

Optimal Environmental Management in the Presence of Irreversibilities

José A. Scheinkman¹

Department of Economics, Princeton University, Princeton, New Jersey 08544-102
joses@princeton.edu

and

Thaleia Zariphopoulou²

Business School and Department of Mathematics, University of Wisconsin, Madison
zariphop@math.wisc.edu

Received December 16, 1998; revised October 9, 1999;
final version received October 22, 1999

We consider an environment of a fixed size that can be converted to another use. This conversion can be made in steps, but it is *irreversible*. The future benefits (per unit) from the original use, and from the alternative use, follow a diffusion process. For a fairly general case, we show that the *value function* must be the unique (viscosity) solution to the associated Hamilton-Jacobi-Bellman equation. We also exhibit several properties of the solution for the case of constant relative risk aversion between 0 and 1, and a log-linear diffusion for the benefits. *Journal of Economic Literature* Classification Numbers: C61, D90, Q30. © 2001 Academic Press

1. INTRODUCTION

In this paper we study a class of dynamic optimization problems, inspired by questions on the economics of environmental management, which are characterized by uncertainty and irreversibility. We consider an environment of a fixed size that can be converted to another use. Though this conversion can be made in steps, it is *irreversible*. The future flow of benefits from each possible use is uncertain. We model these future *flows* as following a diffusion. The utility flow in turn depends on these benefit flows. The presence of irreversibilities, together with the increase over time

¹ Research partially funded by the Fondazione Eni Enrico Mattei and the National Science Foundation.

² Research partially supported by a Romnes Fellowship and the Graduate School of the University of Wisconsin, Madison and the Alfred Sloan Foundation.

in the quality of information about the future values of the benefits, gives rise to what Arrow and Fisher [2] named a “quasi-option value” and Henry [12] the “irreversibility effect”³, that typically creates a bias towards a more conservative conversion policy. In fact our model is basically a generalization to many periods of the example used by Arrow and Fisher to discuss the “quasi-option” value.

We first consider a fairly general type of optimization problem and show that its *value function* must be the unique (viscosity) solution to the associated Hamilton–Jacobi–Bellman equation. We then specialize to utility functions that exhibit constant relative risk aversion parameter $1 - p$ satisfying $0 \leq p \leq 1$, and a log-linear diffusion for the benefits. In this case we can characterize the solution much more explicitly. First, the optimal action becomes simply a function of the fraction θ of the environment that has not yet been converted and of the *ratio* z of the benefit flows from the alternative use to the benefit flows from the original environment. Hence our problem becomes essentially one dimensional. As in [15], the optimal policy is fully determined by an *exercise boundary*, that is a map ψ from the ratio of benefits z to the “optimal” fraction of the environment to be preserved. If $\theta_0 \leq \psi(z)$, the optimal policy is to immediately “jump” to $\psi(z)$. Of course, since conversion is not reversible, if $\theta_0 \geq \psi(z)$, it is optimal to “stay put”. The optimal policy after time zero is to continually “deflect” from the region, $\theta \leq \psi(z)$. We discuss this particular case in Section 4. As usual, we start by using heuristic arguments. We first write down the optimization problem and proceed in Subsection 4.1 to show that the exercise boundary must satisfy a particular equation.

In Subsection 2, we present several qualitative properties of the exercise boundary. We show that there is a value $\underline{z}(p) > 1$, such that if $z \leq \underline{z}(p)$, the optimal policy is to stay put, even if $\theta = 0$. That is, development should not start at all until the benefits of the alternative project exceed that of the original environment by a certain amount. This is one of the implications of the presence of the quasi-option effect. Property 4.3 then characterizes the amount $\bar{z}(p)$ such that if $z \geq \bar{z}(p)$ then the optimal policy is to set $\theta = 1$ (full conversion). In the case of risk-neutrality ($p = 1$), $\underline{z}(1) = \bar{z}(1)$ that is, the quasi-option effect implies that a minimum threshold $\underline{z}(1) > 1$ must be reached before conversion starts, but once this threshold is reached full conversion occurs. When risk aversion is present ($p < 1$) the level of $\bar{z}(p)$ that leads to full conversion is finite, but strictly exceeds the level of $\underline{z}(p)$ required to start conversion.

