OPTIMAL INVESTMENT AND CONSUMPTION WITH TRANSACTION COSTS PART II: TWO BONDS

by

S. E. Shreve & H. M. Soner
Department of Mathematics
Carnegie Mellon University
Pittsburgh, PA 15213

G.-L. Xu
Global Portfolio Research Department
Daiwa Securities Trust Company
Jersey City, NJ 07302

Research Report No. 90-93 a

November 1990

OPTIMAL INVESTMENT AND CONSUMPTION WITH TRANSACTION COSTS PARTI: TWO BONDS

by

S.E. Shreve*,
H.M. Soner»,
and
G.-L. Xu+

June 1990

•Department of Mathematics, Carnegie Mellon University, Pittsburgh, PA 15213

•Global Portfolio Research Department, Daiwa Securities Trust Company, Jersey City, NJ 07302

Work supported by the National Science Foundation under grant DMS-90-02588.

Abstract

An agent can invest in a high–yield bond and a low–yield bond, holding either long or short positions in either asset. Any movement of money between these two assets incurs a transaction cost proportional to the size of the transaction. The low–yield bond is liquid in the sense that wealth invested in this bond can be consumed directly without a transaction cost; wealth invested in the high–yield bond can be consumed only by first moving it into the low–yield bond.

The problem of optimal consumption and investment on an infinite planning horizon is solved for a class of utility functions larger than the class of power functions.

1. Introduction.

This paper studies the optimal consumption and investment behavior for a single agent who attempts to maximize his life—time discounted utility of consumption. The utility function for consumption is not necessarily a power function. The assets available to the agent are two bonds, whose interest rates are two different constants r and R with $0 \leq r < R$. Thus, the prices of the low- and high-yield bonds, denoted by $P_0(t)$ and $P_1(t)$, respectively, evolve according to the differential equations

$$dP_0(t) = r\, P_0(t)dt, \tag{1.1}$$

$$dP_1(t) = R\, P_1(t)dt. \tag{1.2}$$

In our model, the agent's <u>position</u> consists of his holdings in both bonds, which he can adjust. He must also choose at each time $t \geq 0$ an instantaneous <u>consumption rate</u> $C_t \geq 0$. This consumption reduces the wealth held in the low-yield bond, which is the "liquid" asset.

The interesting feature here is the market friction. First, there are transaction costs, which we assume to depend on the type of the transaction and to be proportional to the size of the transaction. Second, the agent is allowed to hold a short position in one of the bonds, but his position vector must remain in the <u>solvency region</u> defined in Section 2. The above frictions are necessary for a viable model. Without them, borrowing from the low—yield bond to invest in the high—yield bond would present an arbitrage opportunity.

We will study this problem by solving explicitly the associated Bellman equation, which turns out to be a first-order nonlinear free boundary problem. The free boundary divides the solvency region into two parts. In one part, the agent should jump to the free boundary by trading low—yield bond for high—yield bond. In the other part, the agent should not trade, but should consume. In the latter part, the optimal position process follows a trajectory to the boundary of the solvency region, and then moves along this boundary. If the agent's utility for consumption is a <u>power function</u>, i.e., of the form $U(c) = \frac{1}{p}c^p$ for some $p \in (0,1)$ the free boundary is linear and is explicitly determined, as are the values of positions on and below the free boundary (see equation (13.16)). In particular, the dependence on the model parameters of these values is exhibited in a closed, albeit complicated, form. This permits a comparison of the merits of different risk–free investment opportunities at interest rates above the rate r for the liquid asset, a comparison which takes transaction costs into account.

By replacing our high–yield bond by a stock, whose price P_ε evolves according to the Brownian–motion–driven equation

$$dP_\varepsilon(t) = P_\varepsilon(t)[Rdt + \varepsilon dw_t], \tag{1.3}$$

we get a stochastic transaction cost problem. This problem was formulated by Magill & Constantinides [8], who conjectured for power utility functions that the no–transaction region was a cone in the two–dimensional space of position vectors. This fact was proved by Constantinides [3] in a discrete–time setting, and also in continuous–time under a strong restriction on the class of admissible consumption processes. Recently Davis & Norman [4] provided a rigorous proof for the continuous–time model without Constantinides' restriction. A closely related transaction cost model was studied by Taksar, Klass & Assaf [13]; their objective was to maximize expected long–run average growth of wealth. In all these models, the optimal cumulative purchases of the assets were found to be singularly continuous processes, causing the position process to reflect at the boundaries of the no–transaction region. Thus these problems belong to the class of *singular stochastic control problems* which have received much attention recently (e.g., [1, 2, 5, 11]).

The problem studied in this paper arose from our attempt to permit non–power utility functions in the stochastic transaction cost problem. The no–transaction region will no longer be a cone, and the study of its boundary is difficult. The simpler model with two bonds offers insight into the stochastic model, and can be used as an approximation to the stochastic model with small volatility (ε in (1.3)).

A non–power stochastic model driven by a finite–state Markov chain rather than a Brownian motion has been investigated by Zariphopoulou [14]. She obtained regularity results on the value function, and characterized this function as the solution to the Bellman equation.

This paper is organized as follows. The mathematical formulation of our problem is given in Section 2. In Section 3 we present the Bellman equation of our model and state our main theorem, which characterizes the value function as a concave solution to the Bellman equation satisfying certain boundary conditions. Sections 4–9 are devoted to the explicit construction of a solution V to the Bellman equation satisfying the given boundary conditions In Section 10 we show that V is concave. Section 11 considers properties of the control laws associated with V. In Section 12 we show that V is the value function, and thereby prove the main theorem stated in Section 3. Section 13 develops the special case of power utility functions. This is a completely deterministic paper, and it uses little more than multivariate calculus.

2. Formulation of the Model.

To simplify nomenclature, we henceforth refer to the high–yield bond as the "stock" and to the low–yield bond as the bond. This is consistent with the origin of the model, as explained in Section 1. We shall also refer to position (state) and control "processes", even though these are deterministic.

The agent chooses a consumption rate process C_t, which must be a nonnegative, Borel–measurable function defined on $[0,\infty)$. He also chooses nondecreasing, nonnegative, right–continuous processes M_t and N_t. We adopt the convention

$$M_{0^-} = 0\ , \quad N_{0^-} = 0\ , \tag{2.1}$$

so M_0 (respectively, N_0) is the "size of the jump" in M (respectively, N) at time 0. The number M_t (respectively, N_t) denotes the cumulative purchase (respectively, sale) of "stock") up to time t.

Let the constants $\lambda \in (0,\infty)$ and $\mu \in (0,1)$ represent the proportional transaction costs in the following sense. In order to invest one dollar in the "stock", the agent must divest $(1+\lambda)$ dollars of bond; in order to invest $(1-\mu)$ dollars in the bond, the agent must divest one dollar of "stock". We denote by X_t (respectively, Y_t) the dollars invested in bond (respectively, "stock") at time t, and refer to (X_t, Y_t) as the agent's position at time t. The position evolves according to

$$dX_t = (rX_t - C_t)\,dt - (1+\lambda)\,dM_t + (1-\mu)dN_t, \quad 0 \le t < \infty, \tag{2.2}$$

$$dY_t = R\,Y_t\,dt + dM_t - dN_t, \quad 0 \le t < \infty. \tag{2.3}$$

Following Davis & Norman [4], we define the solvency region to be

$$\mathscr{S} \triangleq \{(x,y) \mid x + (1+\lambda)\,y \ge 0,\ x + (1-\mu)\,y \ge 0\} \tag{2.4}$$

and we partition its boundary into

$$\partial_1\mathscr{S} \triangleq \{(x,y) \mid y > 0, x + (1-\mu)y = 0\},\ \partial_2\mathscr{S} \triangleq \{(x,y) \mid y \le 0, x + (1+\lambda)y = 0\}. \tag{2.5}$$

Roughly speaking, the solvency region is the set of positions for which the net wealth of the agent is nonnegative. We assume that our agent is given an initial endowment $(x,y) \in \mathscr{S}$, i.e.,

$$X_{0^-} = x, \quad Y_{0^-} = y. \tag{2.6}$$

If he chooses the consumption/investment policy (C,M,N), then his position will evolve according to (2.2), (2.3), (2.6). The policy is admissible for (x,y) if $(X_t,Y_t) \in \mathscr{S}$ for every $t \ge 0$, and we denote by $\mathscr{A}(x,y)$ the set of all such policies.

2.1 REIIAU. If $(X_{0}.,Y_{0}-)$ G #2<^> the only admissible policy is to jump immediately to the origin and remain there. This is because $X_t + (1 + A)Y_t$ must remain nonnegative, but

$$d(X_t + (1 + A)Y_t) = r(X_t + (1 + A)Y_t)dt + (R - r)(l + A)Y_t dt$$

$$- C_t dt - (A + \mu)dN_t,$$

and all the terms appearing on the right—hand side are nonpositive if (X_t,Y_t) 6 #2<^. Indeed, $(R - r)(l + A)Y_t$ is strictly negative unless $Y_t = 0$.

Now we introduce the agent's utility function U : [O,OD) -• [O,OD). We assume that U is strictly increasing, strictly concave, twice differentiate, and satisfies

$$U(0) = 0, U'(0) 4 \lim_{c \downarrow 0} U'(c) = a_{»}, U'(») 4 \lim_{c \to \infty} U'(c) = 0. \quad (2.7)$$

The life—time discounted utility of consumption is

$$J((x,y); C.M.N) 4 \int e"^ U(C_t)dt, \quad (2.8)$$

where /? is a positive discount factor. The value function is

$$V^*(x,y) 4 " \sup_{(C,M,N) \in \mathcal{A}(x,y)} J((xj); C.M.N), \quad V (x,y) 6 •. \quad (2.9)$$

In light of Remark 2.1, we must have $V^* = 0$ on $\partial_2\mathcal{S}$.

A remarkable difference between this model and the stochastic model can be observed on the boundary *di<&**. In the stochastic case, if $(X_{0}.,Y_{0}.)$ *e d\<$f* and the agent does not intervene, the stock volatility will cause the position process to immediately exit $\mathscr{S}$. Thus, t agent must sell stock and buy bond so as to bring (X_0,Y_0) to $(0,0)$. This is an absorbing stat with value zero, and so the stochastic model value function is zero on all of $\partial_1\mathscr{S}$.

In the deterministic model, if $(X_{0^-},Y_{0^-}) = (x,y) \in \partial_1\mathscr{S}$, the agent can choose $M \equiv 0$, $N \equiv 0$ and

$$C_t = ye^{Rt}(R-r)(1-\mu), \quad t \geq 0. \tag{2.10}$$

The resulting position process is

$$X_t = -(1-\mu)\, ye^{Rt}, \quad Y_t = ye^{Rt}, \tag{2.11}$$

which is in $\partial_1\mathscr{S}$ for all $t \geq 0$. Therefore,

$$V^*(-(1-\mu)y, y) \geq W(y) \triangleq \int_0^\infty e^{-\beta t}\, U[ye^{Rt}(R-r)(1-\mu)]dt \quad \forall\, y \geq 0 \tag{2.12}$$

We shall eventually discover (see Remark 3.3) that equality holds in (2.12), i.e., (2.11) describes the optimal trajectory for an initial condition on $\partial_1\mathscr{S}$. In any case, (2.12) shows t it is necessary to make Assumption I below. In fact, we make Assumptions I, II and III belc throughout the paper.

