An Exact Solution to a Dynamic Portfolio Choice Problem under Transactions Costs
Author(s): Bernard Dumas and Elisa Luciano
Source: *The Journal of Finance*, Vol. 46, No. 2 (Jun., 1991), pp. 577-595
Published by: Blackwell Publishing for the American Finance Association
Stable URL: http://www.jstor.org/stable/2328837
Accessed: 25/11/2009 12:07

THE JOURNAL OF FINANCE • VOL. XLVI, NO. 2 • JUNE 1991

An Exact Solution to a Dynamic Portfolio Choice Problem under Transactions Costs

BERNARD DUMAS and ELISA LUCIANO*

ABSTRACT

The presence of any friction in financial markets qualitatively changes the nature of the optimization problem faced by an investor. It requires one to either act or do nothing, an issue which, of course, does not arise in frictionless situations. The investor considered here accumulates wealth without consuming until some terminal point in time when he consumes all. His objective is to maximize the expected utility derived from that terminal consumption. We postpone the terminal point far into the future to obtain a stationary portfolio rule. The portfolio policy is in the form of two control barriers between which portfolio proportions are allowed to fluctuate. We show how to calculate them.

MUCH OF FINANCIAL THEORY neglects transactions costs. Perhaps the most successful implementation of it—i.e., continuous-time portfolio choice and option pricing—is downright inconsistent with the existence of any transactions cost at all. For instance, the hedging strategy proposed by Black and Scholes (1973) to value derivative assets is not rationally feasible when portfolio adjustment is not costless. Similarly, other reasonings adopted to justify alternative pricing methods do not seem to be immediately applicable when transaction costs impinge on investment returns.[1] When they are applied, straightforward continuous adjustment of the portfolio composition would lead to infinite transactions costs in a finite time period.

Nonetheless *prima facie* evidence from the trade is that transactions costs are a source of concern for portfolio managers. Included in this category of costs are not only brokerage fees but also any cost of analysis, information

* Dumas is with the Wharton School of the University of Pennsylvania, HEC School of Management, and National Bureau of Economic Research. Luciano is with the Università di Torino, Italy. Dumas gratefully acknowledges the financial support of the A. Shoemaker term chain at the Wharton School. Both authors are grateful to an anonymous referee of this Journal, to Jérôme Detemple, Sanford Grossman, Bruce Lehman, Jean-Luc Vila, participants of workshops at the NBER, Columbia University, the University of Warwick, and the University of Stockholm for helpful comments, and to Francisco Delgado for his diligent and effective research assistance. This paper has evolved from the first part of a document entitled: "Bid-ask Option Pricing under Transactions Costs" which was first presented at an International Conference held at HEC School of Management in June 1988.

[1] See for example Cox, Ross, and Rubinstein (1979). See, however, Boyle and Vorst (1990).

cost, and any expense incurred in the process of deciding upon and placing an order. Delays in execution which cause prices at which one trades to be different from those at which one planned to trade may in some way be included as well. Recent literature on these questions includes Perold (1988), Amihud and Mendelson (1988), and Hasbrouck and Schwartz (1988).

The assumption made is that these costs are of a proportional nature, although there would be no added analytical difficulty in assuming that they are fixed in nature, or of a mixed character. Proportionality means that the cost incurred is proportional to the value of the trade.

The presence of practically any friction in financial markets qualitatively changes the nature of the optimization problem for it produces the need sometimes to do nothing and sometimes to act, an issue which, of course, does not arise in frictionless situations. Recently Constantinides (1986) has proposed an approximate solution to the portfolio choice problem under transactions costs.[2] The investor in Constantinides' problem maximizes the expected value of his infinite-horizon utility function. Portfolio policies are computed numerically under the assumption (this is the nature of the approximation) that the investor in each period consumes a fixed proportion of his wealth.

The problem studied in the present paper is somewhat simpler.[3] Our investor does not consume along the way. He accumulates wealth until some terminal point in time. At that point he consumes all. His objective is to maximize the expected utility derived from that terminal consumption. We postpone the terminal point infinitely far into the future to obtain a stationary portfolio rule. In contrast to Constantinides, our alternative leads to an exact solution.[4] The assumption that the investor does not consume as he goes and cares about the achieved terminal wealth seems quite realistic when modeling the behavior of a trader employed by a financial institution.

Our paper is organized as follows: in Section I we set up the portfolio-selection model. Section II examines the portfolio choice problem in the absence of transactions costs, which will serve as a reference. In Section III we write the necessary conditions which must be satisfied when it is optimal to refrain from trading, as well as the necessary conditions which must prevail at the random times when trading takes place. In Section IV we obtain an analytical solution. In Section V we run some comparative statics exercises. Section VI concludes and suggests applications.

[2] See also Kamin (1975), Duffie and Sun (1990), and Davis and Norman (1990) who prove existence and optimality of the solution.

[3] A similar formulation has been adopted independently in Grossman and Vila (1989). See also Taksar et al. (1988) which provides a proof of optimality in the logarithmic special cases. Fleming et al. (1990) have now produced proofs of optimality under the exact same specification which we adopt here.

[4] Our solution is not a special case of Constantinides' solution. Since he imposes the condition that the investor's rate of consumption out of wealth be constant over time, one might think that setting the rate of consumption equal to zero would produce our solution. It does not; Constantinides' solution is degenerate in that case.

I. The Portfolio Selection Model

We analyze the behavior of an investor whose planning horizon is infinite and whose objective is to maximize the expected utility of his final consumption. No consumption takes place along the way, and it is assumed that his preferences are represented by a utility function $u(c)$, where c is terminal consumption. Since consumption takes place at one point of time only, *no rate of time preference comes into the picture.*

Assumption 1: The utility function is of the power form:[5]

$$u(c) = c^{\gamma}/\gamma, \tag{1}$$

with γ constant, $\gamma < 1$.

Given this functional form, we can tentatively formulate the investor's problem as:

$$\lim_{T \to +\infty} \max E_t\left[\frac{1}{\gamma}\left[c(T)\right]^{\gamma}\right], \tag{2}$$

where E_t is the expectation operator conditional on time-t information, T is the horizon.

The economy is characterized on the real side by a unique consumption good, used as numeraire, and on the financial side by two assets:

- A riskless asset (cash), whose accumulated amount at time t, denoted $x(t)$, increases in value when trading does not take place, according to the following equation

Assumption 2:

$$\mathrm{d}x(t) = rx(t)\,\mathrm{d}t, \tag{3}$$

where r is the rate of interest, assumed to be constant;

- A risky asset with an accumulated value[6] $y(t)$ which follows—again in the absence of portfolio adjustments—the stochastic differential equation

Assumption 3:

$$\mathrm{d}y(t) = \alpha y(t)\,\mathrm{d}t + \sigma y(t)\,\mathrm{d}z, \tag{4}$$

with α and σ constants, $\mathrm{d}z$ a white noise. One may question whether such a process for asset prices would be compatible with general equilibrium in a world where financial transactions entail costs. But this issue will have to await further research. The risky asset is assumed to pay no cash dividend,

[5] This functional form implies constant relative risk aversion equal to $1 - \gamma$. The characteristics of HARA functions are listed, for example, in Merton (1971).