Several results comparing solutions are also established. An increase in the expected rate of growth of X (or a decrease on the expected rate of growth of Y) leads to a more conservative policy. The decrease in the rate

³ Cf. also [11] and [16].

of growth of X makes it less likely that the irreversibility constraint is ever binding in the future and hence, as argued earlier by Kolstad [16], lowers the “quasi-option” value. Similarly, an increase in the discount factor leads to a less conservative policy. We also show that an increase in the correlation between innovations in X and in Y , lowers the option value of remaining at a given θ and hence leads to a less conservative policy. Finally we argue that an increase in risk-aversion, because it makes diversification more valuable, leads to a *less* conservative policy when θ is small (that is, when much of the original environment is still preserved) and to a *more* conservative policy when θ is large.

The effect of irreversibilities and changes in information have also been studied in the literature on irreversible investments (e.g. [1], [19]). In fact, the optimal investment policies in the models of Pindyck [20], Bertola [3] or Dixit [8] exhibit the same “deflection” behavior of our model in Section 4 and Kobila [15] analyzes models of irreversible investments in a singular stochastic control set up similar to ours. One novelty in our analysis is that we allow for two capital stocks instead of one. In addition, we admit risk aversion⁴. Finally, the economic problem that we study leads naturally to different comparative statics questions and results.

Even though our formulation is motivated by questions about the economics of environmental management, our results should prove useful in other set-ups in which one can convert at a fixed price from one project, or asset, to another but this conversion is irreversible. We hope to develop these connections in future work.

2. THE MODEL

As discussed in the introduction, we model an environment of fixed size which can be developed into an alternative use. The development is *irreversible* and the benefits (per unit) of the original environment and of the alternative project are random. More precisely, we suppose the per unit benefits of the environment, at instant $t \geq 0$, are given by a diffusion that satisfies

$$dX_t = \mu_1(X_t) dt + \sigma_1(X_t) dW_t^1. \quad (1)$$

The alternative project has a flow of (per unit) benefits Y_t that is also a diffusion and satisfies

$$dY_t = \mu_2(Y_t) dt + \sigma_2(Y_t) dW_t^2. \quad (2)$$

⁴ Cf. also Remark 2.2 below.

Here, W_t^1 and W_t^2 are Brownian motions defined on a probability space $\{\Omega, \mathcal{F}, P\}$ and have a correlation coefficient $\gamma \in (-1, 1)$. We assume that $X_0 \geq 0, Y_0 \geq 0$ and that the coefficients μ_i and σ_i satisfy conditions that guarantee that $X_t \geq 0$ and $Y_t \geq 0$.⁵ We will also assume that for $i = 1, 2$ the functions μ_i and σ_i are Lipschitz, that is:

$$|\mu_1(X) - \mu_1(X')| \leq K_1^\mu |X - X'|, \tag{3}$$

and

$$|\mu_2(Y) - \mu_2(Y')| \leq K_2^\mu |Y - Y'|, \tag{4}$$

and similarly for the functions σ_1 and σ_2 . Finally we assume that the dynamics are “concave” i.e. for each (t, ω) the functions that map the initial conditions X_0 into $X_t(\omega)$ and Y_0 into $Y_t(\omega)$ are concave.

