2(x))$  be as above, and suppose that  $(x, f_2(x), h_2(x))$  is a type I trajectory. Then  $(x, f_1(x), h_1(x))$  is also of type I. Thus the terminal value  $f_T \in (f_m, 1)$  is a monotonic decreasing function of  $f_0$ .

Next, define  $f_\infty := \inf\{f_0 : (x_T, f(x_T), h(x_T)) \in \tilde{\mathcal{Q}}\}$ . In view of the above results, all trajectories with starting points  $f_0$  such that  $f_m > f_0 > f_\infty$  are of type I. We need the following condition:

**Condition B.**  $f_\infty > 0$ .

If  $f_m > \frac{1}{2}$ ,  $0 \leq \gamma < 1$ , then Condition B automatically holds, since the trajectory with  $h(x_0) = Q(1)$  is necessarily of type II (see Figure 5(i) and (iii)). Numerical work convinces us that  $f_\infty > 0$  in the remaining cases, but we are unable at present to give a proof.

<sup>7</sup>The  $\mu = 1$  case.

<sup>8</sup>When  $\lambda = \mu = 0$ , this corresponds to the solution of the Merton problem, and indeed we find that  $x_0 = f_m/(1 - f_m)$ , corresponding to the Merton proportion  $\pi^*$  of (2.6), namely  $\pi^* = 1/(\pi_0 + 1) = (\alpha - r)/(1 - \gamma)\sigma^2$ .

The next lemma completes the argument by establishing that for any specified  $\lambda > -\mu$ , there is a (type I) trajectory for which  $q(\lambda_T) = 1 + \lambda$ .

LEMMA A.3. Suppose Condition B holds. Then as  $f_0$  decreases from  $f_m$  to  $f_\infty$ ,

- (i)  $q(x_T)$  increases from  $1 - \mu$  to  $\infty$ , and
- (ii)  $x_T$  increases from  $(1 - \mu)(1 - f_m)/f_m$  to  $\infty$ .

PROOF OF LEMMA A.1. The two cases are:

(a)  $\gamma \in \Gamma$ . Then

$$\begin{aligned} \frac{d}{dx}(\beta_3 x f h') \Big|_{x=x_0} &= \frac{d}{dx} \left( \frac{\gamma}{1-\gamma} h[h - Q(f)] \right) \Big|_{x=x_0} \\ &= \frac{\gamma}{1-\gamma} h'[h - Q(f)] + \frac{\gamma}{1-\gamma} h[h' - Q'(f')] \Big|_{x=x_0} = 0. \end{aligned}$$

Also,

$$\frac{d}{dx}(\beta_3 x f h') \Big|_{x=x_0} = h' \cdot \frac{d}{dx}(\beta_3 x f) + \beta_3 x f h'' \Big|_{x=x_0} = 0.$$

But  $f(x_0) \neq 0$ , so  $h''(x_0) = 0$ .

Next,

$$\begin{aligned} \frac{d^2}{dx^2}(\beta_3 x f h') \Big|_{x=x_0} &= \beta_3((2f' + x f'')h' + 2(f + x f')h'' + x f h''') \Big|_{x=x_0} \\ &= h'''(x_0) \beta_3 x_0 f_0. \end{aligned}$$

Another expression for this derivative is

$$\begin{aligned} &\frac{d}{dx} \left( \frac{\gamma}{1-\gamma} h'[h - Q(f)] \right) + \frac{\gamma}{1-\gamma} h[h' - Q'(f')] \Big|_{x=x_0} \\ &= \frac{\gamma}{1-\gamma} (h''[h - Q] + 2h'[h' - Q'(f')] + h[h'' - Q''f'^2 + Q'(f'')]) \Big|_{x=x_0} \\ &= \frac{-\gamma}{1-\gamma} h(x_0) \cdot 2\beta_3(1-\gamma) \left( \frac{R(f(x_0)) - h(x_0)}{\beta_3 x_0} \right)^2 < 0, \end{aligned}$$

showing that  $h'''(x_0) < 0$ .

(b)  $\gamma = 0$ . In this case,

$$\begin{aligned} \frac{d}{dx}(\beta_3 x f h') \Big|_{x=x_0} &= \delta \frac{d}{dx}(h - Q_1(f)) \Big|_{x=x_0} = \delta(h' - Q_1'(f')) \Big|_{x=x_0} \\ &= \beta_3 x_0 f_0 h''(x_0). \end{aligned}$$

So  $h''(x_0) = 0$ . Further,

$$\begin{aligned} \frac{d^2}{dx^2}(\beta_3 x f h') \Big|_{x=x_0} &= \beta_3((2f' + x f'')h' + 2(f + x f')h'' + x f h''') \Big|_{x=x_0} \\ &= h'''(x_0) \beta_3 x_0 f_0 \\ &= \delta(h'' - Q_1'' f'^2 - Q_1' f'') \Big|_{x=x_0} \\ &= -2\delta\beta_3 \left( \frac{R_1(f(x_0)) - h(x_0)}{\beta_3 x_0} \right)^2 < 0, \end{aligned}$$

showing that  $h'''(x_0) < 0$ . This completes the proof.