[6] Quantity y is the number of shares held at time t multiplied by the price of each share.

an extension to the case of dividend paying assets would be of some interest.[7]

We choose parameter values such that

$$0 < \frac{\alpha - r}{(1 - \gamma)\sigma^2} < 1, \tag{5}$$

thereby guaranteeing that x and y would be chosen to be strictly positive in the absence of transactions costs. We assume

Assumption 4: Condition (5) suffices also to guarantee $x, y > 0$ in the presence of transactions costs.

The investor acts as a trader in the markets for borrowing, for lending, and for the risky asset y. In these markets he takes prices as given and chooses quantities without restrictions but incurs transaction costs.

As far as the cost of exchanging financial assets is concerned, the following assumption is made:

Assumption 5: The conversion ratio denoted s ($s < 1$) is taken to be the same in either direction: an amount of sq^* of y can be purchased by giving up an amount q^* of x; similarly, an amount sq of x can be obtained by selling an amount of q of y.

II. Finite-Horizon vs. Infinite-Horizon Solutions in the Absence and in the Presence of Transactions Costs

If there were no transactions costs, Problem (2) could be dealt with according to Merton (1969). One would first solve it with a finite horizon T and then take the limit as $T \to +\infty$. Define the value (or Bellman) function at time t:

$$K^*(W, t; T) \equiv \max E_t\left[\frac{1}{\gamma}[W(T)]^{\gamma}\right], \tag{6}$$

where W would be the investor's wealth:[8] $W \equiv x + y$, with dynamics given by:

$$\mathrm{d}W = (rx + \alpha y)\,\mathrm{d}t + \sigma y\,\mathrm{d}z. \tag{7}$$

$K^*(W, t; T)$ must satisfy the Hamilton-Jacobi partial differential equation:

$$\max_{\substack{x,\, y \\ \text{s.t. } x+y=W}} \left\{K_t^* + K_W^*(rx + \alpha y) + K_{WW}^*\sigma^2 y^2/2\right\} \equiv 0, \tag{8}$$

with the terminal condition: $K^*(W, T; T) \equiv W^{\gamma}/\gamma$. Considering the homogeneity of the objective function (6), of the constraint (7), and of the terminal

[7] The dividend yield, if variable, would appear as an additional state variable.

[8] From now on, we will almost always write x and y instead of $x(t)$ and $y(t)$ for the asset holdings.

condition, it follows that K^* must be proportional to W^γ/γ. Writing K^* in the form: $K^*(W, t; T) \equiv H(t; T) \cdot W^\gamma/\gamma$ and substituting this function into the homogeneous differential equation (8), one obtains, by a standard procedure, the optimal portfolio strategy which is to keep y/x at a value θ^* equal to:

$$\theta^* = \frac{\alpha - r}{(1-\gamma)\sigma^2 - \alpha + r}. \tag{9}$$

One obtains also the differential equation to be satisfied by the function H:

$$-\frac{H'}{H} \equiv r\gamma - \delta^*, \tag{10}$$

where:

$$\delta^* = -\frac{1}{2}\gamma\frac{(\alpha - r)^2}{(1-\gamma)\sigma^2}. \tag{11}$$

The terminal condition for $H(t; T)$ is $H(T; T) = 1$. Equation (10) implies that the finite-horizon solution to the portfolio problem is:

$$K^*(W, t; T) = e^{(r\gamma - \delta^*)(T-t)} W^\gamma/\gamma. \tag{12}$$

Letting $T \to +\infty$ in order to address the infinite-horizon problem, observe from equation (12) that the value of K^* approaches zero or infinity.

However, using the number $\beta^* = r\gamma - \delta^*$ as a discount rate, define the discounted value function:

$$J^*(W, t; T) \equiv e^{-\beta^*(T-t)} K^*(W, t; T), \tag{13}$$

which solves the modified Hamilton-Jacobi equation:

$$\max_{\substack{x,\, y \\ \text{s.t. } x+y=W}} \left\{ J_t^* - \beta^* J^* + J_W^*(rx + \alpha y) + J_{WW}^* \sigma^2 y^2/2 \right\} \equiv 0, \tag{14}$$

with the same terminal condition $J^*(W, T; T) \equiv W^\gamma/\gamma$ as before. The solution for J^* is evidently $J^*(W, t; T) \equiv W^\gamma/\gamma$, whether or not T is driven to infinity. Discounting at the rate β^* has rendered the value function stationary (i.e., independent of time t) and therefore finite when $T \to +\infty$. Furthermore, $\beta^* = r\gamma - \delta$ is the only value of the discount rate for which there exists a stationary solution $J^*(W, t; T) \equiv J^*(W)$ $[J_t^* \equiv 0]$ of the differential equation (14).

Made wiser by the solution of the investment problem without transactions costs, *consider now the same portfolio problem with transactions costs.* The first step is to define the *finite-horizon* value function of the problem:

$$K(x, y, t; T) \equiv \max E_t\left[\frac{1}{\gamma}[c(T)]^\gamma\right], \tag{15}$$

subject to an appropriate modification of equations (3) and (4) incorporating trading activity.[9] Inside a domain to be described at length in Section III,

[9] See equations (19) and (20) below.

this function K must satisfy the Hamilton-Jacobi equation:[10]

$$K_t + K_x rx + K_y \alpha y + K_{yy} y^2 \sigma^2/2 \equiv 0 \tag{16}$$

and a terminal condition such as: $K(x, y, T; T) \equiv (x + sy)^\gamma/\gamma$. Such a terminal condition reflects the idea that, at the horizon point, the investor would turn all his assets into the riskless one (cash) in order to consume out of it.

We can surmise with some assurance that the function K does not reach a finite value as $T \to +\infty$. However, consider a discount rate β and the corresponding discounted value function, which would be defined as:

$$\begin{aligned} J(x, y, t; T) &\equiv e^{-\beta(T-t)} K(x, y, t; T) \\ &\equiv e^{-\beta(T-t)} \max E_t\left[\frac{1}{\gamma}\left[c(T)\right]^\gamma\right]. \end{aligned} \tag{17}$$

The function J would be a solution of the following modified Hamilton-Jacobi equation:

$$-\beta J + J_t + J_x rx + J_y \alpha y + J_{yy} y^2 \sigma^2/2 \equiv 0. \tag{18}$$

Assume now

Assumption 6: As was true in the absence of costs, there does exist a number β (to be determined) such that $J(x, y, t; T)$ reaches a well-defined limit, noted $J(x, y)$ as $T \to +\infty$. Regard the unknown function $J(x, y)$ as the solution to the infinite-horizon portfolio problem.