Remark 2.1. This concavity holds if e.g. for each $i = 1, 2$ μ_i is concave and σ_i is linear. This can be shown by letting $X_t^{x_i}$ denote the solution to equation (1) with initial condition x_1, x_2 or $x_3 = \lambda x_1 + (1 - \lambda) x_2, \lambda \in (0, 1)$. Let $\tilde{X}_t = \lambda X_t^{x_1} + (1 - \lambda) X_t^{x_2}$. Finally, let $A_t = \tilde{X}_t - X_t^{x_3}$. Using the concavity of μ , the linearity of the volatility term and standard iteration arguments for establishing classical comparison results for the solutions of stochastic differential equations (e.g. [14], Chapt. 5, Propositions 2.13 and 2.18), we get that $E[\max\{A_t, 0\}] = 0$ a.s. for all $t \geq 0$. Hence, $\tilde{X}_t \leq X_t^{x_3}$ a.s.

If at time t a fraction θ_t of the environment has been transformed, then the utility flow is given by $U((1 - \theta_t) X_t, \theta_t Y_t)$, where U is a continuous, concave function that is increasing in its arguments, with $U(0, 0) = 0$.

The initial data is given by $X_0 = x, Y_0 = y$ and the fraction of the environment that has been transformed $\theta_0 = \theta$. We will write \mathcal{F}_t for the σ -field generated by the realizations of the pair of Brownian motions up to t . The objective is to choose an \mathcal{F}_t -progressively measurable non-negative, non-decreasing process of cumulative development M_t , to maximize

$$E \int_0^\infty e^{-\beta t} U((1 - \theta_t) X_t, \theta_t Y_t) dt, \tag{5}$$

where X_t solves (1) with $X_0 = x, Y_t$ solves (2) with $Y_0 = y$ and

$$0 \leq \theta_t = \theta + M_t \leq 1. \tag{6}$$

⁵ There are many alternative standard conditions that guarantee this property. For instance we could assume that the origin is absorbing, i.e. for $i = 1, 2, \mu_i(0) = \sigma_i(0) = 0$, or we could assume properties on μ_i and σ_i that imply that the origin is not attracting, which insures the positivity of X_t and Y_t , if we start at positive initial conditions.

For convenience we assume that $M_0 = 0$. Note that the above equation indicates that θ is a state variable, and M_t is the control process that we have to specify in order to maximize the payoff given in (5). The fact that M_t is non-decreasing captures the *irreversibility* of the development process and the constant β measures the time rate of discounting. We will write \mathcal{A}_θ for the set of non-negative, non-decreasing \mathcal{F}_t -measurable processes M_t that satisfy (6).

Remark 2.2. The formulation of the objective in equation (5) is justified if we assume that the benefits of the natural environment and of the new project are not spanned by existing assets. This is different than what is done in the irreversible investment literature where spanning is typically assumed; but seems more natural in our context. Further, we set the cost of development to zero. Equivalently, we may assume that the costs of the development are linear and that Y is the flow of *net* benefits, already accounting for the cost of development. In any case, we could easily accommodate a constant marginal cost of development.

We define the value function associated with the expected payoff (5) via:

$$v(x, y, \theta) = \sup_{\mathcal{A}_\theta} E \int_0^\infty e^{-\beta t} U((1 - \theta_t) X_t, \theta_t Y_t) dt. \quad (7)$$

PROPOSITION 2.1. *If $\beta > \max\{K_1^\mu, K_2^\mu\}$, then the value function is concave, non-decreasing in x and y and non-increasing in θ .*

Proof. Since U is concave, and $0 \leq \theta_t \leq 1$, we know that there are constants M, C_1 and C_2 such that:

$$E \int_0^\infty e^{-\beta t} U((1 - \theta_t) X_t, \theta_t Y_t) dt \leq E \int_0^\infty e^{-\beta t} (M + C_1 X_t + C_2 Y_t) dt.$$

Since $X_t \geq 0$ and $\mu_1(\cdot), \sigma_1(\cdot)$ are Lipschitz functions we have $E(X_t) = x + \int_0^t E\mu_1(X_s) ds \leq [x + \mu_1(0) t] + K_1^\mu \int_0^t E(X_s) ds$ and hence, from Gronwall's inequality, $E(X_t) \leq c_1 e^{K_1^\mu t}$ for some c_1 . Similarly, $E(Y_t) \leq c_2 e^{K_2^\mu t}$. Hence the value function is finite. The concavity follows from the concavity of U and the "concave dynamics". The fact that U is increasing and $0 \leq \theta_t \leq 1$ guarantees that v is non-decreasing with respect to x and y . Finally, since $\mathcal{A}_{\theta_1} \subseteq \mathcal{A}_{\theta_2}$ if $\theta_2 \leq \theta_1$, v is non-increasing in θ .