**PROOF OF LEMMA A.2.** Consider two trajectories evaluated at common  $f_1 = f_2 = f$ . When  $\gamma \in \Gamma$ , we have

$$\frac{dh}{df} = \frac{\gamma}{1-\gamma} \frac{h}{f} \left( \frac{h - Q(f)}{R(f) - h} \right).$$

Thus if  $h_1, h_2$ , lie in  $D$  (see Figure 5(i)), then

$$\frac{d}{df}(h_2 - h_1) = \frac{1}{1-\gamma} \frac{1}{f} \left( \gamma h_2 \left( \frac{h_2 - Q(f)}{R(f) - h_2} \right) - \gamma h_1 \left( \frac{h_1 - Q(f)}{R(f) - h_1} \right) \right).$$

When  $\gamma = 0$ , the corresponding expressions are:

$$\begin{aligned} \frac{dh}{df} &= \frac{\delta}{f} \left( \frac{h - Q_1(f)}{R_1(f) - h} \right), \\ \frac{d}{df}(h_2 - h_1) &= \frac{\delta}{f} \left( \left( \frac{h_2 - Q_1(f)}{R_1(f) - h_2} \right) - \left( \frac{h_1 - Q_1(f)}{R_1(f) - h_1} \right) \right). \end{aligned}$$

In either case, if  $h_2 > h_1$ , then  $d(h_2 - h_1)/df > 0$ . Since  $h_2$  increases in  $D$ , and  $h_{20} > h_{10}$  then  $h_2 > h_1$  for all  $f$ , as claimed.

**PROOF OF LEMMA A.3.** (i) Consider the case  $\gamma \in \Gamma$  first. We saw that  $q(x_T) = 1 - \mu$  when  $f_0 = f_m$ . Take two trajectories with  $h_2 > h_1$ , as before. Now  $q_2$  increases since  $q' = [h - Q(f)]/\beta_3 f^2 \geq 0$ . So, initially  $q_2 > q_1 = 1 - \mu$ . With a common  $f$  as in Lemma A.2, the corresponding values of  $x$  are  $x_1, x_2$  given by:

$$\begin{aligned} f &= \frac{x_1}{x_1 + q_1(x_1)} = \frac{x_2}{x_2 + q_2(x_2)}, \\ \frac{q_1(x_1)}{x_1} &= \frac{q_2(x_2)}{x_2} \quad \text{i.e.} \quad \frac{q_2(x_2)}{q_1(x_1)} = \frac{x_2}{x_1}. \end{aligned}$$

Then

$$\begin{aligned} \frac{d}{df}(q_2 - q_1) &= \frac{1}{f^2} \left( x_2 \left( \frac{h_2 - Q(f)}{R(f) - h_2} \right) - x_1 \left( \frac{h_1 - Q(f)}{R(f) - h_1} \right) \right) > 0 \\ &= \frac{x_1}{f^2} \left( \frac{q_1}{f^2} \left( \frac{h_2 - Q(f)}{R(f) - h_2} \right) - \left( \frac{h_1 - Q(f)}{R(f) - h_1} \right) \right) > 0 \end{aligned}$$

whenever  $q_2 > q_1$  ( $x_1 > 0$ ).

In case  $\gamma = 0$ , the expressions are the same, with  $Q, R$  replaced by  $Q_1, R_1$ . In both cases,  $q_2 > q_1$  for all (common)  $f$ . Finally, since  $q_2$  increases, we have  $q_2(x_{2T}) > q_1(x_{1T})$ , showing that  $q(x_T)$  increases as  $f_0$  decreases.

To see that  $q(x_T) \rightarrow \infty$  as  $f_0$  decreases to  $f_\infty$ , use the fact that  $q$  is a continuous function of  $f_0$  and  $x$  and Lemma A.4 below which establishes that  $q(x) \rightarrow \infty$  along the type III trajectory with  $f_0 = f_\infty$ .

(ii) Since  $q_2/q_1 = x_2/x_1 > 1$ , we have  $x_2 > x_1$  for all  $f$ , and since  $x_2(f)$  is increasing, we have that  $x_{2T} > x_{1T}$ . That is,  $x_T$  increases as  $f_0$  decreases. When  $f_0 = f_m$ , we have  $f_T = f_0$  and  $x_T = x_0 = (1 - \mu)(1 - f_m)/f_m$ , since  $f_m = f_0 = x_0/(x_0 + 1 - \mu)$ . Finally, since  $f_T = x_T/(x_T + q(x_T))$  and  $q(x_T)$  is increasing and positive, it follows  $x_T$  increases to infinity as  $f_T$  increases to 1, proving (ii).