It is very important to realize that the number β is not a given parameter. It is not a behavioral parameter reflecting some dimension of the preferences of the investor (such as his subjective discount rate). Indeed, one look at the objective function (17) indicates that, under a finite horizon T, the choice of β would be irrelevant to the determination of optimal portfolio behavior since the discount rate is a constant relative to the portfolio decisions.

Instead, the number β is to be determined endogenously, from the solution of the infinite horizon problem with transactions costs, for given rate of return parameters, transactions costs, and risk aversion parameter. Discounting utility at the rate of β is only a technical device which serves to keep the objective finite.

While every element of the family $\{J(x, y, t; T);\ T \in R_+,\ T > t\}$ satisfies equation (18), it does not necessarily follow that the limit $J(x, y)$ satisfies this equation as well. In order to avoid technicalities, we simply assume

Assumption 7: The infinite-horizon solution $J(x, y)$, which we are seeking is a solution of (18), and it is the only stationary solution ($J_t \equiv 0$) of that equation.

If that is true, the search for the limit is equivalent to the search for the stationary solution of (18).

[10] As we will see, boundary conditions apply on the frontier of this domain. See equations (21) and (22) below.

In what follows, our goal is to obtain the stationary solution $J(x, y)$ of (18) along with a number β which causes $J(x, y)$ to exist.

III. Regulating the Portfolio

When trading costs of the type outlined above (*Assumption* 5) *are present*, Constantinides (1979) has shown in a discrete-time framework and assumed in a continuous-time one that the optimal investment policy is of the following type: instantaneously exchange a *small amount* of the riskless for the risky asset whenever the portfolio ratio $\theta \equiv y/x$ falls below level l to be chosen optimally, and do the reverse if it exceeds the control level u. That the optimal trade size in continuous time should be infinitely small is a postulate that makes sense considering the proportional nature of the transactions costs. It means that the optimal regulator of the portfolio belongs to the class of instantaneous regulators.

We intend to verify that, among instantaneous regulators, one which would be optimal would involve fixed control barriers u and l. We also intend to determine the level of these barriers by a heuristic argument. This is done below by means of the theory of optimal regulated Brownian motion, as exposited by Harrison (1985).[11]

Portfolio rebalancing is modeled mathematically as follows. Define two processes U and L with these properties:

- U and L are continuous non decreasing processes;
- U increases only when θ takes the value of u;
- L increases only when θ takes the value of l.

In order to incorporate portfolio readjustments, equations (3) and (4) are amended as:

$$\mathrm{d}x = rx\,dt + s\,\mathrm{d}U - \mathrm{d}L, \tag{19}$$

$$\mathrm{d}y = \alpha y\,\mathrm{d}t + \sigma y\,\mathrm{d}z + s\,\mathrm{d}L - \mathrm{d}U, \tag{20}$$

where $1 - s$ is the rate of transactions costs and U and sL are interpreted as cumulative sales and purchases of equity since time 0. The stochastic differential equations (19) and (20) reduce to (3) and (4) at such times when no rebalancing takes place ($\theta \neq u$ or l).

The discounted value function $J(x, y)$ of the optimization problem, defined in Section II, is homogeneous of degree γ and is assumed

[11] This approach is very much related to the optimal-stopping literature (Krylov (1980)) which has already been applied to the field of finance by Grossman and Laroque (1990). The mathematical theory of the instantaneous control of Brownian motion can be found in Benes, Shepp, and Witsenhausen (1980) and Harrison and Taksar (1983). Regulated Brownian motion has been applied to problems in economics by Bertola (1988), Dixit (1989a, b), and Dumas (1988, 1990a).

Assumption 8: $J(x, y)$ is twice continuously differentiable in y over the domain: $\ln < y - ux$ and once continuously differentiable in x.

A transposition of Harrison's[12] results to the present setting leads to the following properties of the value function $\dot{J}$:[13]

$$J_x(x, y) = sJ_y(x, y) \qquad \text{when } \theta \equiv y/x = l, \text{ and} \tag{21a}$$

$$sJ_x(x, y) = J_y(x, y) \qquad \text{when } \theta \equiv y/x = u. \tag{21b}$$

Conditions (21a) and (21b) mean that the indirect utility function must present a marginal rate of substitution between y and x equal to $1/s$ when $y = lx$, and equal to s when $y = ux$. They are a mechanical result of the postulate that the optimal policy is of the instantaneous-regulator type; i.e., they hold the minute regulation is applied at u and l, regardless of whether these quantities are chosen optimally or not.

Considering the homogeneity of the function J, properties (21a) and (21b) are satisfied along two rays $y = lx$ and $y = ux$ whose slopes are independent of y and x.

The choice of the control limits u and l will be made by means of two necessary conditions labeled "Super-contact conditions" by Dumas (1990b):

$$0 = -J_{xx}(x, y) + sJ_{xy}(x, y) = -sJ_{yx}(x, y) + s^2 J_{yy}(x, y) \qquad \text{at } \theta \equiv y/x = l, \text{ and} \tag{22a}$$

$$0 = s^2 J_{xx}(x, y) - sJ_{xy}(x, y) = sJ_{yx}(x, y) - J_{yy}(x, y) \qquad \text{at } \theta \equiv y/x = u. \tag{22b}$$

Similar conditions were obtained by Harrison and Taksar (1983) for a linear one-dimensional special case.[14] They are extensions to the case of the instantaneous regulator of the traditional "smooth-pasting conditions" of Samuelson (1965), McKean (1965), Merton (1973, footnote 60) and Krylov (1980). A footnote provides a heuristic argument in favor of conditions (22a, b).[15] Formally, what will ultimately be needed is a "Verification Theorem" along the lines of Benes, Shepp, and Witsenhausen (1980), Taksar et al.

[12] See Harrison (1985, Chapter 5). See also the notes Dumas (1990b) and Dixit (1990).

[13] When $\theta = l$, x is reduced by $\mathrm{d}L$ and y is increased by $s\mathrm{d}L$. When this happens, there can be no jump in the value of the problem ("Value matching"):

$$J(x, y) = J(x - dL, y + s\,\mathrm{d}L).$$

Expanding the right-hand side leads to Equation (21a).

[14] Proposition 5.11, page 449. See also the discussion which precedes proposition 5.11.

[15] Smooth pasting says roughly that the derivatives of the value function must take the same value at the point where the regulator is applied (trigger point) and at the point to which one arrives as a result of the regulation (target point). For instance, when $\theta = l$, x is reduced by $\mathrm{d}L$ and y increased by $s\,\mathrm{d}L$:

$$J_x(x, y) = sJ_y(x, y) = J_x(x - \mathrm{d}L, y + s\,\mathrm{d}L) = sJ_y(x - \mathrm{d}L, y + s\,\mathrm{d}L).$$

Expanding terms directly leads to condition (22a).