Notice that the role of the bound on β given by the Lipschitz constants is to insure the finiteness of the value function. In order to establish uniqueness or, more generally, comparison results, it is more convenient to

strengthen the assumption on the size of the discount factor relative to the Lipschitz constant slightly to:

$$\beta > \max\{K_1^\mu, K_2^\mu\} + 1. \tag{8}$$

This assumption is only *apparently* stronger. If $\beta > \max\{K_1^\mu, K_2^\mu\}$ we may always (linearly) change the measurement of time so that (8) holds.

Remark 2.3. Since θ_t is non-decreasing, $v(x, y, 1) = E \int_0^\infty e^{-\beta t} U(0, Y_t) dt$.

Remark 2.4. If $\mu_1(0) = \sigma_1(0) = 0$, then if $x = 0$, $X_t \equiv 0$ for each $t \geq 0$. Consequently,

$$v(0, y, \theta) = \sup_{\mathcal{A}_\theta} E \int_0^\infty e^{-\beta t} U(0, \theta_t Y_t) dt = E \int_0^\infty e^{-\beta t} U(0, \theta Y_t) dt.$$

Remark 2.5. Similarly, if $\mu_2(0) = \sigma_2(0) = 0$, then if $y = 0$, $Y_t \equiv 0$ for each $t \geq 0$. Hence $v(x, 0, \theta) = \sup_{\mathcal{A}_\theta} E \int_0^\infty e^{-\beta t} U((1 - \theta_t) X_t, 0) dt = E \int_0^\infty e^{-\beta t} U((1 - \theta) X_t, 0) dt$.

Recall that a *stopping time* τ is a non-negative random variable such that the event $\{\tau \leq t\}$ is in \mathcal{F}_t , for each $t \geq 0$. We next state the *dynamic programming principle*:

THEOREM 7. *If τ is a stopping time then:*

$$v(x, y, \theta) = \sup_{\mathcal{A}_\theta} E \left[\int_0^\tau e^{-\beta s} U((1 - \theta_s) X_s, \theta_s Y_s) ds + e^{-\beta \tau} v(X_\tau, Y_\tau, \theta_\tau^*) \right] \tag{9}$$

where $\theta_\tau^* = \theta + M_\tau^*$ with M_τ^* , the optimal control process at time τ .

Proof. See [10].

As is standard in stochastic control problems, we first assume that the value function is smooth and derive the associated *Hamilton–Jacobi–Bellman* equation heuristically. This equation turns out to be a non-linear variational inequality with a gradient constraint. Due to inherent degeneracies in the HJB equation, it is convenient to work with weak solutions in the appropriate class. These turn out to be the so-called *viscosity* solutions, presented in subsequent sections. It is in this class that the value function turns out to be the unique solution of the HJB equation.

In the case of a HARA utility function and log-linear laws of motion, the homogeneity of the problem allows us to reduce the dimensionality and obtain a solution to the HJB equation in closed form. This solution is smooth and thus it is a viscosity solution. Therefore this solution coincides

with the value function because of the uniqueness of viscosity solutions to the HJB equation (see Section 3).

There are two possibilities at each instant: preserve the environment or develop. If the decision is to preserve for a time period $[0, h]$, then $\theta_t = \theta$ in $[0, h]$. Hence by the dynamic programming principle:

$$v(x, y, \theta) \geq E \left[\int_0^h e^{-\beta s} U((1-\theta) X_s, \theta Y_s) ds + e^{-\beta h} v(X_h, Y_h, \theta_h) \right], \quad (10)$$

with equality if the optimal decision is to keep $\theta_t = \theta$ in $[0, h]$.