Assumption I: $W(y) < \infty \;\; \forall\; y \geq 0$.

In order to analyze our model, we need a condition that the transaction costs are larg enough. The particular form this condition takes is as follows.

Assumption II: W is twice differentiable and

$$(A + /z)U'[y(R-r)(l -//)] - W'(y) > 0 \quad V\ y > 0.$$

Finally, we need a condition on the rate of growth of U' at zero.

ASSUMPTIOI HI: The function g defined by

$$g(y) \triangleq U'[y(R-r)(1-\mu)]y^{\frac{r-p}{R}} \quad V\ y > 0 \tag{2.13}$$

is strictly decreasing and satisfies

$$\lim_{ylo} g(y) = \infty \tag{2.14}$$

2.2 REMIM. If U(c) = $\frac{c^p}{j}$ for some p e (0,1) then Assumption I, II and III become β - Rp > 0, 0 - Rp > $\frac{(R - r)(1 - \mu)}{\lambda + \mu}$ and 0 - Rp > -(R - r), respectively. Thus, Assumption II implies the other assumptions. We see in Remark 13.1 that Assumption II is necessary for power utility functions.

2.3REMAM. The concavity of U and (2.7) imply that y U'(y) $\leq$ U(y) V y > 0. Therefore, Assumption I and the Dominated Convergence Theorem imply that W is finite and can be computed by differentiation under the integral, i.e.,

$$W'(y) = (R-r)(1-\mu)\int_0^\infty e^{-(\beta-11)t}\, U'[ye^{at}(R-r)(l-/z)]dt$$

$$= \frac{1}{R}(R-r)(1-/i)y^{\frac{ft-Jt}{}} \int_y a^{-\frac{0}{}} U'[\sigma(R-r)(1-\mu)]da. \tag{2.15}$$

This implies that

$$\lim_{y\to\infty} y^{\frac{R-\beta}{R}} U'[y(R-r)(1-\mu)] = 0, \tag{2.16}$$

since

$$0 = \lim_{y\to\infty} \int_y^{\infty} \sigma^{-\frac{\beta}{R}} U'[\sigma(R-r)(1-\mu)]d\sigma$$

$$\geq \lim_{y\to\infty} \int_y^{2y} \sigma^{-\frac{\beta}{R}} U'[\sigma(R-r)(1-\mu)]d\sigma$$

$$\geq \lim_{y\to\infty} y\,(2y)^{-\frac{\beta}{R}} U'[2y(R-r)(1-\mu)].$$

Furthermore, integration by parts of the dt integral in (2.15) shows that

$$W'(y) = \frac{1}{Ry}[\beta W(y) - U(y(R-r)(1-\mu))] \quad \forall\, y > 0. \tag{2.17}$$

3. The Bellman Variational Inequality.

Because U' is a strictly decreasing function mapping $[0,\infty]$ onto $[0,\infty]$ (see (2.7)), U' has a strictly decreasing inverse I mapping $[0,\infty]$ onto $[0,\infty]$. We define the convex Legendre transform

$$\tilde{U}(v) \triangleq \inf_{c\geq 0} \{cv - U(c)\} = vI(v) - U(I(v)) \quad \forall\, v > 0. \tag{3.1}$$

In terms of $\tilde{U}$, we can write the Bellman variational inequality for the deterministic control problem formulated in Section 2 as

$$\min\{\beta V - rxV_x - RyV_y + \tilde{U}(V_x), (1+\lambda)V_x - V_y,$$
$$-(1-\mu)V_x + V_y\} = 0 \quad \text{on } \mathscr{S}\backslash\partial_2\mathscr{S}. \tag{3.2}$$

One consequence of (3.2) is

$$\beta V - rxV_x - RyV_y + c\, V_x - U(c) \geq 0 \quad \forall\ c \geq 0. \tag{3.3}$$

3.1 Lemma. Assume that V is a nonnegative, concave, continuously differentiable solution to the variational inequality (3.2). Then V dominates the value function defined by (2.9), i.e.,

$$V(x,y) \geq V^*(x,y) \quad \forall\ (x,y) \in \mathscr{S}. \tag{3.4}$$

Proof: Fix $(x,y) \in \mathscr{S}$. For any $(C,M,N) \in \mathscr{A}(x,y)$, let (X_t,Y_t) be the corresponding position process and define

$$\sigma \triangleq \inf\{t \geq 0 \mid (X_t,Y_t) \in \partial_2\mathscr{S}\}.$$

Applying the chain rule rule for finite–variation functions (Rogers & Williams [9], p. 29), we get

$$V(x,y) - e^{-\beta(t\wedge\sigma)}V(X_{t\wedge\sigma}, Y_{t\wedge\sigma})$$

$$= \int_0^{t\wedge\sigma} e^{-\beta s}[\beta V(X_s,Y_s) - rX_sV_x(X_s,Y_s) - RY_sV_y(X_s,Y_s) + C_sV_x(X_s,Y_s)]ds$$

$$+\int_0^{t\wedge\sigma} e^{-\beta s}[(1+\lambda)V_x(X_{s^-},Y_{s^-}) - V_y(X_{s^-},Y_{s^-})]dM_s$$

$$+\int_0^{t\wedge\sigma} e^{-\beta s}[-(1+\mu)V_x(X_{s^-},Y_{s^-}) + V_y(X_{s^-},Y_{s^-})]dN_s \tag{3.5}$$

$$+\sum_{0\le s\le t\wedge\sigma} e^{-\beta s}[-V(X_s,Y_s)+V(X_{s^-},Y_{s^-}) + V_x(X_{s^-},Y_{s^-})\Delta X_s + V_y(X_{s^-},Y_{s^-})\Delta Y_s$$

From (3.2) we know that the second and third integrals are nonnegative. The sum is nonnegative because of the concavity of V. Because of (3.3), the first integral dominates $\int_0^{t\wedge\sigma} e^{-\beta s}\, U(C_s)ds$. Letting t go to infinity, recalling that $V \ge 0$, and observing from Rema 2.1 that $U(C_s) = 0 \;\forall\; s \ge \sigma$, we obtain $V(x,y) \ge \int_0^{\infty} e^{-\beta s}\, U(C_s)ds$. Maximization of the right–hand side over $(C,M,N) \in \mathscr{A}(x,y)$ results in (3.4). □

We next seek to understand how equality can occur in (3.4). For a model in which the is a bond and a risky stock, one would expect there to be two free boundaries, which would split the region $\mathscr{S}$ into three parts. On one part, denoted by NT, no transaction should occur, and the nonlinear differential equation

$$\beta V - rx\, V_x - RyV_y + \tilde{U}(V_x) = 0 \quad \text{on NT} \tag{3.6}$$

would be satisfied. On another region, denoted by SB (Sell Bond), the agent holds too much bond and should transfer some wealth into stock. Thus, his position should jump to the boundary of NT, moving in the direction $(-(1+\lambda),1)$. We would have

$$(1+\lambda)V_x - V_y = 0 \quad \forall\, (x,y) \in SB. \tag{3.7}$$

On the last part, denoted by SS (Sell Stock), the agent should transfer wealth from stock to bond, causing his position jump in the direction $((1-p),-1)$ to the boundary of NT. On this region, we would have

$$-(1-\mu)V_x + V_y = 0 \quad V(x,y) e SB. \tag{3.8}$$

However, in our model, the "stock" is riskless. It would make no sense to sell such a "stock", since it has a higher interest rate than the bond. This leads us to conjecture for our model that there is no region SS and only one free boundary, defined by

$$x = h(y) , \quad y \geq 0 , \tag{3.9}$$

for some continuously differentiable function h. The no—transaction region is defined as

$$NT = \{(x,y) \; e \; (\wedge \mid y \geq 0, x \leq h(y)\} , \tag{3.10}$$

and the sell—bond region is defined by

$$SB = \{(x,y) \; € <\wedge \mid y \geq 0, x > h(y)\} \; U \; \{(x,y) \; e \; \wedge \mid y < 0\}. \tag{3.11}$$

We will establish the above conjecture, and further show that

$$h(0) = 0, \;\; h(y) > -(1-M)y, \;\; h'(y) > -(1+A) \;\; V \; y > 0. \tag{3.12}$$

These conditions guarantee that except for the initial point $(h(0),0)$, the graph of h lies in the interior of *df*. In the Lemma 9.1 we use these conditions to show that for each (x,y) e SB, there is a unique point $(h(\bar{y}),\bar{y})$ on the graph of h which can be reached by moving from (x,y) in the $(-(1+A),1)$ direction. This point is determined by the equation

$$h(\bar{y}) + (1+\lambda)\bar{y} = x + (1+\lambda)y. \tag{3.13}$$

The main theorem of this paper, whose proof proceeds throughout Sections 4 - 12, is the following.

3.2 THEOREM. Under Assumptions I, II and III, the value function V^* is a nonnegative, concave, continuously differentiable solution to the variational inequality (3.2) satisfying the boundary conditions

$$V^*Hl\text{-}/*)y,y) = W(y),\ V^*_xHl\text{-}_M)y,y) = U'(y(R-r)(l\text{-}/x)),\ y > 0, \tag{3.14}$$

$$V^*H1+A)y,y) = 0\ ,\quad y \le 0. \tag{3.15}$$

Furthermore, there is a continuously differentiable function h : [0,o>) -»[0,a>) such that (3.12), (3.6) and (3.7) hold, where NT and SB are defined by (3.10) and (3.11). If *(X*$_o$., Y*$_o$.)* = (x,y) e NT, the optimal control processes in feedback form are

$$M^*_t \equiv 0,\ N^*_t \equiv 0,\ C^*_t = I(V^*_x(X^*_{t)}Y^*_t)),\ t \ge 0. \tag{3.16}$$

If $(X^*_0.,Y^*_0.) = (x,y)$ € SB, the optimal control processes in feedback form are

$$M^*_t = \bar{y} - y,\ N^*_t = 0,\ C^*_t = I(V^*_x(X^*_{t)}Y^*t)),\ t \ge 0, \tag{3.17}$$

where $\bar{y}$ is determined by (3.13). (Recall that $M^*_o. = 0$, so (3.17) mandates a jump in position at time zero. Furthermore, V^*_x = <» on *dttf* and (3.17) is to be interpreted as $C^*_t = 0$ if *(X*$_t$,Y*$_t$)* € $\partial_2\mathcal{S}$.)

3.3 REMAIK. The boundary condition (3.14) on V^*_x is chosen so that (3.16) is consistent with

(2.10), (2.11).

3.4 Remark. Because we construct V^*, the uniqueness of the solution to (3.2) subject to the boundary conditions (3.14), (3.15) is not a major concern. It is to be expected that under condition (3.15), either of the equalities in (3.14), and a growth condition, uniqueness can be proved by the methods employed in Soner [10] and Soner and Shreve [12].

4. Linearization of the Bellman Equation

We first tackle the nonlinear equation (3.6) in NT by a linearization technique introduced by Karatzas, Lehoczky, Sethi and Shreve [7]. We proceed formally in this section and make the arguments precise beginning in Section 5.