(1988), and Davis and Norman (1990) stating that a solution to the partial differential equation (23) below and subject to boundary conditions (21) and (22) is the value function for the problem. In the present work we regard our proposed solution as a candidate for optimality.[16]

On the basis of steady-state Assumption 7, we restrict our search to the stationary solution, defined as the solution for which the value function J is independent of time and the control limits l and u are independent of time as well, in addition to being independent of y and x. It is then possible to provide an explicit solution for $J(x, y)$ over the domain $lx \leq y \leq ux$; this will be done in the next section.[17]

Inside the cone, the Hamilton-Jacobi equation for J—reflecting the conditionally expected drift of the value of J when x and y obey equations (3) and (4)—is (see equation (18) and Assumption 7):

$$-\beta J + J_x r x + J_y \alpha y + J_{yy} y^2 \sigma^2 / 2 \equiv 0 \qquad lx \leq y \leq ux. \tag{23}$$

The boundary conditions applying to equation (23) are those on the marginal rates of substitution and on the second derivatives written above as Conditions (21) and (22).

IV. Analytical Solution for the Candidate Value Function

Exploit the homogeneity of degree γ of the value function in order to state that:

$$J(x, y) \equiv x^{\gamma} I(\theta), \tag{24}$$

where θ stands for $\theta(t)$, with $\theta(t) \equiv y(t)/x(t)$. Substituting (24) into equation (23), it is found that the I function must satisfy the following reduced Hamilton-Jacobi equation:

$$\delta I + \theta(\alpha - r) I'(\theta) + \frac{1}{2}\theta^2 \sigma^2 I''(\theta) \equiv 0, \qquad l \leq \theta \leq u, \tag{25}$$

where δ stands for:

$$\delta = -\beta + r\gamma.$$

We reserve till later the choice of an appropriate value for β (and δ).

[16] The needed theorems and proofs are now available in Fleming et al. (1990).

[17] Outside the cone the indirect utility function is a power function (with γ exponent) of $x + sy$ or $y + sx$—depending on the side being considered. This indirect utility function is relevant only if initial conditions place the portfolio outside the cone. In that case the portfolio mix is instantaneously steered back inside the cone; during the later course of the optimal program, the portfolio will never be positioned outside the cone.

The boundary conditions (21) and (22) transform into:

$$I'(l) = \frac{\gamma I(l)}{1+\frac{l}{s}}\frac{1}{s}; \qquad I'(u) = \frac{\gamma I(u)}{1+us}s; \qquad (26a); (26b)$$

$$I''(l) = \frac{(\gamma-1)I'(l)}{1+\frac{l}{s}}\frac{1}{s}; \qquad I''(u) = \frac{(\gamma-1)I'(u)}{1+us}s. \quad (27a); (27b)$$

A change of variable from θ to $ln(\theta)$ transforms equation (25) into a differential equation with constant coefficients. The corresponding characteristic equation is:

$$0 = \delta + k\left(\alpha - r - \sigma^2/2\right) + k^2\sigma^2/2, \tag{28}$$

where k is an unknown; the solutions (real or imaginary) are called k_1 and k_2. The discriminant of (28) is:

$$\Delta = \left(\alpha - r - \sigma^2/2\right)^2 - 2\delta\sigma^2. \tag{29}$$

Define the critical value of δ_c of δ for which the discriminant Δ changes sign:

$$\delta_c = \frac{\left(\alpha - r - \sigma^2/2\right)^2}{2\sigma^2} \tag{30}$$

If later we are led to choose $\delta < \delta_c$, this will cause the discriminant Δ to be positive and the roots k_1 and k_2 to be real. If, to the opposite, we choose $\delta > \delta_c$, the roots will be imaginary.

The general solution of (25) is:

$$I(\theta) = C_1\theta^{k_1} + C_2\theta^{k_2}, \tag{31}$$

where C_1 and C_2 are two integration constants to be chosen in such a way as to satisfy the boundary conditions (26) and (27).

It may be more convenient to avoid imaginary numbers. The usual groupings defining trigonometric functions yield:

$$I(\theta) = \theta^{-\lambda}\{A\,\mathrm{si}[\nu ln(\theta)] + B\,\mathrm{co}[\nu ln(\theta)]\}, \tag{32}$$

where:

$$\lambda = \frac{\alpha - r - \sigma^2/2}{\sigma^2}, \quad \nu = \frac{\sqrt{|\Delta|}}{\sigma^2},$$

and where si[ ] and co[ ] stand for the sine (resp. hyperbolic sine) and cosine (resp. hyperbolic cosine) functions used when Δ is negative (resp. positive). A and B are two new constants of integration.

The borderline case where $\delta = \delta_c$, or $\Delta = 0$, —which is unlikely to obtain—would generate a different kind of solution:

$$I(\theta) = C_1\theta^k + C_2 ln(\theta)\theta^k, \tag{33}$$

where k would be the single root of equation (28).

The four boundary conditions, (26a, 26b, 27a, and 27b) would appear to be all that is needed to solve the four unknowns A, B, l, and u. In fact, as is evident from the conditions (26) and (27) and from the form of the general solution (32), the corresponding system of equations is homogeneous in A and B. In order to avoid the trivial solution $A = B = 0$, some determinant must be set equal to zero. This will provide the *needed restriction* on the as yet arbitrary constant δ (see below, equation (36)). The two unknowns A and B will remain determined only up to an arbitrary factor.[18]

It is easy, *in order to obtain an equation system for l and u*, to eliminate A and B between the pair of equations (26a) and (27a) on the one hand and the pair (26b) and (27b) on the other. But the same equation system can be obtained by substituting the boundary conditions into the partial differential equations. Defining, for convenience:

$$\varepsilon_l = \frac{l/s}{1 + l/s}, \quad \varepsilon_u = \frac{us}{1 + us}, \tag{34}$$

we obtain, actually, the same second-degree equation to be satisfied by ε_u and by ε_l:

$$0 = \delta + \gamma\varepsilon(\alpha - r) + \frac{1}{2}\gamma(\gamma - 1)\sigma^2\varepsilon^2. \tag{35}$$

The discriminant of this equation must be positive, as there can be no physical interpretation to portfolio control limits being complex numbers.[19] The lower of the two real roots is interpreted as ε_l and the higher one as ε_u. The control limits l and u are then obtained from ε_l and ε_u by inverting the definitions (34).

At this stage, we have been able to determine, for any given value of the constant δ, the form of the general solution which depends on two unknown constants A and B, as well as the corresponding values of l and u. We must now choose δ in such a way that there exists a nonzero solution for A and B. This is accomplished by writing that the determinant of the linear system made up of the two boundary conditions (26) is equal to zero. Rewriting (26a) and (26b) in the form:

$$lI'(l) = \gamma I(l)\varepsilon_l, \quad uI'(u) = \gamma I(u)\varepsilon_u,$$

[18] As in the frictionless case, the value of that factor would fall out of the terminal condition of the finite-horizon problem as one postpones the horizon to infinity (see Section II). The procedure is not implementable in the case where transactions costs are present since we do not have an explicit solution for the finite-horizon problem. Fortunately, the missing factor has no bearing on the determination of the control limits l and u.