Assuming that Ito's rule applies to $e^{-\beta h} v(X_h, Y_h, \theta_h)$ in (10), dividing both sides by h and taking the limit as $h \rightarrow 0$, we obtain:

$$\beta v \geq \mathcal{L}v + U((1-\theta)x, \theta y), \quad (11)$$

with equality if $\theta_t = \theta$ in $[0, h]$.

Here, \mathcal{L} is the differential operator associated with the diffusion (X_t, Y_t) , that is:

$$\begin{aligned} \mathcal{L}f &= \mu_1(x) f_x + \mu_2(y) f_y + \\ &\quad \frac{1}{2} [\sigma_1^2(x) f_{xx} + \sigma_2^2(y) f_{yy}] + \gamma \sigma_1(x) \sigma_2(y) f_{xy}. \end{aligned} \quad (12)$$

Next, assume that the optimal decision is to develop part of the environment and that $\Delta\theta$ represents the instantaneous share that is developed. Then, $\theta_{0+} = \theta + \Delta\theta$ and (9) implies:

$$v(x, y, \theta) \geq E v(x, y, \theta_{0+}).$$

Clearly, if the optimal decision at the point (x, y, θ) is to move instantaneously to $(x, y, \theta + \Delta\theta)$ with $\Delta\theta > 0$, then a singular transaction of magnitude $\Delta\theta - (\tilde{\theta} - \theta)$ is also optimal for any point $\tilde{\theta} \in (\theta, \theta + \Delta\theta)$. Dividing the above equality by $\Delta\theta > 0$ and sending $\Delta\theta \downarrow 0$ yields at the point (x, y, θ)

$$v_\theta = 0.$$

In general, an instantaneous admissible displacement $\tilde{\Delta}\theta > 0$ from the point (x, y, θ) to $(x, y, \theta + \tilde{\Delta}\theta)$ is suboptimal which, in view of (10) yields

$$v(x, y, \theta) \geq E v(x, y, \theta + \tilde{\Delta}\theta).$$

Dividing by $\tilde{\Delta}\theta$ and passing to the limit as $\tilde{\Delta}\theta \downarrow 0$ gives

$$v_\theta \leq 0.$$

In any case we obtain:

$$v_\theta \leq 0. \quad (13)$$

Writing (11) and (13) in a concise form gives the Hamilton–Jacobi–Bellman (HJB) equation:

$$\min\{\beta v - \mathcal{L}v - U((1 - \theta)x, \theta y), -v_\theta\} = 0 \quad (14)$$

for $x > 0$, $y > 0$ and $0 \leq \theta < 1$.

3. VISCOSITY SOLUTIONS OF THE HJB EQUATION

A classical approach in the theory of stochastic control is to derive information about the value function and the optimal policies via the associated HJB equation. In fact, if it is a priori known that the value function is smooth then it is a classical solution of the HJB equation and, moreover, the optimal policies can be obtained in a feedback form via the first order conditions.

In many applications, like the one we study here, the value function may fail to be smooth and therefore the notion of a solution to the HJB equation needs to be relaxed. It turns out that the “right” class of weak solutions are the so called *viscosity solutions*, introduced by Crandall and Lions [7] for first order equations and by Lions [18] for second order equations⁶.

The main advantage of employing viscosity solutions is that under fairly general assumptions the value function turns out to be the *unique* viscosity solution of the HJB equation. Furthermore, the *stability* properties of viscosity solutions ensures that the value function and optimal policies of more tractable approximations to the original problem, approach the value function and optimal policies of this original problem. This convergence is valid even in cases where the original value function is merely continuous. The theorems presented in this section elaborate on these ideas.

When *state and control constraints* are binding, as for instance the state constraint (6) above, the value function can be characterized as a *constrained* viscosity solution of the HJB equation. The following notion of constrained viscosity solutions was introduced by Soner [21], and Capuzzo-Dolcetta and Lions [5]⁷.