Suppose V is a strictly concave, smooth solution to (3.6) satisfying the boundary condition (3.14). Then for each $y > 0$, $V_x(\cdot,y)$ is a strictly decreasing function and so has a strictly decreasing inverse $\mathscr{X}(\cdot,y)$, defined on some subinterval of $\mathbb{R}$. In other words,

$$V_x(\mathscr{X}(\delta,y),y) = \delta, \quad \mathscr{X}(V_x(x,y),y) = x \tag{4.1}$$

for all $y > 0$ and all δ and x in intervals to be specified later. Differentiation of (4.1) with respect to δ and y leads to the formulas

$$V_{xx}(\mathscr{X}(\delta,y),y) = \frac{1}{\mathscr{X}_\delta(\delta,y)}, \quad V_{xy}(\mathscr{X}(\delta,y),y) = -\frac{\mathscr{X}_y(\delta,y)}{\mathscr{X}_\delta(\delta,y)}. \tag{4.2}$$

Differentiating (3.6), we get

$$(\beta - r)V_x - rxV_{xx} - Ry\,V_{xy} + I(V_x)V_{xx} = 0. \tag{4.3}$$

Substituting (4.2) into (4.3), we obtain the linear equation

$$(\beta - r)\delta \mathscr{X}_\delta + Ry\mathscr{X}_y - r\mathscr{X} + I(\delta) = 0, \tag{4.4}$$

whose general solution is

$$\mathscr{X}(\delta,y) = y^{\frac{r}{R}} \theta(\delta y^{\frac{r-\beta}{R}}) + \psi(\delta), \tag{4.5}$$

where θ is a C^1 function and ψ is a particular solution to (4.4). We specify ψ by

$$\psi(\delta) \triangleq \begin{cases} \frac{1}{r} I(\delta) & \text{if } \beta = r, \\ \frac{1}{r} I(\delta) + \frac{1}{r}\int_{I(\delta)}^{A} \left[\frac{U'(c)}{\delta}\right]^{\frac{r}{r-\beta}} dc, & \text{if } \beta \neq r, \end{cases} \tag{4.6}$$

where

$$A = \begin{cases} 0 & \text{if } \beta > r, \\ \infty & \text{if } \beta < r. \end{cases}$$

4.1 Remark. The integral in (4.6) is finite. If $\beta > r$, then the integrand is bounded at zero. Suppose $\beta < r$. Equation (2.16) implies $U'(c) \leq c^{\frac{\beta - R}{R}}$ for sufficiently large c, so

$$(U'(c))^{\frac{r}{r-\beta}} = U'(c)(U'(c))^{\frac{\beta}{r-\beta}} \leq U'(c)\, c^{-\frac{\beta}{R}(\frac{R-\beta}{r-\beta})}$$

$$\leq U'(c)\, c^{-\frac{\beta}{R}}$$

for sufficiently large c. The last expression is integrable at infinity because of the finiteness of W' (see (2.15)). □

We use the boundary condition (3.14) to determine θ in (4.5). Define

$$\delta^L(y) \triangleq U'(y(R-r)(1-\mu)) , \quad \forall\, y > 0 , \tag{4.7}$$

$$g(y) \triangleq \delta^L(y)\, y^{\frac{r-\beta}{R}} , \quad \forall\, y > 0. \tag{4.8}$$

Then $V_x(-(1-\mu)y,y) = \delta^L(y)$, or equivalently,

$$\mathscr{X}(\delta^L(y),y) = -(1-\mu)y, \quad \forall\, y > 0. \tag{4.9}$$

Comparing this with (4.5), we obtain

$$\theta(g(y)) = \begin{cases} -\frac{R}{r}(1-\mu)y^{\frac{R-r}{R}} & \text{if } \beta = r, \\ -\frac{R}{r}(1-\mu)\, y^{\frac{R-r}{R}} - \frac{1}{r}\, y^{-\frac{r}{R}} \int_{y(R-r)(1-\mu)}^{A} \left[\frac{U'(c)}{\delta^L(y)}\right]^{\frac{r}{r-\beta}} dc & \text{if } \beta \neq r. \end{cases} \tag{4.10}$$

By Assumption III and (2.16), g is a strictly decreasing function mapping $(0,\infty)$ onto $(0,\infty)$, so it has a strictly decreasing inverse $f : (0,\infty) \xrightarrow{\text{onto}} (0,\infty)$, i.e.,

$$f(g(y)) = y , \quad g(f(\alpha)) = \alpha , \quad \forall\, y > 0,\ \alpha > 0. \tag{4.11}$$

Note that $\delta^L(f(\alpha)) = \alpha(f(\alpha))^{\frac{\beta-r}{R}}$. We may rewrite (4.10) as

$$\theta(\alpha) = \begin{cases} -\frac{R}{r}(1-\mu)(f(\alpha))^{\frac{R-r}{R}} & \text{if} \quad \beta = r, \\ -\frac{R}{r}(1-\mu)(f(\alpha))^{\frac{R-r}{R}} - \frac{1}{r}\int_{f(\alpha)(R-r)(1-\mu)}^{A} \left[\frac{U'(c)}{\alpha}\right]^{\frac{r}{r-\beta}} dc & \text{if } \beta \neq r. \end{cases} \tag{4.12}$$

Direct computation shows that

$$(\beta-r)\alpha\,\theta'(\alpha) - r\theta(\alpha) = R(1-\mu)(f(\alpha))^{\frac{R-r}{R}} \quad \forall\ \alpha > 0. \tag{4.13}$$

5. Precise Definition and Properties of $\mathcal{X}$.

The preceding section, particularly (4.5), (4.6) and (4.12), suggest the precise definition of a function $\mathcal{X}$ as

$$\mathcal{X}(\delta,y) \triangleq \begin{cases} \frac{1}{r}I(\delta) - \frac{R}{r}(1-\mu)y^{\frac{r}{R}}\left[\frac{I(\delta)}{(R-r)(1-\mu)}\right]^{\frac{R-r}{R}}, & \text{if } \beta = r, \\ \frac{1}{r}I(\delta) - \frac{R}{r}(1-\mu)y^{\frac{r}{R}}\left[f(\delta y^{\frac{r-\beta}{R}})\right]^{\frac{R-r}{R}} & \end{cases} \tag{5.1}$$

$$+\frac{1}{r}\int_{I(\delta)}^{(R-r)(1-\mu)f(\delta y^{(r-\beta)/R})} \left(\frac{U'(c)}{\delta}\right)^{\frac{r}{r-\beta}} dc \quad \text{if } \beta \neq r,$$

defined for all $y > 0$ and $0 < \delta \le \delta^L(y)$. We use δ for the first argument to remind us that the derivative V_x^* should appear in this position. Note that the definition of $\mathcal{X}$ is independent of A.

Our intention is to use $\mathcal{X}$ to construct a function V satisfying (3.6). With ψ and θ defined by (4.6) and (4.12), the equations (4.4), (4.5) and (4.9) hold. Differentiation of (4.5), coupled with (4.13), yields

$$\mathscr{X}_y(\delta,y) = -(1-\mu)y^{\frac{r-R}{R}}[f(\delta y^{\frac{i-0}{R}})]^{\frac{R-r}{R}}. \tag{5.2}$$

If P = r, then differentiation of (5.1) yields

$$\mathscr{X}_\delta(\delta,y) = \frac{1}{r}I'(\delta)\{1-[\frac{I(\delta)}{y(R-r)(1-\mu)}]^{-\frac{r}{R}}\}; \tag{5.3}$$

if 0 + r, then (4.4), (5.1) and (5.2) imply

$$\mathscr{X}_\delta(\delta,y) = \frac{1}{(\beta-r)\delta}[r\,\mathscr{X}(\delta,y) - Ry\mathscr{X}_y(\delta,y) - I(\delta)]$$

(5.4)

$$= \frac{1}{(\beta-r)\delta}\int_{1(6)}^{(R-r)(1-\mu)f(\delta y^{(r-\beta)/R})}(\frac{U'(c)}{\delta})^{\frac{r}{r-\beta}}dc.$$

5.1 LEMMA. For each y > 0, the function ^(-,y) : (0,^(y)] -* [-(l-/x)y,oD) satisfies

$$\mathscr{X}_\delta(\delta^L(y),y) = 0\,, \tag{5.5}$$

$$\mathscr{X}_\delta(\delta,y) < 0 \quad \forall \quad \delta \in (0,\delta^L(y)). \tag{5.6}$$

PROOF: We first consider the case /? = r. Since I is decreasing, we have *1(5)_> l(6^L(y))* = y(R-r)(l-/i) for tf 6 (0,^(7)] and the inequality is strict if *S* < *6^L(y)*. The result follows from (5.3).

We next consider the case /3 $ r. Let *6 e* (0,£^L(y)) be given, and note that y(R-r)(l-M) = I(£^{L}(y)) < I(^)- $^{Fr<>m\ thi}$s follow the equivalences

$$\beta < r \Leftrightarrow \delta\, y^{\frac{r-\beta}{R}} < \delta\left(\frac{I(\delta)}{(R-r)(1-\mu)}\right)^{\frac{r-\beta}{R}} = g\left(\frac{I(\delta)}{(R-r)(1-\mu)}\right)$$

$$\Leftrightarrow f(\delta\, y^{\frac{r-\beta}{R}}) > \frac{\text{Iffl}}{(R-r)(1-\mu)} \qquad \textbf{(5.7)}$$

$$\Leftrightarrow (R-r)(1-\mu)\, f(\delta\, y^{\frac{r-\beta}{R}}) > I(\delta).$$

Inequality (5.6) follows from (5.4). If *6* = ^(y), except for /? < r the inequalities in (5.7) become equalities, and (5.5) is obtained. **o**

6. Inversion of *£"(•,y).

For each y > 0, the strictly decreasing function ^(..y) has an inverse A(- ,y). We will eventually identify A(x,y) with V$\overset{*}{x}$(x,y), at least on part of the domain of the former. In light of (4.9), -(1—*p,)y* is one endpoint of the domain of A(- ,y), and is itself in the domain. The other endpoint is ^"(0,y) ≜ lim <#(£,y), which is not a member of the domain. Thus, the
6[0
domain of A is

3 4 {(x,y) | y > 0, -(H*)y ≤ x < jr(0,y)], (6.1)

which is a subset of G/, and

Δ(𝒳(δ,y),y)= *6,* jr(A(x,y),y)= x V y > 0, 0 < S < ^(y), -(1-M)y < x < *S(0j).* (6.2)

In particular,

$$\Delta(-(1-\mu)y,y) = \delta^L(y) \quad \forall\ y > 0. \tag{6.3}$$

7. Construction of $\hat{V}$ on $\bar{\mathscr{D}}$.

In this section we construct a function $\hat{V}$ on $\bar{\mathscr{D}} \triangleq \mathscr{D} \cup \{0,0)\}$, a set which will turn out to contain NT.

Restricted to NT, $\hat{V}$ will agree with the value function V^*. We define $\hat{V}(0,0) = 0$ and

$$\hat{V}(x,y) \triangleq W(y) + \int_{-(1-\mu)y}^{x} \Delta(\xi,y)d\xi \ \forall\ (x,y) \in \mathscr{D}. \tag{7.1}$$

Because Δ is not bounded near $(0,0)$, it is not immediately clear that $\hat{V}$ is continuous at $(0,0)$; however, we establish this continuity in the summary at the end of the next section.