[19] This restriction sets an upper bound on the unknown value of δ. This upper bound turns out to be equal to δ^*, defined in equation (11), which is the value of δ that applies in the frictionless case. One is not surprised to find that the value of δ should always be on one side of the frictionless value.

we obtain the following entries for this determinant:[20]

$$\left|\begin{array}{c:c} -\lambda\,\mathrm{si}[\nu\,ln(l)] + \nu\,\mathrm{co}[\nu\,ln(l)] & -\lambda\,\mathrm{co}[\nu\,ln(l)] - \nu\,\mathrm{si}[\nu\,ln(l)] \\ -\gamma\,\mathrm{si}[\nu\,ln(l)]\varepsilon_l & -\gamma\,\mathrm{co}[\nu\,ln(l)]\varepsilon_l \\ \hdashline -\lambda\,\mathrm{si}[\nu\,ln(u)] + \nu\,\mathrm{co}[\nu\,ln(u)] & -\lambda\,\mathrm{co}[\nu\,ln(u)] - \nu\,\mathrm{si}[\nu\,ln(u)] \\ -\gamma\,\mathrm{si}[\nu\,ln(u)]\varepsilon_u & -\gamma\,\mathrm{co}[\nu\,ln(u)]\varepsilon_u \end{array}\right| \tag{36}$$

Equating this determinant to zero generally produces two roots. One of them is the trivial solution $\delta = \delta_c$ which should be disregarded. It should be disregarded on the grounds that, when $\delta = \delta_c$, the solution is of the type (33) not (31). Such a solution would generate a determinant different from (36). In order for this other determinant to be equal to zero precisely for the value $\delta = \delta_c$, a very special combination of parameter values would have to be chosen. The analysis of this borderline situation is uninteresting.

Later research will seek to prove that equating the determinant (36) to zero necessarily produces a second, nontrivial solution for δ. The nontrivial root can be found by numerical iteration. Figure 1 depicts the value of the determinant as a function of δ for the following *numerical base case*:[21]

$$\alpha - r = 0.05; \quad \gamma = -1; \quad \sigma^2 = 0.04. \tag{37}$$

This base case, in the absence of transactions ($s = 1$), would produce an optimal policy $\theta^* = 1.67$ (or a weight on the risky asset equal to $\theta^*/(1 + \theta^*) = 0.625$) and a value of $\delta^* = 0.015625$.

Under 10% transactions costs ($s = 0.9$), the root is found to be in the imaginary domain ($\Delta < 0$; $\delta_c < \delta < \delta^*$; $\delta_c = 0.01125$):

$$\delta = 0.014713; \quad l = 0.811094; \quad u = 3.848747,$$

implying a weight in the portfolio allocated to the risky asset fluctuating between $l/(1 + l) = 0.44785$ and $u/(1 + u) = 0.79376$, which is a tolerance zone of 27% on the upper side and 28% on the lower side of the frictionless optimum.

V. Comparative Analysis

We now examine deviations from the numerical base case (37) in the dimensions of increasing transactions costs, increasing risk aversion, and increasing risk.

A. Varying the Size of Transactions Costs

Not surprisingly, as illustrated in Figure 2, increasing the size of transactions costs (lowering the parameter s) widens the region of no transactions.

[20] The determinant is shown here correctly for the case of imaginary roots ($\Delta < 0, \delta > \delta_c$). Minor sign differences would appear in the case of real roots.

[21] We choose here the same base case as in Constantinides (1986).

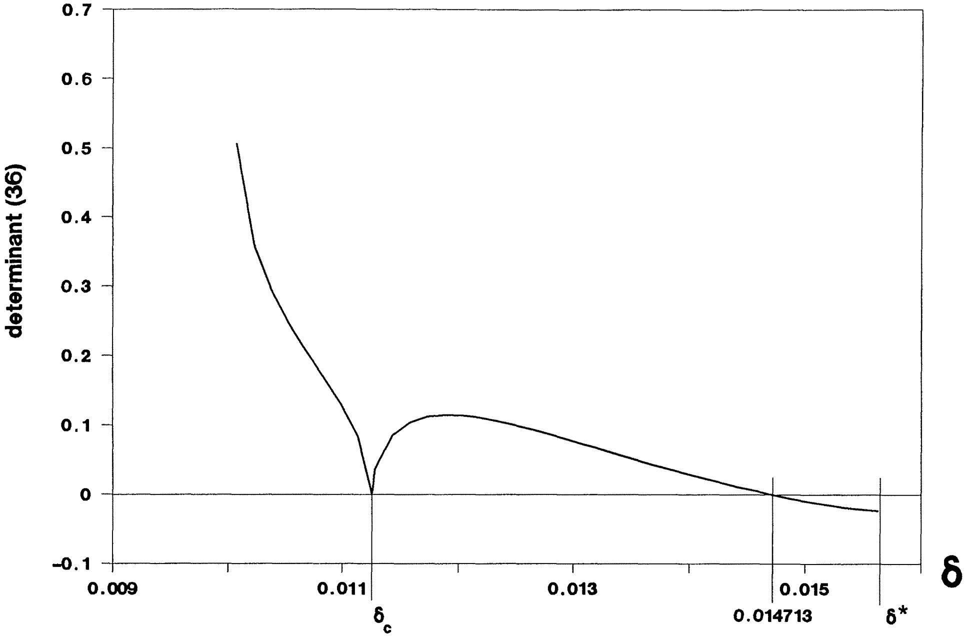

Figure 1. Locating the root of the system. The value of the endogenous discount rate β ($\beta = \gamma r - \delta$) which makes the value function stationary is determined numerically by equating the determinant (36) to zero. The solution is found to be: $\delta = 0.014713$. δ^* is the value of δ corresponding to the costless case. δ_c is the critical value of δ separating real roots (on the left) from imaginary roots. Numerical values of parameters are: excess return on the risky asset $(\alpha - r) = 5\%$; risk aversion = 2 ($\gamma = -1$); variance of return on risky asset = 0.04 ($\sigma = 20\%$); transactions costs: 10% ($s = 0.9$).

This is especially true for low transactions costs; even a value of s extremely close to 1 already causes portfolio control limits to separate appreciably. The reason for this phenomenon is well known; because the Brownian motion is an infinite variation process, costs accompanying frequent rebalancing, even levied at a small rate, would quickly outweigh the benefits of precisely optimal diversification. The applicable *caveat* is, of course, that there is no reason to believe in the first place that asset prices should follow a process with unbounded variation when trading is subject to transactions costs.