⁶ For a general overview of the theory see [6] or [10].

⁷ See also [13].

Consider a nonlinear second order partial differential equation

$$F(Z, v, Dv, D^2v) = 0 \quad \text{for } Z \in \mathcal{D}, \quad (15)$$

where \mathcal{D} is an open subset of R^n , $F: \bar{\mathcal{D}} \times R \times R^n \times M(n, n) \rightarrow R$, and $M(n, n)$ denotes the set of $n \times n$ matrices. We assume that F is continuous, and if $0 \leq B \in M(n, n)$, then

$$F(Z, t, p, A + B) \leq F(Z, t, p, A)$$

i.e. F is (possibly degenerate) elliptic. Here $Dv \in R^n$ denotes the gradient of v and $D^2v \in M(n, n)$ the matrix of second-order derivatives of v . We denote by $\bar{\mathcal{D}}$ the closure of the open domain \mathcal{D} .

DEFINITION 3.1. A continuous function $v: \bar{\mathcal{D}} \rightarrow R$ is a constrained viscosity solution to (15) if and only if:

(i) For any function $\phi \in \mathcal{C}^2(\bar{\mathcal{D}})$, (i.e. for any ϕ that can be extended as a \mathcal{C}^2 function to an open set containing $\bar{\mathcal{D}}$) and any point $Z_0 \in \bar{\mathcal{D}}$ where $v - \phi$ achieves a maximum,

$$F(Z_0, v(Z_0), D\phi(Z_0), D^2\phi(Z_0)) \leq 0.$$

(ii) For any function $\phi \in \mathcal{C}^2(\bar{\mathcal{D}})$, and any point $Z_0 \in \bar{\mathcal{D}}$ where $v - \phi$ achieves a minimum,

$$F(Z_0, v(Z_0), D\phi(Z_0), D^2\phi(Z_0)) \geq 0.$$

If only condition (i) (or condition (ii)) is satisfied we call v a *viscosity subsolution* (resp. *supersolution*) of equation (15).

Let

$$\mathcal{O} = \{(x, y, \theta) \in R^3 : x > 0, \quad y > 0, \quad 0 < \theta < 1\},$$

and $\bar{\mathcal{O}}$ its closure.

THEOREM 3.1. *The value function v is a constrained viscosity solution of the HJB equation (14) on $\bar{\mathcal{O}}$.*

The proof of this result is essentially as in [22] or as in Theorem VIII.5.1 of [10]. For completeness, we present some of the main steps assuming the existence of an optimal trajectory⁸. The proof will require the following Lemma:

⁸ The technically inclined reader should consult [23] or [17] for general results on the existence of optimal policies in singular stochastic control problems.

To ease the presentation we introduce the generic argument $Z = (x, y, \theta)$ and we will use—by slight abuse of notation— $v(z)$ to denote $v(x, y, \theta)$.

LEMMA 3.1. *Suppose $Z_0 = (x_0, y_0, \theta_0) \in \bar{\mathcal{O}}$ and $\phi \in \mathcal{C}^2(\bar{\mathcal{O}})$ are such that $v \leq \phi$, $\phi(Z_0) = v(Z_0)$ and, $\phi_\theta(Z) > 0$ for each $Z \in \mathcal{B}(Z_0)$, a neighborhood of Z_0 . Consider the optimal trajectory Z_t^* starting at $t=0$ at Z_0 . Let ε be such that $(x_0, y_0, \theta_0 + \varepsilon) \in \mathcal{B}(Z_0)$ and let A be the event that the optimal trajectory has a jump of size at least ε at time $t=0$. Then $P(A) = 0$.*

Proof. By the dynamic programming principle, we have that

$$v(x_0, y_0, \theta_0) = \int_A v(x_0, y_0, \theta_0 + \varepsilon) dP + \int_{\Omega \setminus A} v(x_0, y_0, \theta_0) dP. \tag{16}$$