In the proof of Lemma 3 we needed the following fact concerning the behaviour of the special (type III) trajectory with  $f_0 = f_\infty$ ,  $f_T = 1$  (see Figures 4, 5).<sup>9</sup>

LEMMA A.4. Along the type III (asymptotic) trajectory,  $q(x) \rightarrow \infty$  as  $x \rightarrow \infty$ .

PROOF. (a) For  $\gamma \in \Gamma$ ,

$$\frac{q'}{q} = \frac{1}{x} \left( \frac{h - Q}{R - Q} \right).$$

We know that  $h \rightarrow Q(1)$ ,  $f \rightarrow 1$  as  $x \rightarrow \infty$ . By L'Hôpital's rule,

$$\lim_{x \rightarrow \infty} \left( \frac{h - Q}{R - Q} \right) = \lim_{x \rightarrow \infty} \left( \frac{h' - Q'f'}{(R' - Q')f'} \right) \in [0, 1].$$

Write  $a = h - Q$ ,  $b = R - h$ . Then

$$\begin{aligned} &\lim_{x \rightarrow \infty} \left( \frac{a}{a + b} \right) \\ &= \lim \left( \frac{\frac{\gamma}{1 - \gamma} \cdot \frac{h}{\beta_3 x f} \cdot [h - Q(f)] - [-\beta_2 + 2\beta_3(1 - \gamma)f] \frac{1}{\beta_3 x} [R - h]}{\beta_3(1 - 2f) \cdot \frac{1}{\beta_3 x} \cdot [R - h]} \right) \\ &= \frac{-\gamma}{1 - \gamma} \cdot \frac{Q(1)}{\beta_3} \lim \left( \frac{a}{b} \right) + \frac{-\beta_2 + 2\beta_3(1 - \gamma)}{\beta_3} \\ &= \lim \left( \frac{a/b}{(a/b) + 1} \right) = K, \text{ say.} \end{aligned}$$

<sup>9</sup>This corresponds to a no-transaction region in which the lower barrier lies on the  $x$ -axis, i.e. the cost of buying into the stock is prohibitive.

So,

$$\beta_3 K = 2\beta_3(1 - \gamma) - \beta_2 - \frac{\gamma}{1 - \gamma} \cdot Q(1) \frac{K}{1 - K},$$

and  $K \in [0, 1]$ , i.e.

$$\begin{aligned} H(K) &:= \beta_3 K^2 + \left( -\beta_3 - 2\beta_3(1 - \gamma) + \beta_2 - \frac{\gamma}{1 - \gamma} \cdot Q(1) \right) K + 2\beta_3(1 - \gamma) - \beta_2 \\ &= 0 \\ &= \frac{\sigma^2}{2} K^2 + \left( \frac{-\sigma^2}{2} - (\alpha - r) - \frac{\gamma}{1 - \gamma} \left( \frac{\delta}{\gamma} - r \right) \right) K + \alpha - r. \end{aligned}$$

Since

$$H(0) = \alpha - r > 0 \quad \text{and} \quad H(1) = -\gamma/(1 - \gamma)(\delta/\gamma - r) < 0,$$

it follows that the roots satisfy  $0 < K_1 < 1 < K_2$ . So

$$\begin{aligned} K = K_1 &= \frac{1}{\sigma^2} \left( \frac{1}{2} \sigma^2 + \alpha - r + \frac{\gamma}{1 - \gamma} \left( \frac{\delta}{\gamma} - r \right) \right. \\ &\quad \left. - \sqrt{\left( \frac{1}{2} \sigma^2 + \alpha - r + \frac{\gamma}{1 - \gamma} \left( \frac{\delta}{\gamma} - r \right) \right)^2 - 2\sigma^2(\alpha - r)} \right). \end{aligned}$$

(b) When  $\gamma = 0$ ,

$$\frac{q'}{q} = \frac{1}{x} \left( \frac{h - Q_1}{R_1 - Q_1} \right).$$

Writing

$$\bar{K} = \lim_{x \rightarrow \infty} \left( \frac{h - Q_1}{R_1 - Q_1} \right),$$

the corresponding quadratic is

$$H_1(\bar{K}) = \frac{1}{2} \sigma^2 \bar{K}^2 + \left[ -\frac{1}{2} \sigma^2 - \delta - (\alpha - r) \right] \bar{K} - \alpha - r = 0,$$

$$H_1(0) = \alpha - r > 0, \quad H_1(1) = -\delta < 0.$$

So,  $0 < \bar{K}_1 < 1 < \bar{K}_2$ , and

$$\bar{K} = \bar{K}_1 = \frac{1}{\sigma^2} \left( \frac{1}{2} \sigma^2 + \delta + (\alpha - r) - \sqrt{\left( \frac{1}{2} \sigma^2 + \delta + (\alpha - r) \right)^2 - 2\sigma^2(\alpha - r)} \right).$$

Then, for large  $x$ ,  $q$  is asymptotic to  $x^{\bar{K}}$  (respectively  $x^{\bar{K}_2}$ ). So,  $q \rightarrow \infty$  as  $x \rightarrow \infty$  along the special trajectory, as claimed.