7.1 Lemma. We have for all $(x,y) \in \mathscr{D}$,

$$\hat{V}_x(x,y) = \Delta(x,y) \tag{7.2}$$

$$\hat{V}_y(x,y) = -(1-\mu)y^{\frac{\beta-R}{R}} \int_{f(\Delta(x,y)y^{(r-\beta)/R})}^{\infty} \sigma^{\frac{R-r}{R}} g'(\sigma)d\sigma. \tag{7.3}$$

Proof: Differentiation of (7.1) yields (7.2). To obtain (7.3), we make the substitution $\xi = \mathscr{X}(\delta,y)$ in (7.1) and integrate by parts to obtain

$$\hat{V}(x,y) = W(y) + \int_{\delta^L(y)}^{\Delta(x,y)} \delta \mathscr{X}_\delta(\delta,y)d\delta$$

$$= W(y) + x\Delta(x,y) + (1-\mu)y\delta^L(y) - \int_{\delta^L(y)}^{\Delta(x,y)} \mathscr{X}(\delta,y)d\delta.$$

Therefore,

$$\hat{V}_y(x,y) = W'(y) + (1-\mu)\delta^L(y) - \int_{\delta^L(y)}^{\Delta(x,y)} \mathscr{X}_y(\delta,y)d\delta. \tag{7.4}$$

Denote the right–hand side of (7.3) by $H(x,y)$. After integration by parts and an application of (2.16), we see that

$$\begin{aligned} H(x,y) = {} & (1-\mu)y^{\frac{r-R}{R}} \Delta(x,y) \left[f(\Delta(x,y)y^{\frac{r-\beta}{R}})\right]^{\frac{R-r}{R}} \\ & + \frac{1}{R}(R-r)(1-\mu)\, y^{\frac{\beta-R}{R}} \int_{f(\Delta(x,y)y^{\frac{r-\beta}{R}})}^{\infty} \sigma^{-\frac{\beta}{R}} U'(\sigma(R-r)(1-\mu))d\sigma. \end{aligned} \tag{7.5}$$

From (6.3), (4.11), (2.15) and (7.4), we have

$$H(-(1-\mu)y,y) = (1-\mu)\ \delta_0(y) + W'(y) = \hat{V}_y(-(1-\mu)y,y). \tag{7.6}$$

Differentiation of (7.5), taking (4.11), the definition of g, and (5.2) into account, results in

$$H_x(x,y) = (1-\mu)\, y^{\frac{r-R}{R}}\, \Delta_x(x,y)\, [f(\Delta(x,y)\, y^{\frac{r-\beta}{R}})]^{\frac{R-r}{R}}$$

$$= -\mathscr{X}_y(\Delta(x,y),y)\, \Delta_x(x,y). \tag{7.7}$$

From (6.2) and (7.2) we have

$$-\mathscr{X}_y(\Delta(x,y),y)\Delta_x(x,y) = \Delta_y(x,y) = \hat{V}_{xy}(x,y). \tag{7.8}$$

Equation (7.3) follows from (7.6) – (7.8). □

7.2 Lemma. The function

$$G(\alpha) \triangleq (\beta-R) \int_{f(\alpha)}^{\infty} \sigma^{\frac{R-r}{R}} g'(\sigma)d\sigma + (\beta-r)\, \alpha\, (f(\alpha))^{\frac{R-r}{R}}$$

is positive for every $\alpha > 0$.

Proof: Integration by parts using (2.16) reveals that

$$G(\alpha) = (R-r)[-\frac{(\beta-R)}{R} \int_{f(\alpha)}^{\infty} \sigma^{-\frac{r}{R}} g(\sigma)d\sigma + \alpha(f(\alpha))^{\frac{R-r}{R}}],$$

which is positive if $\beta \leq R$. If $\beta > R$, then

$$\frac{\beta-R}{R}\int_{f(\alpha)}^{\infty}\sigma^{-\frac{r}{R}}g(\sigma)d\sigma = \frac{\beta-R}{R}\int_{f(\alpha)}^{\infty}\sigma^{-\frac{\beta}{R}}U'(\sigma(R-r)(1-\mu))d\sigma$$

$$< \frac{\beta-R}{R}U'((R-r)(1-\mu)f(\alpha))\cdot\int_{f(\alpha)}^{\infty}\sigma^{-\frac{\beta}{R}}d\sigma$$

$$= g(f(\alpha))\,(f(\alpha))^{\frac{R-r}{R}} = \alpha(f(\alpha))^{\frac{R-r}{r}},$$

and again $G(\alpha) > 0$. □

We would like to show that $\hat{V}$ is concave, but it is not clear that its domain $\mathscr{D} \cup \{(0,0)\}$ is convex. We content ourselves, therefore, with the following local convexity result.

7.3 Proposition. The Hessian matrix $\nabla^2\hat{V}$ is negative definite on $\mathscr{D}$.

Proof: Recall that $\mathscr{D}$ is the range of the mapping $\mathscr{X}$. Differentiation of (6.2) implies

$$\Delta_y(\mathscr{X}(\delta,y),y) = -\Delta_x(\mathscr{X}(\delta,y),y)\mathscr{X}_y(\delta,y) = -\frac{\mathscr{X}_y(\delta,y)}{\mathscr{X}_\delta(\delta,y)}. \tag{7.9}$$

Differentiation of (7.2) yields

$$\hat{V}_{xx}(\mathscr{X}(\delta,y),y) = \Delta_x(\mathscr{X}(\delta,y),y) = \frac{1}{\mathscr{X}_\delta(\delta,y)}, \tag{7.10}$$

$$\hat{V}_{xy}(\mathscr{X}(\delta,y),y) = \Delta_y(\mathscr{X}(\delta,y),y) = -\frac{\mathscr{X}_y(\delta,y)}{\mathscr{X}_\delta(\delta,y)}. \tag{7.11}$$

Differentiation of (7.3) yields

$$\hat{V}_{yy}(x,y) = -\frac{(\beta-R)}{Ry}(1-\mu)y^{\frac{\beta-R}{R}}\int_{f(\Delta(x,y)y^{(r-\beta)/R})}\sigma^{\frac{R-r}{R}}g'(\sigma)d\sigma$$

$$+ (1-\mu)y^{\frac{r-R}{R}}\left(i(\delta y^{*})\right)^{\frac{R-r}{R}}[A_y(x,y) + \wedge A(x,y)].$$

Using (5.2) and (7.9), we may rewrite this as

$$\hat{V}_{yy}(\mathscr{X}(\delta,y),y) = -\frac{(1-\mu)}{Ry}y^{\frac{\beta-R}{R}}G(\delta y^{\frac{r-\beta}{R}}) + \frac{\mathscr{X}_y^2(\delta,y)}{\mathscr{X}_\delta(\delta,y)}. \tag{7.12}$$

Now $\hat{V}_{xx} < 0$ on $\mathscr{D}$ because of (7.10) and Lemma 5.1. Moreover

$$\det \nabla^2\hat{V}(\mathscr{X}(\delta,y),y) = -\frac{(1-\mu)}{Ry}y^{\frac{\beta-R}{R}}G(\delta,y^{\frac{r-\beta}{\beta}})\hat{V}_{xx}(\mathscr{X}(\delta,y),y) > 0$$

because of Lemma 7.2. D

7.4 THEOREM. The function $\hat{V}$ solves the nonlinear equation (3.6) and satisfies the boundary conditions (3.14).

PROOF: The boundary condition follows immediately from the definition of $\hat{V}$ and from (7.2) and (6.3). On ^i^\{0,0)}, we have from (3.14), (7.4), (3.1) and (2.17) that (3.6) holds, to wit,

$$\beta\hat{V}(x,y) - rx\hat{V}_x(x,y) - Ry\,\hat{V}_y(x,y) + \tilde{U}(\hat{V}_x(x,y))$$

$$= \beta W(y) + r(1-\mu)y\delta^L(y) - Ry[W'(y) + (1-\mu)\delta^L(y)]$$

$$+ \delta^L(y) I(\delta^L(y)) - U(I(\delta^L(y)))$$

= 0W(y) - ^{R}y $^{w/}$(y) - U(y(R-r)(l-/x)) = 0.

Furthermore, $\hat{V}$ satisfies (4.3) because 3> satisfies (4.4). If we now integrate (4.3) with respect to x, starting at x = -(l-/i)y, we obtain (3.6). D

8. The Free Boundary.

On the free boundary, we should have (1+A)V_X - V_y = 0. The formulas in Lemma 7.1 suggest the introduction of the function

$$F(\rho,y) \triangleq (1+\lambda)\rho\delta^L(y) + (1-\mu)y^{\frac{\beta-R}{irr}} \int_{f(\rho\delta^L(y)y^{(r-\beta)/R})} \sigma^{\frac{R-r}{R}} g'(\sigma)d\sigma, \qquad (8.1)$$

defined for y > 0, p > 0. For each y > 0, we seek p(y) 6 (0,1) such that

$$F(/>(y),y) = 0. \qquad (8.2)$$

The free boundary h will then be characterized by the equation £^{R}(y) $\triangleq$ />(y)£^{L}(y) = A(x,y), or equivalently,

$$h(y) 4 \wedge(\wedge(y),y) \; V \; y > 0. \qquad (8.3)$$

8.1 LEMMA. For each y > 0, F(« ,y) is a strictly convex function satisfying

$$\lim_{\rho i O} F(p,y) = 0, \quad \lim_{\rho i 0} F_\rho(\rho,y) = -\infty, \qquad (8.4)$$

$$F(1,y) = (A+/i)5^L(y)-W'(y)>0. \qquad (8.5)$$

In particular, there exists a C^1 function $\rho\colon (0,\infty) \to (0,1)$ such that (8.2) holds for every $y > 0$.

PROOF: We have

$$F_\rho(\rho,y) = (1+\lambda)\delta^L(y) - (1-\mu)y^{\frac{\beta-R}{R}} [f(\rho\delta^L(y)y^{\frac{r-\beta}{R}})]^{\frac{R-r}{R}} \delta^L(y)y^{\frac{r-\beta}{R}},$$

which satisfies (8.4) and is strictly increasing in ρ. By (7.3), (7.4) and Assumption II,

$$F(1,y) = (1+\lambda)\delta^L(y) - \hat{V}_y(-(1-\mu)y,y)$$

$$= (\lambda+\mu)\delta^L(y) - W'(y) > 0.$$

These facts imply the existence of a function $\rho(\cdot)$ satisfying (8.2); smoothness follows from the Implicit Function Theorem and the fact that

$$F_\rho(\rho(y),y) > 0 \quad \forall\ y > 0. \tag{8.6}$$

□

8.2 COROLLARY. For each $y > 0$, we have

$$F(\rho,y) > 0 \quad \forall\, \rho > \rho(y), \tag{8.7}$$

$$(\lambda+\mu)\rho\, \delta^L(y) - F(\rho,y) > 0 \quad \forall\, \rho \in (0,1]. \tag{8.8}$$

PROOF: Because $W'(y) > 0$ (Remark 2.3), the concave function of ρ in (8.8) is nonnegative at $\rho = 0$ and positive at $\rho = 1$, hence positive on $(0,1]$. □

The function $\delta^R(y) \triangleq \rho(y)\delta^L(y)$ is characterized by the equation

$$(1+\lambda)\delta^{R}(y)+(1-\mu)y^{\frac{\beta-R}{R}}\int_{f(\delta^{R}(y)y^{(r-\beta)/R})}^{\infty}\sigma^{\frac{R-r}{R}}g'(\sigma)d\sigma = 0 \;\forall\; y > 0. \tag{8.9}$$

We shall see that $\delta^{R}(y)$ is the derivative $V_x^*(h(y),y)$ along the free boundary.