Hence, $\int_A [v(x_0, y_0, \theta_0 + \varepsilon) - v(x_0, y_0, \theta_0)] dP = 0$. Since $v - \phi$ has a maximum at Z_0 , we have that:

$$\int_A [\phi(x_0, y_0, \theta_0 + \varepsilon) - \phi(x_0, y_0, \theta_0)] dP \geq 0, \tag{17}$$

and also that the function v is subdifferentiable at Z_0 . This together with the fact that v is non-decreasing in θ , yields $\phi(x_0, y_0, \theta_0) \leq \phi(x_0, y_0, \theta_0 + \varepsilon)$. Letting $\varepsilon \rightarrow 0$, and using Fatou's Lemma, inequality (17) yields:

$$\int_A \liminf_{\varepsilon \rightarrow 0} \frac{\phi(x_0, y_0, \theta_0) - \phi(x_0, y_0, \theta_0 + \varepsilon)}{\varepsilon} dP \leq 0. \tag{18}$$

Which, in turn, implies that $P(A) \phi_\theta(Z_0) \leq 0$. Hence, $P(A) = 0$.

Proof of Theorem 3.1. (a) We first show that v is a viscosity subsolution on $\bar{\mathcal{O}}$ i.e. that if $\phi \in C^2(\bar{\mathcal{O}})$ and $v - \phi$ has a maximum at $Z_0 = (x_0, y_0, \theta_0) \in \bar{\mathcal{O}}$ then:

$$\min[\beta v(Z_0) - \mathcal{L}\phi(Z_0) - U(Z_0), -\phi_\theta(Z_0)] \leq 0, \tag{19}$$

where $U(Z_0) = U((1 - \theta_0)x_0, \theta_0 y_0)$. Without loss of generality we assume that $v(Z_0) = \phi(Z_0)$ and $v \leq \phi$ elsewhere. To prove (19) we argue by contradiction. Let us assume that $\phi_\theta(Z_0) < 0$ and that there exists a $\delta > 0$ such that $\beta\phi(Z_0) - \mathcal{L}\phi(Z_0) - U(Z_0) > \delta$. Since ϕ is smooth there must exist a neighborhood of Z_0 , $\mathcal{B}(Z_0)$ such that for $Z \in \mathcal{B}(Z_0)$,

$$\phi_\theta(Z) < 0, \tag{20}$$

and

$$\beta\phi(Z) - \mathcal{L}\phi(Z) - U(Z) > \delta. \tag{21}$$

We write Z_t^* for the optimal trajectory that starts at $t=0$ at Z_0 and let $\tau(\omega) = \inf\{t > 0 : Z_t^*(\omega) \notin \mathcal{B}(Z_0)\}$. By the preceding lemma, $\tau(\omega) > 0$ P -a.s. and inequalities (20) and (21) imply:

$$E \int_0^{\tau(\omega)} e^{-\beta t} \delta \, dt \leq E \int_0^{\tau(\omega)} e^{-\beta t} [\beta \phi(Z_t^*) - \mathcal{L} \phi(Z_t^*) - U(Z_t^*)] \, dt - E \int_0^{\tau(\omega)} e^{-\beta t} \phi_\theta(Z_t^*) \, dm_t^*, \quad (22)$$

where, $M_t^* = \int_0^t dm_s^*$ is the optimal policy.

By Ito's formula:

$$\begin{aligned} E[e^{-\beta \tau(\omega)} \phi(Z_{\tau(\omega)}^*)] &= \phi(Z_0) + E \int_0^{\tau(\omega)} e^{-\beta t} [-\beta \phi(Z_t^*) + \mathcal{L} \phi(Z_t^*)] \, dt \\ &\quad + E \int_0^{\tau(\omega)} e^{-\beta t} \phi_\theta(Z_t^*) \, dm_t^*. \end{aligned} \quad (23)$$

The dynamic programming principle together with the assumptions for the maximum of $v - \phi$ at Z_0 , yields:

$$\phi(Z_0) \leq E \left[\int_0^{\tau(\omega)} e^{-\beta t} U(Z_t^*) \, dt + e^{-\beta \tau(\omega)} \phi(Z_{\tau(\omega)}^*) \right]. \quad (24)$$

Adding (22) to (23) and (24), we obtain:

$$E \int_0^{\tau(\omega)} e^{-\beta t} \delta \, dt \leq 0,$$

which is a contradiction.