## References

- [1] Anderson, R. F. and Orey, S. (1976). Small Random Perturbations of Dynamical Systems with Reflecting Boundary. *Nagoya Math J.* **60** 189–216.
- [2] Beneš, V. E., Shepp, L. A. and Witsenhausen, H. S. (1980). Some Solvable Stochastic Control Problems. *Stochastics* **4** 134–160.
- [3] Bensoussan, A. and Lions, J. L. (1984). *Impulse Control and Quasi-variational Inequalities*. Gauthier-Villars, Paris.
- [4] Constantinides, G. M. (1986). Capital Market Equilibrium with Transaction Costs, *J. Political Economy* **94** 842–862.
- [5] Davis, M. H. A. (1988). Local Time on the Stock Exchange. in *Stochastic Calculus in Application*. J. R. Norris (Ed.), Pitman Research Notes in Mathematics 197, Longman, London.
- [6] Doléans-Dade, C. (1976). On the Existence and Unicity of Solutions of Stochastic Integral Equations. *Z. Wahrsch. Verw. Gebiete* **36** 93–101.
- [7] Duffie, D. and Sun, T. S. Transaction Costs and Portfolio Choice in a Discrete-Continuous Time Setting. *J. Economic Dynamics and Control*, to appear.
- [8] Fleming, W. H. and Rishel, R. W. (1975). *Deterministic and Stochastic Optimal Control*. Springer-Verlag, New York.
- [9] Harrison, J. M. (1985). *Brownian Motion and Stochastic Flow Systems*. Wiley, New York.
- [10] ——— and Taylor, A. J. (1978). Optimal Control of a Brownian Storage System. *Stochastic Process. Appl.* **6**.
- [11] Karatzas, I., Lehoczky, J. P. and Shreve, S. E. (1987). Optimal Portfolio and Consumption Decisions for a “Small Investor” on a Finite Horizon. *SIAM J. Control Opt.* **25** 1557–1586.
- [12] ———, ———, Sethi, S. P. and Shreve, S. E. (1986). Explicit Solution of a General Consumption/Investment Problem. *Math. Oper. Res.* **11** 261–294.
- [13] ——— and Shreve, S. E. (1986). Equivalent Models for Finite Fuel Stochastic Control. *Stochastics* **18** 245–276.
- [14] ——— and ——— (1984). (1985). Connections between Optimal Stopping and Singular Stochastic Control. I, II. *SIAM J. Control Optim.* **22** 856–877 and **23** 433–451.
- [15] Leland, H. E. (1985). Option Pricing and Replication with Transactions Costs, *J. Finance* **40** 1283–1301.
- [16] Magill, M. J. P. (1976). The Preferability of Investment through a Mutual Fund. *J. Economic Theory* **13** 264–271.
- [17] ——— and Constantinides, G. M. (1976). Portfolio Selection with Transaction Costs. *J. Economic Theory* **13** 245–263.
- [18] Merton, R. C. (1971). Optimum Consumption and Portfolio Rules in a Continuous-Time Model, *J. Economic Theory* **3** 373–413.
- [19] Stroock, D. W. and Varadhan, S. R. S. (1971). Diffusion Processes with Boundary Conditions. *Comm. Pure Appl. Math.* **24** 147–225.
- [20] Taksar, M., Klass, M. J. and Assaf, D. (1988). A Diffusion Model for Optimal Portfolio Selection in the Presence of Brokerage Fees. *Math. Oper. Res.* **13** 277–294.
- [21] Tanaka, H. (1979). Stochastic Differential Equations with Reflecting Boundary Conditions in Convex Regions. *Hiroshima Math. J.* **9** 163–177.
- [22] Varadhan, S. R. S. and Williams, R. J. (1985). Brownian Motion in a Wedge with Oblique Reflection. *Comm. Pure Appl. Math.* **38** 405–443.

DAVIS: DEPARTMENT OF ELECTRICAL ENGINEERING, IMPERIAL COLLEGE, LONDON SW7 2BT, UNITED KINGDOM

NORMAN: COUNTY NATWEST, LONDON, UNITED KINGDOM



Copyright 1990, by INFORMS, all rights reserved. Copyright of Mathematics of Operations Research is the property of INFORMS: Institute for Operations Research and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.