8.3 Lemma. The function $\delta^{R} : (0,\infty) \to \mathbb{R}$ is positive, strictly decreasing, and satisfies

$$\lim_{y\downarrow 0} \delta^{R}(y) = \infty, \quad \lim_{y\downarrow 0} f(\delta^{R}(y)\, y^{\frac{r-\beta}{R}}) = 0. \tag{8.10}$$

Furthermore, for every $y > 0$,

$$1 + \lambda + \mathscr{X}_y(\delta^{R}(y),y) > 0, \tag{8.11}$$

$$\frac{\frac{d}{dy}\delta^{R}(y)}{\delta^{R}(y)} = \frac{(\beta-R)(1+\lambda) + (\beta-r)\mathscr{X}_y(\delta^{R}(y),y)}{Ry[1 + \lambda + \mathscr{X}_y(\delta^{R}(y),y)]}. \tag{8.12}$$

Proof: From (8.6) and (5.2) we have

$$0 < F_\rho(\rho(y),y) = (1+\lambda)\delta^{L}(y) - (1-\mu)y^{\frac{r-R}{R}}[f(\delta^{R}(y)y^{\frac{r-\beta}{R}})]^{\frac{R-r}{R}}\delta^{L}(y)$$

$$= \delta^{L}(y)[1 + \lambda + \mathscr{X}_y(\delta^{R}(y),y)],$$

which proves (8.11). The above inequality also implies

$$f(\delta^{R}(y)y^{\frac{r-\beta}{R}}) \le \left(\frac{1+\lambda}{1-\mu}\right)^{\frac{R}{R-r}}y \quad \forall\; y > 0. \tag{8.13}$$

Applying the function g to both sides of (8.13), we obtain

$$\delta^R(y) \geq \left(\frac{1+\lambda}{1-\mu}\right)^{\frac{r-\beta}{R-r}} U'\left[(R-r)(1-\mu)\left(\frac{1+\lambda}{1-\mu}\right)^{\frac{R}{R-r}} y\right] \quad \forall\, y > 0. \tag{8.14}$$

The limits (8.10) follows from (8.13) and (8.14).

Differentiation of (8.9), followed by substitution based on (8.9), results in the formula

$$0 = (1+\lambda)\frac{d}{dy}\,\delta^R(y) - \left(\frac{\beta-R}{Ry}\right)(1+\lambda)\delta^R(y)$$

$$-\,(1-\mu)\, y^{\frac{r-R}{R}}\left[f\left(\delta^R(y)y^{\frac{r-\beta}{R}}\right)\right]^{\frac{R-r}{R}}\left[\frac{d}{dy}\,\delta^R(y) + \left(\frac{r-\beta}{Ry}\right)\delta^R(y)\right]$$

$$= \frac{d}{dy}\,\delta^R(y)\left[1 + \lambda + \mathscr{X}_y(\delta^R(y),y)\right] - \frac{\delta^R(y)}{Ry}\left[(\beta-R)(1+\lambda) + (\beta-r)\mathscr{X}_y(\delta^R(y),y)\right],$$

which gives us (8.12).

It remains to show that δ^R is strictly decreasing, which we accomplish by proving

$$\delta^R(y)\left[(\beta-R)(1+\lambda) + (\beta-r)\mathscr{X}_y(\delta^R(y),y)\right] < 0 \quad \forall\, y > 0, \tag{8.15}$$

and then appealing to (8.11), (8.12). But with G as defined in Lemma 7.2, we may use (8.9) to write the expression in (8.15) as

$$-(\beta-R)(1-\mu)y^{\frac{\beta-R}{R}} \int_{f(\delta^R(y)y^{(r-\beta)/R})}^{\infty} \sigma^{\frac{R-r}{R}} g'(\sigma)d\sigma$$

$$-(\beta-r)(1-\mu)y^{\frac{\beta-R}{R}} \delta^R(y)y^{\frac{r-\beta}{R}} [f(\delta^R(y)y^{\frac{r-\beta}{R}})]^{R-r}$$

$$= -(1-\mu)\, y^{\frac{\beta-R}{R}}\, G(\delta^R(y)y^{\frac{r-\beta}{R}}) < 0.$$

□

The free boundary is defined by (8.3). We extend h by setting

$$h(0) = 0, \quad h(y) = \mathscr{X}(\delta^R(y),y) \quad \forall\ y > 0. \tag{8.16}$$

8.4 Lemma. The free boundary h satisfies

$$\lim_{y \downarrow 0} h(y) = 0\,, \tag{8.17}$$

i.e., h is continuous at $y = 0$. Furthermore,

$$h(y) > -(1-\mu)y, \quad h'(y) > -(1+\lambda) \quad \forall\ y > 0. \tag{8.18}$$

Proof: For (8.17), we consider only the more difficult case of $\beta \neq r$. According to (5.1),

$$h(y) = \frac{1}{r} I(\delta^R(y)) - \frac{R}{r}(1-\mu)y^{\frac{r}{R}} [f(\delta^R(y)y^{\frac{r-\beta}{R}})]^{\frac{R-r}{R}}$$

$$+ \frac{1}{r} \int_{I(\delta^R(y))}^{(R-r)(1-\mu)f(\delta^R(y)y^{(r-\beta)/R})} \left[\frac{U'(c)}{\delta^R(y)}\right]^{\frac{r}{r-\beta}} dc.$$

As $y \downarrow 0$, the first two terms on the right—hand side approach zero because of (8.10). Because of (5.7), the upper limit of the integral dominates the lower one if and only if $\beta < r$, in which case

$$\left(\frac{U'(c)}{\delta^R(y)}\right)^{\frac{r}{r-\beta}} \le \left(\frac{U'(I(\delta^R(y)))}{\delta^R(y)}\right)^{\frac{r}{r-\beta}} = 1 \tag{8.19}$$

for all c in the integral of integration. In the reverse case $\beta > r$, (8.19) holds because now $c \ge I(\delta^R(y))$ and $\frac{r}{r-\beta} < 0$. Since the integrand is bounded by 1, and both limits of integration approach zero as y does, the integral also has limit zero. This concludes the proof of (8.17).

By construction, $\delta^R(y) < \delta^L(y)$, and the first inequality in (8.18) follows from (4.9) and Lemma 5.1. For the second inequality, we use the strict decrease of $\mathscr{X}(\cdot,y)$ and δ^R and inequality (8.11) to write

$$h'(y) = \mathscr{X}_\delta(\delta^R(y),y)\frac{d}{dy}\delta^R(y) + \mathscr{X}_y(\delta^R(y),y) > \mathscr{X}_y(\delta^R(y),y) > -(1+\lambda).$$

□

8.5 LEMMA. The function $\hat{V}$ satisfies

$$\lim_{y\downarrow 0}\hat{V}(h(y),y) = 0. \tag{8.20}$$

PROOF: Making the change of variables $\xi = \mathscr{X}(\delta,y)$ in (7.1), we obtain

$$\hat{V}(h(y),y) = W(y) + \int_{\delta^L(y)}^{\delta^R(y)} \delta\,\mathscr{X}_\delta(\delta,y)d\delta. \tag{8.21}$$

It is clear from (2.12) that

$$\lim W(y) = 0, \tag{8.22}$$

so we concentrate on the integral in (8.21). This integral is positive; we seek an upper bound which approaches zero as y does.

Let us first consider the case $\beta = r$. Recall (5.3) and use the change of variables $c = I(\delta)$ to write

$$\int_{\delta^L(y)}^{\delta^R(y)} \delta \mathscr{X}_\delta(\delta,y)\,d\delta \le \int_{\delta^R(y)}^{\delta^L(y)} \frac{\delta}{r}\,(-I'(\delta))\,d\delta$$

$$= \frac{1}{r}\int_{I(\delta^L(y))}^{I(\delta^R(y))} U'(c)\,dc$$

$$\le \frac{1}{r}\,U(I(\delta^R(y))).$$

Because of (8.10), $\lim_{y \downarrow 0} I(\delta^R(y)) = 0$, and we have the desired result.

We next consider the case $\beta < r$ and recall from (5.4) that

$$\int_{\delta^L(y)}^{\delta^R(y)} \delta \mathscr{X}_\delta(\delta,y)\,d\delta = \frac{1}{r-\beta}\int_{\delta^R(y)}^{\delta^L(y)} \int_{I(\delta)}^{(R-r)(1-\mu)f(\delta y^{(r-\beta)/R})} \left(\frac{U'(c)}{\delta}\right)^{\frac{r}{r-\beta}} dc\,d\delta.$$

We reverse the order of integration. Note that $I(\delta) \le c$ if and only if $U'(c) \le \delta$, and $c \le (R-r)(1-\mu)f(\delta y^{\frac{r-\beta}{R}})$ if and only if

$$\delta y^{\frac{r-\beta}{R}} \le g\left(\frac{c}{(R-r)(1-\mu)}\right) = U'(c)\left[\frac{c}{(R-r)(1-\mu)}\right]^{\frac{r-\beta}{R}}.$$

Therefore,

$$\int_{\delta^L(y)}^{\delta^R(y)} \delta\mathscr{X}_\delta(\delta,y)\,d\delta \le \frac{1}{r-\beta}\int_{I(\delta^L(y))}^{(R-r)(1-\mu)f(\delta^R(y)y^{(r-\beta)/R})} \int_{U'(c)}^{U'(c)[\cdots]^{(r-\beta)/R}} \left(\frac{U'(c)}{\delta}\right)^{\frac{r}{r-\beta}} d\delta\, dc$$

$$= \frac{1}{\beta}\int_{y(R-r)(1-\mu)}^{(R-r)(1-\mu)f(\delta^R(y)y^{(r-\beta)/R})} U'(c)\left[1-\left(\frac{c}{y(R-r)(1-\mu)}\right)^{R}\right]dc$$

$$\le \frac{1}{\beta}\int_{0}^{(R-r)(1-\mu)f(\delta^R(y)y^{(r-\beta)/R})} U'(c)\,dc$$

$$= \frac{1}{\beta}\, U[(R-r)(1-\mu)f(\delta^R(y)y^{\frac{r-\beta}{R}})] ,$$

and this has limit zero as $y \downarrow 0$ because of (8.10).