At all levels of transactions costs, the region of no transactions is markedly wider than that obtained by Constantinides (1986).[22] The difference between the two results is shown in Figure 2. Also, there is no tendency for increased transactions costs to bias the portfolio one way or the other. As is quite visible in Figure 2, this result contrasts with that of Constantinides who

[22] A caveat is in order concerning this comparison with Constantinides. His formulation of the optimization problem involved a discount rate for utility. This parameter is absent here. It seems unlikely that a change in the value of this parameter would reverse the results of the comparison between the two approaches.

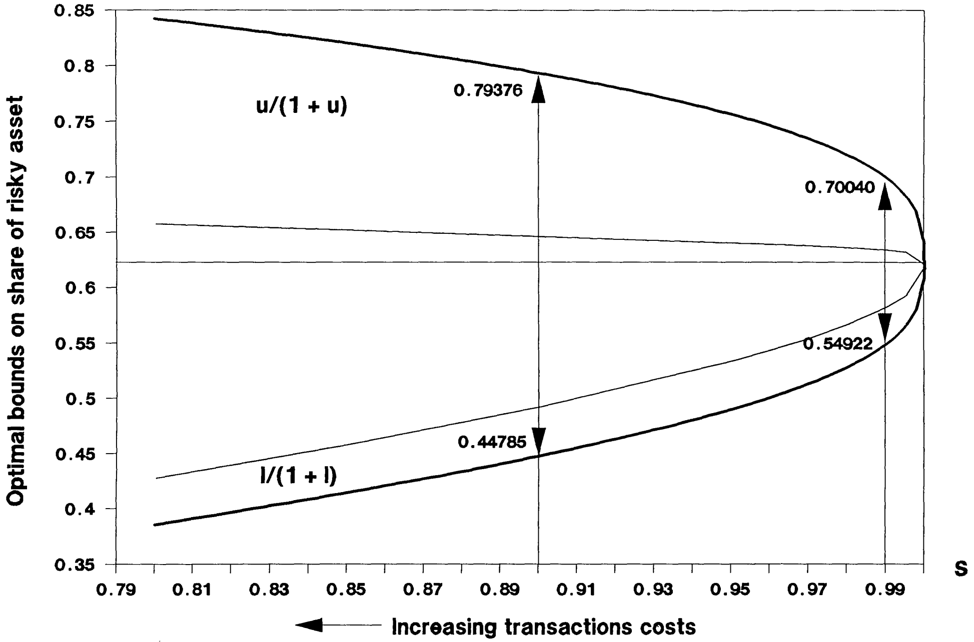

Figure 2. The effect of transactions costs on portfolio choice. The optimal bounds placed on the value-share of the risky asset in the portfolio are shown vertically by the thickly drawn line. For instance, at 10% transactions costs ($s = 0.9$) the optimal policy is to allow the portfolio share of the risky asset to move freely between $l/(1 + l) = 44.755\%$ and $u/(1 + u) = 79.376\%$ (u and l are the limit ratios of risky over riskless asset shares). The thinly drawn line represents similar quantities obtained by Constantinides (1986). Numerical parameter values: excess return on the risky asset ($\alpha - r$) = 5%; risk aversion = 2 ($\gamma = -1$); variance of return on risky asset = 0.04 (σ = 20%); transactions costs: 10% ($s = 0.9$).

found that "transactions costs shift the region of no transactions towards the riskless asset". Both differences in results can be ascribed to the fact that, in Constantinides' model, consumption was taking place along the way and not just at the terminal time and, furthermore, that consumption came out of the riskless asset only. When a steady flow of consumption expenditures must be met out of the existing cash on hand, there is both less room for fluctuations in the amount of cash available and more need to bias the portfolio in favor of cash than when consumption is postponed to infinity.

B. Varying Risk Aversion

Figures 3 and 4 give two similar views of the way in which increasing risk aversion affects portfolio choices. Basically, there is very little interaction between risk aversion and transactions costs. Figure 3 provides a measure of the width of the portfolio weight band [$u/(1 + u)$, $l/(1 + l)$] as a function of risk aversion. It is apparent that the width of that band is nearly constant, in fact slowly increasing as one scans risk aversions ranging from 1.6 to 2.7.

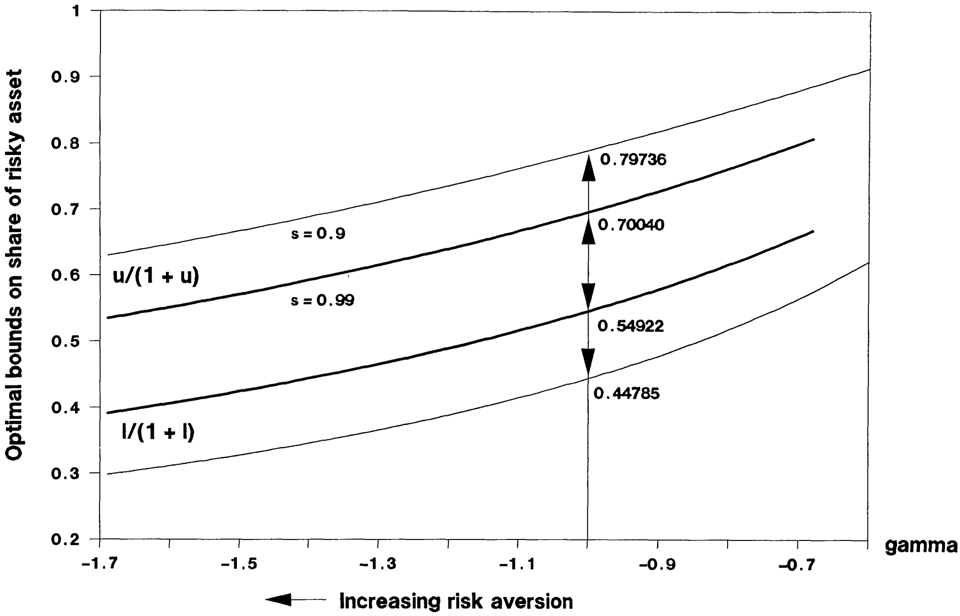

Figure 3. The effect of risk aversion on portfolio choice. The optimal bounds placed on the value-share of the risky asset in the portfolio are shown vertically by the thinly drawn line for the case of 10% transactions costs ($s = 0.9$) and by the thickly drawn line for the case of 1% transactions costs ($s = 0.99$). For instance, at a relative risk aversion of two ($\gamma = -1$), with 10% transactions costs, the optimal policy is to allow the portfolio share of the risky asset to move freely between $l/(1 + l) =$ 44.755% and $u/(1 + u) =$ 79.376%. u and l are the limit ratios of risky over riskless asset shares.

A similar message can be read in Figure 4, which controls for the effect of risk aversion on what would have been the optimal frictionless portfolio (see equation (9)); $\frac{u/(1+u)}{\theta^*/(1+\theta^*)}$ and $\frac{l/(1+l)}{\theta^*/(1+\theta^*)}$ are slowly increasing again as one increases risk aversion.

These results are qualitatively the same as those of Constantinides (1986).

C. Increasing Risk

Broadly, the same conclusions are reached for an increase in risk as for an increase in risk aversion. Figure 5 shows that the width of the band $[u/(1 + u), l/(1 + l)]$ is approximately constant around the optimal frictionless value, as σ^2 is gradually increased.

VI. Conclusion and Applications

The exact solution to the portfolio management problem which we have obtained is in the form of two control barriers. These set the maximum and minimum allowable degrees of imbalance in the portfolio which will be

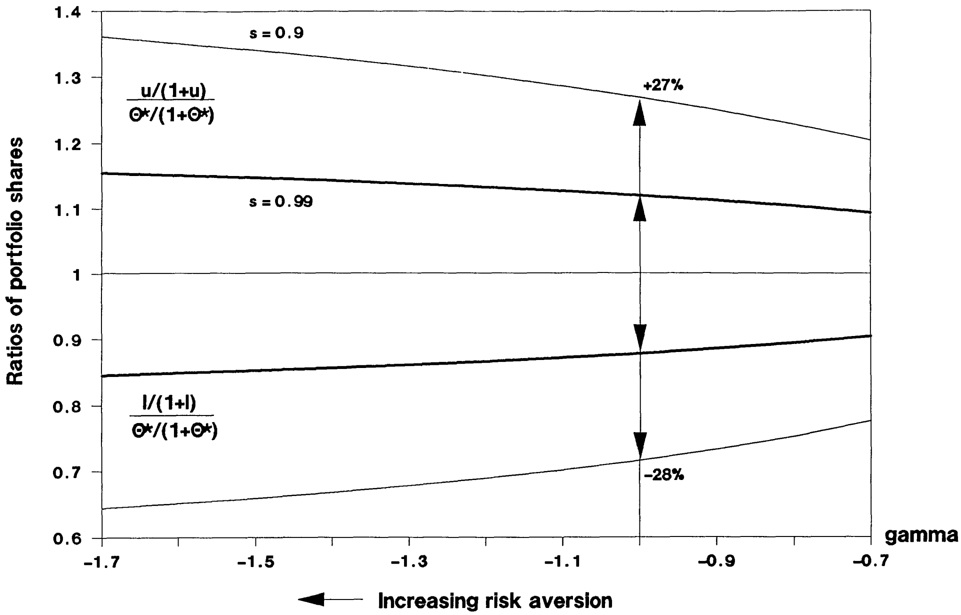

Figure 4. The effect of risk aversion on portfolio choice relative to the frictionless case. The optimal bounds to be placed on the value-share of the risky asset are shown after dividing by the optimal value-share in the costless case. The thinly and thickly drawn lines correspond to the cases of 10% and 1% transactions costs, respectively. For instance, at a relative risk aversion 2 ($\gamma = -1$), with 10% transactions costs ($s = 0.9$), the optimal policy is to allow the portfolio share of risky asset to go as high as 27% above, and as low as 28% below its costless optimal value. u and l are the limit ratios of risky over riskless asset shares; θ^* is optimal ratio of risky over riskless asset in the absence of transactions costs (Merton case).

tolerated before any action is taken. The position of these barriers was computed numerically. The candidate policy identified here calls for wider bands of tolerated imbalance in the portfolio than had been suggested by Constantinides (1986). Also, the band is not biased in favor of cash, the more so as one increases transactions costs as was true in Constantinides' model. The comparative static analysis, however, yields conclusions broadly similar to those of Constantinides. Ours are exact results while previous work has only produced approximate solutions.

Several major applications and extensions of our exact solution can be envisaged. The first would be a contribution to the economics of the dealer function.[23] The investor considered here trades infrequently and in small quantities. For a price, however, he should be willing to absorb (positively or negatively) finite chunks of the risky asset when a customer comes by who desires to make a finite trade. These prices at which he would be willing to

[23] On this count see Treynor (1981), Ho and Stoll (1981), and Treynor (1987).

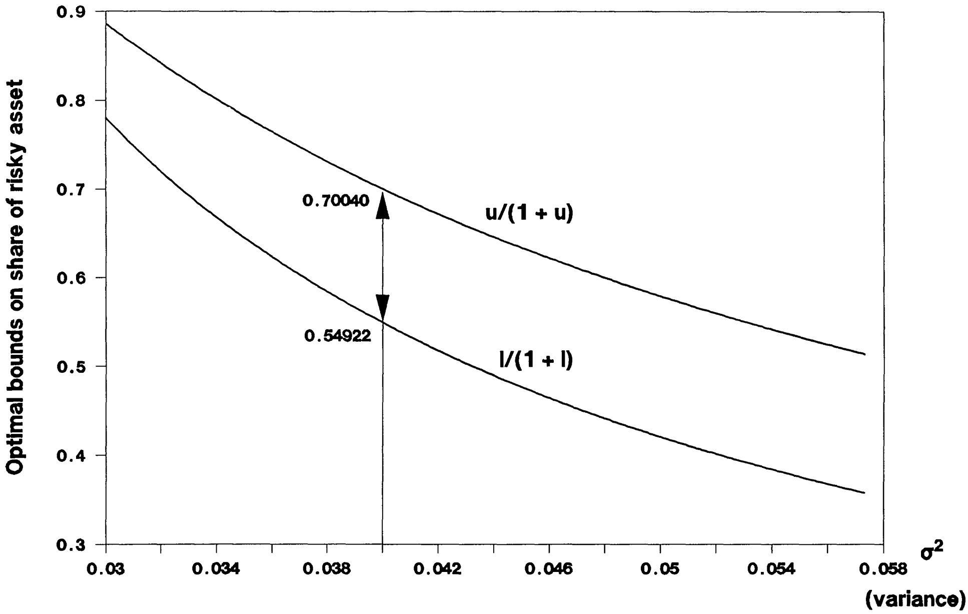

Figure 5. The effect of risk on portfolio choice. The optimal bounds placed on the value-share of the risky asset in the portfolio are shown vertically. For instance, at a level of variance equal to 0.04 ($\sigma = 20\%$), with 1% transactions costs ($s = 0.99$), the optimal policy is to allow the portfolio share of the risky asset to move freely between $l/(1 + l) = 0.54922$ and $u/(1 + u) = 0.70040$. u and l are the limit ratios of risky over riskless asset shares.

trade finite quantities would be his bid- and ask-prices. In this theory the determinant of the bid-ask spread would be purely the instantaneous volatility of returns, the size of transactions costs, and the inventory of the dealer. A missing determinant, which was examined by Ho and Stoll (1981) would be the rate of random customer arrivals. There is hope that this added variable can also be incorporated in the optimization plan.

A second application would be an extension of the application just mentioned. Ultimately it must pay for the customer to place his order with the dealer rather than transact directly in the market. Also, dealers must be allowed to trade with each other. Ultimately the equilibrium dynamics of the price in the presence of frictions must be determined, rather than postulated as they were here.