The case $\beta > r$ can be treated similarly. Using the equality

$$(R-r)(1-\mu)f(\delta^L(y)y^{(r-\beta)/R}) = (R-r)(1-\mu)f(g(y)) = y\,(R-r)(1-\mu),$$

one obtains

$$\int_{\delta^L(y)}^{\delta^R(y)} \delta\mathscr{X}_\delta(\sigma,y)\,d\delta \le \frac{1}{\beta-r}\int_{y(R-r)(1-\mu)}^{I(\delta^R(y))}\int_{U'(c)[\frac{c}{y(R-r)(1-\mu)}]^{(r-\beta)/R}}^{U'(c)} \left(\frac{U'(c)}{\delta}\right)^{\frac{r}{r-\beta}} d\delta\, dc$$

$$\le \frac{1}{\beta}\, U(I(\delta^R(y))) ,$$

and again (8.10) gives the desired result. □

8.6 SUMMARY. By definition, $h(y) \triangleq S(\wedge(y),y) < c5''(0,y)$, so $\bar{3}$ defined at the beginning of Section 7 contains

$$NT\ 4\ \{(x,y)\ e\ \bullet \mid y \geq 0,\ x \leq h(y)\}.$$

Thus $\hat{V}$ is defined on NT. We have

$$\hat{V}x(h(y),y) = A(h(y),y) = \Delta(\mathscr{S}(\delta^{R}(y)_{>y}) = 6\backslash y) \quad V \quad y > 0.$$

In other words, $\hat{V}_x(\bullet,y)$ takes the value $6^L(y)$ at $x = -(1—/x)y$ and decreases to the value $\delta^R(y)$ at the free boundary $x = h(y)$. In particular,

$$W(y) \leq \hat{V}(x,y) \leq \hat{V}(h(y),y)\ \ V\ \ x\ e\ [-(l-\wedge y.hCy)],\ \ V\ y > 0.$$

From (8.20), (8.22), we see that $\hat{V}I$ is continuous at (0,0), i.e.,

$$\lim_{\substack{(x,y)-(0,0) \\ (x,y)\ e\ NT}} \hat{V}(x,y) = 0. \tag{8.23}$$

Because $\lim 6\backslash y) = w$, we have

$$\lim_{\substack{(x,yH0,0) \\ (x,y)\ e\ NT}} \hat{V}_x(x,y) = \infty. \tag{8.24}$$

On the boundary itself,

$$(l+A)\hat{V}_x(h(y),y) - \hat{V}_y(h(y),y) = F(p(y)) = 0\ \ V\ y > 0. \tag{8.25}$$

9. Construction of the Value Function on $\mathscr{S}$.

We partition $\mathscr{S}$ into the regions NT and SB of (3.10) and (3.11).

9.1 Lemma. For each $(x,y) \in SB$, there exists a unique point $(h(\bar{y}),\bar{y}) \in NT$ satisfying (3.13). Moreover, $\bar{y} > y$.

Proof: Because of Lemma 8.4, the function $\varphi(z) \triangleq h(z) + (1+\lambda)z$ satisfies $\varphi(0) = 0$, $\varphi'(z) > 0 \ \forall\, z > 0$, and $\lim_{z\to\infty} \varphi(z) = \infty$. Therefore, there is a unique $\bar{y} \geq 0$ such that $\varphi(\bar{y}) = x + (1+\lambda)y$. Now, $x > h(y)$, so $\varphi(y) < \varphi(\bar{y})$, and thus $y < \bar{y}$. □

For $(x,y) \in SB$, we denote by $(\bar{x}(x,y), \bar{y}(x,y))$ the point on the free boundary constructed in Lemma 9.1. The functions $\bar{y}$ and $\bar{x} = h \circ \bar{y}$ are C^1 on $SB\backslash\partial_2 S$, and

$$\bar{y}_x = \frac{1}{h'(\bar{y}) + 1+\lambda}, \ \bar{y}_y = \frac{1+\lambda}{h'(\bar{y}) + 1+\lambda}, \ \bar{x}_x = h'(\bar{y})\bar{y}_x, \ \bar{x}_y = h'(\bar{y})\bar{y}_y. \tag{9.1}$$

Define V on $\mathscr{S}$ to be

$$V(x,y) \triangleq \begin{cases} \hat{V}(x,y) & \text{if } (x,y) \in NT, \\ \hat{V}(\bar{x}(x,y),\bar{y}(x,y)) & \text{if } (x,y) \in SB. \end{cases} \tag{9.2}$$

We will ultimately show that V is the value function.

9.2 Lemma. The function V is continuous on $\mathscr{S}$, C^1 on $\mathscr{S}\backslash\partial_2\mathscr{S}$, and for every $(x,y) \in \overline{SB}\backslash\partial_2\mathscr{S}$,

$$V_x(x,y) = \hat{V}_x(\bar{x}(x,y), \bar{y}(x,y)), \quad V_y(x,y) = \hat{V}_y(\bar{x}(x,y),\bar{y}(x,y)), \tag{9.3}$$

$$(1+\lambda)\, V_x(x,y) - V_y(x,y) = 0. \tag{9.4}$$

If $(x,y) \in \partial_2\mathscr{S}$, then

$$\lim_{\substack{(\xi,\eta)\to(x,y)\\(\xi,\eta)\in\mathscr{S}}} V_x(\xi,\eta) = \infty. \tag{9.5}$$

Proof: Formulas (9.3) and (9.4) follow immediately from (8.25) and (9.1). From (8.24) and (9.3), we have (9.5). □

9.3 Theorem. The function V is a nonnegative, continuously differentiable solution to the variational inequality (3.2) satisfying the boundary conditions (3.14) and (3.15).

Proof: Because $\mathscr{X}$ satisfies (4.4), V satisfies (4.3) on $NT\backslash\{(0,0)\}$. On $\partial_2\mathscr{S}\backslash\{(0,0)\}$, we have from (7.1), (7.2), (6.3), (7.4), (3.1) and (2.17) that

$$\beta V(x,y) - rxV_x(x,y) - RyV_y(x,y) + \tilde{U}(V_x(x,y))$$

$$= \beta W(y) + ry(1-\mu)\delta^L(y) - Ry[W'(y) + (1-\mu)\delta^L(y)]$$

$$+ \delta^L(y)\, I(\delta^L(y)) - U(I(\delta^L(y)))$$

$$= \beta W(y) - RyW'(y) - U(y(R-r)(1-\mu)) = 0.$$

If we now integrate (4.3) with respect to x, we discover that

$$\beta V - rxV_x - RyV_y + \tilde{U}(V_x) = 0 \text{ on } NT\backslash\{(0,0)\}. \tag{9.6}$$

It remains to show that

$$-(1-\mu)\,V_x + V_y > 0 \qquad \text{on } \mathscr{S}\backslash\partial_2\mathscr{S}, \tag{9.7}$$

$$(1+\lambda)\,V_x - V_y > 0 \qquad \text{on } \partial_1\mathscr{S} \cup \text{int(NT)}, \tag{9.8}$$

$$\beta V - rxV_x - RyV_y - \tilde{U}(V_x) > 0 \text{ on } SB\backslash\partial_2\mathscr{S}. \tag{9.9}$$

From Lemma 7.1 and (8.2) we have that for every $(x,y) \in \partial_1\mathscr{S} \cup \text{int(NT)}$,

$$V_x(x,y) = \Delta(x,y),\; V_y(x,y) = (1+\lambda)\,\Delta(x,y) - F\Big(\frac{\Delta(x,y)}{\delta^L(y)}, y\Big).$$

Therefore,

$$-(1-\mu)V_x(x,y) + V_y(x,y) = (\lambda+\mu)\Delta(x,y) - F\Big(\frac{\Delta(x,y)}{\delta^L(y)}, y\Big), \tag{9.10}$$

$$(1+\lambda)V_x(x,y) - V_y(x,y) = F\Big(\frac{\Delta(x,y)}{\delta^L(y)}, y\Big),$$

and both these expressions are positive in $\partial_1\mathscr{S} \cup \text{int(NT)}$ because of Corollary 8.2 and the fact that

$$\rho(y)\delta^L(y) = \delta^R(y) < \Delta(x,y) \le \delta^L(y) \;\; \forall\, x \in [-(1-\mu)y, h(y))$$

(see Summary 8.6). Indeed, (9.10) is also positive when $x = h(y)$, $y > 0$, and this positivity extends to all of $\mathscr{S}\backslash\partial_2\mathscr{S}$ because of (9.3).

To prove (9.9), we use Lemma 9.2 and (9.6). For $(x,y) \in SB\backslash\partial_2S$, let $\bar{x} = \bar{x}(x,y)$, $\bar{y} = \bar{y}(x,y)$, and write

$$\beta V(x,y) - rxV_x(x,y) - RyV_y(x,y) - \tilde{U}(V_x(x,y))$$

$$= \beta V(\bar{x},\bar{y}) - r\bar{x}V_x(\bar{x},\bar{y}) - R\bar{y}V_y(\bar{x},\bar{y}) - \tilde{U}(V_x(\bar{x},\bar{y}))$$

$$+ r(\bar{x}-x)V_x(\bar{x},\bar{y}) + R(\bar{y}-y)V_y(\bar{x},\bar{y})$$

$$= r(h(\bar{y}) - x)V_x(\bar{x},\bar{y}) + R(\bar{y}-y)(1+\lambda)V_x(\bar{x},\bar{y})$$

$$> r[h(\bar{y}) + (1+\lambda)\bar{y} - x - (1+\lambda)y]\, V_x(\bar{x},\bar{y}) = 0.$$

□

10. Concavity of V on $\mathscr{S}$.

A function Φ defined on a domain $\mathscr{C} \subset \mathbb{R}^n$ is said to be locally concave at a point $\underset{\sim}{x} \in \mathscr{C}$ if there exists $\varepsilon > 0$ such that the open ball $B_\varepsilon(\underset{\sim}{x}) \triangleq \{\underset{\sim}{y} \in \mathbb{R}^n \mid |\underset{\sim}{y}-\underset{\sim}{x}| < \varepsilon\}$ is a subset of $\mathscr{C}$ and Φ is concave on $B_\varepsilon(\underset{\sim}{x})$. If $\mathscr{C}$ is an open subinterval of $\mathbb{R}$ and Φ is locally concave on $\mathscr{C}$, the the right– and left–derivatives of Φ exist on $\mathscr{C}$, are nonincreasing, and are right– and left–continuous, respectively (see, e.g., Karatzas & Shreve [6], Problem 6.19, p. 212 and Solution 6.19, p. 234). This implies, in turn, that Φ is concave on $\mathscr{C}$ (see, e.g. ibid, Problem and Solution 6.20). By applying this one–dimensional result along lines in $\mathbb{R}^n$, we obtain the following lemma.

10.1 Lemma. Let Φ be defined on an open domain $\mathscr{C} \subset \mathbb{R}^n$. If Φ is locally concave at every point in $\mathscr{C}$, and $\mathscr{C}$ is a convex set, then Φ is a concave function.

10.2 Lemma. Let Φ be a C^1 function defined on a convex, open domain $\mathscr{C} \subset \mathbb{R}^n$. The function Φ is concave if and only if

$$\Phi(\underset{\sim}{y}) - \Phi(\underset{\sim}{x}) \le \nabla\Phi(\underset{\sim}{x}) \cdot (\underset{\sim}{y} - \underset{\sim}{x}) \quad \forall\ \underset{\sim}{x},\underset{\sim}{y} \in \mathscr{C}. \tag{10.1}$$

PEOOF: We show that (10.1) implies concavity. Let $\tilde{x},\tilde{y}$ € ^ be given. Then (10.1) implies

$$\frac{1}{2}\Phi(x) + \frac{1}{2}\Phi(y) - \Phi(\frac{1}{2}x + \frac{1}{2}y)$$

= \ [*(?) - *(£ x + 1 y)] + *[*(y) - *(J x + \ y)]

$$\leq \frac{1}{2}\nabla\Phi(\frac{1}{2}x + \frac{1}{2}y) \cdot [x - (\frac{1}{2}x + \frac{1}{2}y) + y - (\frac{1}{2}x + \frac{1}{2}y)]$$

$$= 0.$$ □

10.3 THEOEEM. The function V defined by (9.2) is concave on $\mathscr{S}$.

PEOOF: Since V is continuous on <2f> it suffices to prove concavity on int(G^). For this, we prove local concavity and then appeal to Lemma 10.1. From Proposition 7.3, we already have local concavity of V on int(NT), and it remains to prove local concavity on 5H\5₂<^.

Let (xO)yo) € SF\c?2<^ be given and define $\bar{x}_0 = \bar{x}(x_o,yo)$, $\bar{y}o = \bar{y}(xojo)$- Then $(\bar{x}o,\bar{y}o)$ £ NT, and we can choose $e_0 > 0$ such that $B_{e_0}(\bar{x}_o,\bar{y}o)$ C 3. Now choose $e > 0$ such that $B_e(x_o,yo)$ n NT c $B_{£_o}(\bar{x}_o,\bar{y}_o)$ and $(\bar{x}(x,y), \bar{y}(x,y))$ € $B_{e_o}(\bar{x}_o,\bar{y}_o)$ V (x,y) G $B_e(x_0,y_0)$ n SB.

We shall prove that V is concave on $B_e(x_o,yo)$ by checking the condition of Lemma 10.2.

Let (xiji), $(x_{2)}y_2)$ be in $B_e(x_o,yo)$. For i = 1,2, if (xi,yi) G SB, define $\bar{x}i = \bar{x}(xi,yi)$, $\bar{y}_i = \bar{y}(xi,yi)$; if (xi,yi) G NT, set $\bar{x}i = xi$, $\bar{y}i = yi$. Note from (3.13) that in either case

$$-(x_i - \bar{x}_i) = (1+\lambda)(y_i - \bar{y}_i). \qquad (10.2)$$

From (8.25) we condude that

$$(x_i - \bar{X}i)\ \hat{V}_x(\bar{x}_b\bar{y}i) + (y_i - \bar{y}i)\ \hat{V}_y(\bar{x}i,\bar{y_i}) = 0. \qquad (10.3)$$

Using the convexity of $\hat{V}$ on $B_{\epsilon_0}(\bar{x}_0,\bar{y}_0)$ and (9.3), we have

$$V(x_2,y_2) - V(x_1,y_1) = \hat{V}(\bar{x}_2,\bar{y}_2) - \hat{V}(\bar{x}_1,\bar{y}_1)$$

$$\le (\bar{x}_2 - \bar{x}_1)\hat{V}_x(\bar{x}_1,\bar{y}_1) + (\bar{y}_2 - \bar{y}_1)\hat{V}_y(\bar{x}_1,\bar{y}_1)$$

$$= (x_2 - x_1)\hat{V}_x(\bar{x}_1,\bar{y}_1) + (y_2 - y_1)\hat{V}_y(\bar{x}_1,\bar{y}_1) + [(\bar{x}_2 - x_2)\hat{V}_x(\bar{x}_1,\bar{y}_1)$$

$$+ (\bar{y}_2 - y_2)\hat{V}_y(\bar{x}_1,\bar{y}_1)] - [(\bar{x}_1 - x_1)\hat{V}_x(\bar{x}_1,\bar{y}_1) + (\bar{y}_1 - y_1)\hat{V}_x(\bar{x}_1,\bar{y}_1)]$$

$$= (x_2 - x_1)V_x(x_1,y_1) + (y_2 - y_1)V_y(x_1,y_1).$$

□

11. The Optimal Trajectories

We show in this section that if we start the position processes at $(X_{0-},Y_{0-}) = (x,y) \in \mathcal{NT}$, use the feedback consumption process

$$C_t \triangleq I(V_x(X_t,Y_t)), \tag{11.1}$$

and do not trade ($M \equiv 0$, $N \equiv 0$), then the position process stays in $\mathcal{NT}$, moving in finite time to the boundary $\partial_1\mathcal{NT}$. We have already observed (see Remark 3.3) that once the position process arrives at $\partial_1\mathcal{NT}$, it stays on this boundary.

Let $(X_{0-},Y_{0-}) = (x,y) \in \mathcal{NT}$, and define X and Y by the formulas

$$Y_t \triangleq y\,e^{Rt}, \quad X_t \triangleq x + \int_0^t [rX_s - I(V_x(X_s,Y_s))]ds. \tag{11.2}$$

We also define the *derivative process*

$$S_t \triangleq V_x(X_t,Y_t). \tag{11.3}$$

As long as (X_t,Y_t) remains in NT, we have X_t = c5"(^$_t$,Y_t) and

$$dX_t = \#_b\{6_t,Yt)\ \&5_t + JZ_y\{6t,Y_t)\ R\ Y_t\ dt. \qquad (11.4)$$

Comparison of (11.2) with (11.4) shows that

$$\mathscr{X}_\delta(\delta_t,Y_t)\,d\delta_t = [r\mathscr{X}(\delta_t,Y_t) - l\{6_t) - J?T_y(\pounds_t)Yt)RY_t]dt. \qquad (11.5)$$

Invoking (4.4), we obtain the first—order linear differential equation

$$\mathscr{X}_\delta(\delta_t,Y_t)d\delta_t = (\beta-r)\ \delta_t\ \mathscr{X}_\delta(\delta_t,Y_t)dt. \qquad (11.6)$$

Define

$$T = \inf\{t \geq 0 \mid X_t = -(l-/i)Yt \text{ or } X_t = h(Y_t)\}. \qquad (11.7)$$

Because of (5.6), we can cancel c£"$_6$(£$_t$,Yt) from (11.6) and thereby obtain

$$S_t = 6ae(P\sim^T)^X, \qquad 0 \leq t \leq r, \qquad (11.8)$$

where $So = V_x(x,y)$.

If r is finite and X, = h(Y,), then

$$\delta_\tau = V_x(h(Y_\tau),Y_\tau) = \delta^R(Y_\tau).$$

On the other hand, if r is finite and X, = -(l-^)Y,, then

$$\delta_\tau = V_x(-(1-\mu)Y_\tau,Y_\tau) = \delta^L(Y_\tau).$$

Finally, (X_t,Y_t) e NT if and only if $\wedge(Y_t) \leq$ «t $\leq$ $*^L(Y_t)$. In particular, $0 < 6\backslash y) < 6_0 \leq \delta^L(y)$.

11.1 LEMMA. We have $r <$ a> and (X,,Y,,) e $d\backslash df$, i.e.,

$$\delta_t > «*(Yt), \quad 0 < t \leq r, \tag{11.9}$$

$$\delta_\tau = \delta^L(Y_\tau). \tag{11.10}$$

PIOOF: Let Z_t = fc - ^"(Y_t). From (8.12), (8.11) we have

$$\frac{d}{dt} Z_t = (3 - r \cdot \delta_t - \frac{(6-R)(\backslash+X) + (\beta-r)\mathcal{S}_y(\delta^R(Y_t),Y_t)}{1 + A + \mathcal{S}_y(\delta^R(Y_t),Y_t)} \delta^R(Y_t)$$

$$= (\beta-r)Z_t + \frac{(R-ilQ+Al}{1 + A + Sj\{S\backslash Y_t),Y_t)} \delta^R(Y_t)$$

$$> (/5-r)Z_t.$$

Therefore, $\frac{d}{dt}$J (e"XK~r^tZt) > 0 and (11.9) follows from the initial condition $So \geq \delta^R(y)$.

To show that (11.10) holds for some finite r, we write

$$\delta_t - \delta^L(Y_t) = \delta_0 e^{(\beta-r)t} - U'(y\ e^{at}(R-r)(l-/x))$$

$$= e^{(\beta-r)t} [\delta_0 - (e^{at})^{\frac{r}{l}}\ U'(y\ e^{lt}(R-*)(H*))].$$

This quantity is nonpositive at $t = 0$ but becomes positive as t -» OD because of (2.16).

□

11.2 REMAU. For $t > r$, formula (11.8) cannot be derived from (11.6) because $\wedge_6(\&_T,Y_1) = 0$

(see (5.5)). Indeed the boundary condition (3.14) shows that

$$\delta_t = U'(y\, e^{Rt}(R-r)(1-\mu)), \quad t > \tau.$$

12. Proof of Theorem 3.2.

We show in this section that the function V defined by (9.2) is the value function V^* of (2.9). We already know from Lemma 3.1 and Theorems 9.3, 10.3 that $V \geq V^*$. Because of Remark 2.1, $V^* \equiv 0$ on $\partial_2\mathscr{S}$, and this agrees with the definition $V(x,y) = \hat{V}(0,0) = 0$ $\forall\ (x,y) \in \partial_2\mathscr{S}$. It remains to show that

$$V(x,y) \leq V^*(x,y) \quad \forall\ (x,y) \in \mathscr{S}\backslash\partial_2\mathscr{S}. \tag{12.1}$$

We first consider the case $(x,y) \in \partial_1\mathscr{S}$ and show that when $X^*_{0-} = x$, $Y^*_{0-} = y$ and M^*, N^*, C^* are as in Theorem 3.2, then

$$\lim_{t\to\infty} e^{-\beta t}\, V(X^*_t, Y^*_t) = 0. \tag{12.2}$$

Because V satisfies the boundary condition (3.14) (Theorem 9.3), C^*, X^* and Y^* are given by (2.10), (2.11), and

$$e^{-\beta t}\, V(X^*_t, Y^*_t) = e^{-\beta t}\, W(ye^{Rt})$$

$$= \int_0^\infty e^{-\beta(t+s)}\, U(ye^{R(t+s)}\,(R-r)(1-\mu)]ds$$

$$= \int_t^\infty e^{-\beta s}\, U[ye^{Rs}\,(R-r)(1-\mu)]ds.$$

This expression has limit zero as $t \to \infty$ because of Assumption I. Now suppose $(X^*_{0-},Y^*_{0-}) \in \mathscr{S}$. Then $(X^*_0,Y^*_0) \in NT$, and the discussion in Section 11 shows that (X^*,Y^*) reaches $\partial_1\mathscr{S}$ in finite time. Therefore, (12.2) holds when M^*, N^*, C^* are as in Theorem 3.2, regardless of the initial position in $\mathscr{S}$.