A third application of this exact solution would be to price derivative assets in the presence of transactions costs, when investors adopt optimal portfolio strategies, in continuous time. Our objective would be to find the bid- and ask-prices of a European call option in an intertemporal setting.

As has been amply demonstrated by Figlewski (1989), it is not true, in the presence of transactions costs, that a replicating argument would provide the right option price.[24] Instead it would be appropriate to analyze the portfolio strategy of an investor who acts as a trader (and pays transactions costs) on the primitive asset market and as a dealer in the option market.

[24] Recently, Leland (1985) has proposed a replicating strategy for option pricing which implies finite transaction costs and still generates, with probability one, a payoff equal to that of the option. Unfortunately, the frequency of portfolio revisions in his model is exogenously given instead of being optimally chosen with respect to transactions costs. As a consequence, his formulation does not match the optimal investment strategy with proportional transactions costs which has been outlined by Constantinides (1979, 1986). Merton (1989) also studies the problem of option pricing with transactions costs; he formulates a two-period replicating strategy in a binomial context but does not extend it to an arbitrary number of periods and, consequently, is not in a position to determine the limiting value of the option price when time is allowed to become continuous. But see Boyle and Vorst (1990). At any rate, pricing by replication is questionable in the current context; when transactions costs are present, an exact replication is not generally the most efficient method of manufacturing an option. Following the purchase or sale of an option, an intermediary would not typically choose to fully offset that transaction by means of sales and purchases of the underlying asset. A strategy of full offset would be unnecessarily costly and would not be pursued. It cannot consequently provide the price of the option under pure competition; it can only provide a lower or an upper bound on the price. In our view the option cannot be priced outside an optimal portfolio-investment framework.

REFERENCES

Amihud, Y. and H. Mendelson, 1988, Liquidity and asset prices: Financial management implications, *Financial Management* Spring, 5-15.

Benes, V. E., L. A. Shepp, and H. S. Witsenhausen, 1980, Some solvable stochastic control problems, *Stochastics* no. 4, 39-83.

Bertola, G., 1988, Irreversible investment, Working paper, Princeton University.

Black, F. and M. Sholes, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy* 81, 637-654.

Boyle, P. P. and T. Vorst, 1990, Option pricing in discrete time with transactions costs, Working paper, Erasmus University.

Constantinides, G. M., 1979, Multiperiod consumption and investment behavior with convex transactions costs, *Management Science* 25, 1127-1137.

——, 1986, Capital market equilibrium with transactions costs, *Journal of Political Economy* 94, 842-862.

Cox, J. C., S. A. Ross, and M. Rubinstein, 1979, Option pricing: A simplified approach, *Journal of Financial Economics* 7, 229-263.

Davis, M. H. A. and A. R. Norman, 1990, Portfolio selection with transactions costs, *Mathematics of Operations Research*, Forthcoming.

Dixit, A., 1989a, Entry and exit decisions under uncertainty, *Journal of Political Economy* 97, 620-638.

——, 1989b, Hysteresis, import penetration, and exchange rate pass-through, *Quarterly Journal of Economics* CIV, 205-228.

——, 1990, A simplified treatment of some results concerning regulated Brownian motion, *Journal of Economic Dynamics and Control*, Forthcoming.

Duffie, D. and T. Sun, 1990, Transaction costs and portfolio choice in a discrete-continuous time setting, *Journal of Economic Dynamics and Control* 14, 35-51.

Dumas, B., 1988, Perishable investment and hysteresis in capital formation, Working paper, University of Pennsylvania.

——, 1990a, Dynamic equilibrium and the real exchange rate in a spatially separated world, *Review of Financial Studies*, Forthcoming.

——, 1990b, Super contact and related optimality conditions, *Journal of Economic Dynamics and Control*, Forthcoming.

Figlewski, S., 1989, Options arbitrage in imperfect markets, *Journal of Finance*, 44, 1289–1312.

Fleming, W. H., S. G. Grossman, J.-L. Vila, and T. Zariphopoulou, 1990, Optimal portfolio rebalancing with transactions costs, Working paper, Brown University.

Grossman, S. J. and J.-L. Vila, 1989, Optimal dynamic trading with leverage constraints, Working paper, University of Pennsylvania.

—— and G. Laroque, 1990, Asset pricing and optimal portfolio choice in the presence of illiquid durable consumption goods, *Econometrica*, 58, 25–52.

Harrison, J. M. and M. I. Taksar, 1983, Instantaneous control of Brownian motion, *Mathematics of Operations Research* 8, 439–453.

——, 1985, *Brownian Motion and Stochastic Flow System*, (John Wiley & Sons, Somerset, NJ).

Hasbrouck, J. and R. A. Schwartz, 1988, Liquidity and execution costs in equity markets, *The Journal of Portfolio Management* Spring, 10–16.

Ho, T., Stoll, H., 1981, Optimal dealer pricing under transactions and returns uncertainty, *Journal of Financial Economics* 9, 47–73.

Kamin, J. H., 1975, Optimal portfolio revision with a proportional transactions costs, *Management Science* 21, 1263–1271.

Krylov, N. V., 1980, *Controlled Diffusion Processes*, (Springer Verlag, New York).

Leland, H., 1985, Option pricing and replication with transactions costs, *The Journal of Finance* 40, 1283–1031.

Magill, M. J. P. and G. Constantinides, 1976, Portfolio selection with transactions costs, *Journal of Economic Theory* 13, 245–263.

McKean, H. P., Jr., 1965, Appendix: A free boundary problem for the heat equation arising from a problem in mathematical economics, *Industrial Management Review* 6, 32–39.

Merton, R. C., 1969, Lifetime portfolio selection under uncertainty: The continuous-time case, *The Review of Economics and Statistics* 51, 247–257.

——, 1971, Optimum consumption and portfolio rules in a continuous-time model, *Journal of Economic Theory* 3, 373–413.

——, 1973, The theory of rational option pricing, *The Bell Journal of Economics and Management Science* 4, 141–183.

——, 1989, On the application of the continuous-time theory of finance to financial intermediation and insurance, *The Geneva Papers in Insurance* 14, 225–261.

Perold, A. F., 1988, The implementation shortfall: Paper versus reality, *The Journal of Portfolio Management* Spring 4–9.

Samuelson, P. A., 1965, Rational theory of warrant pricing, *Industrial Management Review* 6, 41–50.

Taksar, M., M. J. Klass, and D. Assaf, 1988, A diffusion model for optimal portfolio selection in the presence of brokerage fees, *Mathematics of Operations Research* 13, 277–294.

Treynor, J. L., 1981, What does it take to win the trading game? *Financial Analysts Journal* January -February, 55–60.

——, 1987, The economics of the dealer function, *Financial Analysts Journal* November–December, 27–34.