Let $(x,y) \in \mathscr{S}\backslash\partial_2\mathscr{S}$ be given, and let $X^*_{0-} = x$, $Y^*_{0-} = y$ and M^*, N^*, C^* be as in Theorem 3.2. Then $(X_t,Y_t) \in NT$ for all $t \geq 0$ (even if $(x,y) \notin NT$, because in this case, there is an immediate jump to the free boundary). As in the proof of Lemma 3.1, we may use the chain rule rule to write

$$V(x,y) - e^{-\beta t}\, V(X^*_t,Y^*_t)$$

$$= \int_0^t e^{-\beta s}[\beta V(X^*_s,Y^*_s) - rX^*_s V_x(X^*_s,Y^*_s) - RY^*_s V_y(X^*_s,Y^*_s) + C^*_s V_x(X^*_s,Y^*_s)]ds$$

$$+ \int_0^t e^{-\beta s}[(1+\lambda)V_x(X^*_{s-},Y^*_{s-}) - V_y(X^*_{s-},Y^*_{s-})]dM^*_s$$

$$+ \sum_{0\leq s\leq t} e^{-\beta s}[-V(X^*_s,Y^*_s)+ V(X^*_{s-},Y^*_{s-})+ V_x(X^*_{s-},Y^*_{s-})\Delta X^*_s+ V_y(X^*_{s-},Y^*_{s-})\Delta Y^*_s].$$

The first term on the right–hand side is equal to $\int_0^t e^{-\beta s}\, U(C^*_s)ds$ because V satisfies (3.6) (see Theorem 9.3, especially (9.6) in its proof, and also (3.1)). The second term is zero because after the initial time, M^* is constant, and if M^* jumps at the initial time, then

$$(1+\lambda)V_x(X^*_{0-},Y^*_{0-}) - V_y(X^*_{0-},Y^*_{0-}) = 0$$

(see Lemma 9.2). The third term is zero because if M^* jumps at the initial time, then

$$V(X_0^*,Y_0^*) = V(X_{0-}^*,Y_{0-}^*),$$

$$V_x(X_{0-}^*,Y_{0-}^*)\Delta X_0^* + V_y(X_{0-}^*,Y_{0-}^*)\Delta Y_0^*$$

$$= \Delta M_0^*[-(1+\lambda)V_x(X_{0-}^*,Y_{0-}^*) + V_y(X_{0-}^*,Y_{0-}^*)] = 0.$$

Therefore,

$$V(x,y) - e^{-\beta t}\, V(X_t^*,Y_t^*) = \int_0^t e^{-\beta s} U(C_s^*)ds,$$

and taking $t \to \infty$, using (12.2), we we obtain

$$V(x,y) = \int_0^\infty e^{-\beta s}\, U(C_s^*)ds \le V^*(x,y).$$

□

13. Power Utility Functions.

In this section we specialize the model to the case that the utility function for consumption is of the form $U(c) = \frac{c^p}{p}$ for some $p \in (0,1)$. Then $U'(c) = c^{p-1}$ and $I(\delta) = \delta^{\frac{1}{p-1}}$. The function g defined by (2.13) becomes

$$g(y) = [(R-r)(1-\mu)]^{p-1}\, y^{-\frac{R-r\ +\ \beta-Rp}{R}}, \quad y > 0, \tag{13.1}$$

and its inverse is

$$f(\alpha) = [(R-r)^{1-p}(1-\mu)^{1-p}\alpha]^{-\frac{R}{R-r\ +\ \beta-Rp}}, \quad \alpha > 0. \tag{13.2}$$

As observed in Remark 2.2, Assumption II implies that

$$\beta - Rp > \frac{(R-r)(1-\mu)}{\lambda+\mu}, \tag{13.3}$$

an assumption which is in force throughout this section.

A straight–forward but tedious evaluation of (5.1) results in the key formula, valid both when $\beta = r$ and $\beta \neq r$,

$$\mathscr{S}(\delta,y) = \left(\frac{1-p}{\beta-rp}\right)\delta^{\frac{1}{p-1}} - \frac{1}{k(R-r)}[(1-\mu)(R-r)y]^{k}\,\delta^{\frac{1-k}{p-1}}, \; \delta > 0, \tag{13.4}$$

where

$$k \triangleq \frac{\beta-rp}{R-r+\beta-Rp}. \tag{13.5}$$

The function F of (8.1) becomes

$$F(\rho,y) = [(R-r)(1-\mu)y]^{p-1}\rho\left\{(1+\lambda)-(1-\mu)\left(\frac{R-r+\beta-Rp}{\beta-Rp}\right)\rho^{\frac{1-k}{p-1}}\right\}, \tag{13.6}$$

and this function takes the value zero when ρ is

$$\rho(y) \equiv \left[\frac{(1-\mu)(R-r+\beta-Rp)}{(1+\lambda)(\beta-Rp)}\right]^{\frac{p-1}{k-1}}. \tag{13.7}$$

In particular,

$$\delta^{L}(y) = [(R-r)(1-\mu)y]^{p-1}, \; \delta^{R}(y) = [(R-r)(1-\mu)y]^{p-1}\left[\frac{(1-\mu)(R-r+\beta-Rp)}{(1+\lambda)(\beta-Rp)}\right]^{\frac{p-1}{k-1}}, \tag{13.8}$$

and the free boundary is (see (8.3))

$$h(y) = ay, \quad y \geq 0, \tag{13.9}$$

where

$$a \triangleq \frac{1-\mu}{k}\left[(1-k)\, b^{\frac{1}{1-k}} - b\right], \tag{13.10}$$

$$b \triangleq \frac{(1+\lambda)(\beta - Rp)}{(1-\mu)(R - r + \beta - Rp)}. \tag{13.11}$$

Just as in Davis & Norman [4], the region NT is a cone.

13.1 REMALK. Assumption II in the form (13.3) is a necessary condition for the free boundary to be in the interior of *df*. To see this, define

$$\nu(\gamma) \triangleq \frac{(1+\lambda)\gamma}{(1-\mu)(R-r+\gamma)}, \ \forall\ \gamma > 0,$$

$$\varphi(\nu) \triangleq \frac{1-\mu}{k}\left[(1-k)\, \nu^{\frac{1}{1-k}} - \nu\right], \ \forall\ \nu > 0,$$

so $b = \nu(\beta - Rp)$ and $a = \varphi(b)$. Note that

$$\nu\left(\frac{(R-r)(1-\mu)}{\lambda+\mu}\right) = 1, \ \ \varphi(1) = -(1-\mu),$$

$$\nu'(\gamma) = \frac{(1+\lambda)(R-r)}{(1-\mu)(R-r+\gamma)^2} > 0 \quad \forall\ \gamma > 0$$

$$\varphi'(\nu) = \frac{1-\mu}{k}\left[\nu^{\frac{k}{1-k}} - 1\right] > 0, \ \forall\ \nu > 1,$$

because $\frac{k}{1-k} = \frac{(R-r)(1-p)}{R-r+\beta-Rp} > 0$. Therefore,

$$(\varphi \circ \gamma)(\nu) > -(1-\mu) \quad \forall \ \gamma > \frac{(R-r)(1-\mu)}{\lambda+\mu}.$$

In particular, a > -(1-/*) because of (13.3). □

Just as in the proof of Lemma 7.1, one can change variables in (7.1) to obtain the formula

$$V(x,y) = W(y) + \int_{L(x,y)}^{\Delta(x,y)} SJt\wedge SjW \quad V(x,y) \in NT. \tag{13.12}$$

However, the inversion of ^(*,y) in (13.4), necessary to determine A(«,y), is not algebraically possible, and it thus does not seem possible to obtain a closed—form formula for V in NT. One can obtain such a formula for V on the free boundary, because A(h(y),y) = 6^{L}(y) (see Summary 8.6). We have

$$V(x,y) = W(y) + \int_{\delta^{L}(x,y)}^{\delta^{R}(x,y)} \$J?T_6(£,y)d£=cyP, \qquad y > 0, \tag{13.13}$$

where

$$c = \frac{(R-r)^p(1-\mu)^p}{p(\beta-Rp)} + \frac{(R-r)^p(1-\mu)^p}{(\beta-rp)}\left\{\frac{1-p}{1-p}\, b^{\frac{p}{1-k}} + \frac{1-p}{p-k}\, b^{\frac{p-k}{1-k}} + \frac{p-1}{p} + \frac{p-1}{p-k}\right\}. \tag{13.14}$$

For (x,y) e SB, $(\bar{x},\bar{y})$ constructed in Lemma 9.1 is given by the formulas (see (3.13))

$$\bar{y} = \frac{x}{1+A} + \frac{\beta+AW}{1+A}, \quad \bar{x} = a\,\bar{y}. \tag{13.15}$$

Because $V(x,y) = V(\bar{x},\bar{y})$, we obtain

$$V(x,y) = c\left[\frac{x}{a} + \frac{(1+\lambda)y}{1+\lambda}\right]^p \quad V(x,y) \in SB, \tag{13.16}$$

where c is given by (13.14), and k, a, b are given by (13.5), (13.10), (13.11).

13.2 COICLUSIOI. If the agent is given an initial position (x,y) e SB, he should sell the bond and buy "stock" so as to immediately bring his position to $(\bar{x},\bar{y})$ given by (13.15). He should do no further trading. He should set *S(6j)* defined by (13.4) equal to $\bar{x}$ and solve for *S*. Let S_o satisfy $S(6_{0y}\bar{y}) = \bar{x}$. The agent's consumption should be $C_t \triangleq [S_o e^{(\beta-r)t}]^{\frac{1}{p-1}}$ (see (11.1), (11.3), (11.8)) until time r when his position reaches *d\df*, where X^ = — (l ɇ)Y r Thereafter his consumption should be $C_t \triangleq Y_1 e^{E\{t\sim r}$)(R--r)(l--/z) (see Remark 11.2).

If the utility function U(c) is not of the form $\frac{1}{p}c^p$, the above description of the optimal policy is still correct, except the explicit formulas (13.15) and (13.4) must be replaced by their more general versions (3.13) and (5.1), and, of course, $\bar{x} \triangleq h(\bar{y})$.

14. References.

[1] Beneš, V.E., Shepp, L.A. and Witsenhausen, H.S., Some solvable stochastic control problems, *Stochastics* 4 (1980), 39-33.

[2] Chow, P.L., Menaldi, J.L., and Robin, M., Additive control of stochastic linear systems with finite horizon, *SIAM J. Control and Optimization* 23 (1985), 858-899.

[3] Constantinides, G.M., Capital market equilibrium with transaction costs, J. *Political Economy* 94 (1986), 842-362.

[4] Davis, M.H.A. and Norman, A.R., Portfolio selection with transaction costs, *Math. Operations Research* (1989).

[5] Harrison, J.M., *Brownian Motion and Stochastic Flow Systems,* John Wiley and Sons, New York, 1985.

[6] Karatzas, I. and Shreve, S.E., *Brownian Motion and Stochastic Calculus*, Springer–Verlag, New York, 1988.

[7] Karatzas, I., Lehoczky, J., Sethi, S., and Shreve, S., Explicit solution of a general consumption/investment model, *Math. Operations Research* 11 (1986), 261–294.

[8] Magill, M.J.P. and Constantinides, G.M., Portfolio selection with transaction costs, *J. Econ. Theory* **13** (1976), 245–263.

[9] Rogers, L.C.G. and Williams, D., *Diffusions, Markov Processes, and Martingales. Volume 2: Itô Calculus*, Wiley, New York, 1987.

[10] Soner, H.M., Optimal control with state–space constraint I, *SIAM J. Control & Optimization* **24** (1986), 552–562.

[11] Soner, H.M. and Shreve, S.E., Regularity of the value function for a two–dimensional stochastic control problem, *SIAM J. Control and Optimization* **27** (1989), 876–907.

[12] Soner, H.M. and Shreve, S.E., A free boundary problem related to singular stochastic control: the parabolic case, Dept. of Mathematics, Carnegie Mellon Univ., Tech. Report, 1990.

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

[14] Zariphopoulou, Th., Investment–consumption model with transaction costs and Markov–chain parameters, Technical report LCDS/CSS #89–16, Brown University, 1989.