(b) We now show that v is a viscosity supersolution of (14) in \mathcal{O} *i.e.* that if $\phi \in C^2(\bar{\mathcal{O}})$ and $v - \phi$ has a minimum at $Z_0 = (X_0, Y_0, \theta_0) \in \mathcal{O}$ then:

$$\min[\beta v(Z_0) - \mathcal{L} \phi(Z_0) - U(Z_0), -\phi_\theta(Z_0)] \geq 0. \quad (25)$$

Again we assume without loss of generality that $v \geq \phi$, $v(Z_0) = \phi(Z_0)$, and we will show that both arguments in (25) are non-negative.

Since choosing $\theta_s = \theta_0$ for $0 \leq s \leq t$ is feasible, the dynamic programming principle implies that:

$$v(Z_0) \geq E \left[\int_0^t e^{-\beta s} U((1 - \theta_0) X_s, \theta_0 Y_s) \, ds + e^{-\beta t} \phi(X_t, Y_t, \theta_0) \right]. \quad (26)$$

Applying Ito's formula to $e^{-\beta t}\phi(X_t, Y_t, \theta_0)$ and using the assumptions on the minimum of $v - \phi$ at Z_0 , yields:

$$E \left\{ \int_0^t e^{-\beta s} (-\beta\phi(X_s, Y_s, \theta_0) + \mathcal{L}\phi(X_s, Y_s, \theta_0) + U((1 - \theta_0) X_s, \theta_0 Y_s)) ds \right\} \leq 0.$$

Dividing by t and letting $t \rightarrow 0$, we get that $\beta v(Z_0) - \mathcal{L}\phi(Z_0) - U(Z_0) \geq 0$ (for details see [22]).

Suppose now that $\theta_0 < 1$ and that $\theta_0 < \theta_0 + M_0 \leq 1$. Again the dynamic programming principle implies that:

$$\phi(x_0, y_0, \theta_0) = v(x_0, y_0, \theta_0) \geq v(x_0, y_0, \theta_0 + M_0) \geq \phi(x_0, y_0, \theta_0 + M_0).$$

Dividing by M_0 and letting $M_0 \rightarrow 0$, since ϕ is smooth, we obtain $\phi_\theta(Z_0) \leq 0$ and thus (25) is established.

Our next task is to characterize the value function as the unique constrained viscosity solution of the HJB equation (14). This characterization will be used in the next section to establish that the solution we obtain for (14), actually coincides with the value function.

Remark 3.1. As it is usually the case in the literature of nonlinear partial differential equations, we present the uniqueness result via a comparison theorem. This is done only to provide more general properties of the solutions of (14), since the uniqueness property follows easily from the theorem below. In fact, suppose that besides the value function v , there exists another viscosity solution to (14) say u which belongs to the same class of solutions that are continuous, concave in x and y and nondecreasing in θ . Since both u and v are viscosity solutions of (14), they are both sub- and supersolutions as the Definition 3.1 requires. In other words, v is a subsolution and u is a supersolution of (14). Then the comparison result stated below yields that $v \leq u$ on $\bar{\mathcal{O}}$. Reversing the arguments yields that $v \geq u$ and the uniqueness follows.

THEOREM 3.2. *Let $u, v: \bar{\mathcal{O}} \rightarrow \mathbb{R}$ be continuous functions which are also concave in x and y and non-increasing with respect to θ . Also assume that v is bounded from below and that, u is a viscosity subsolution of (14) on $\bar{\mathcal{O}}$ and v is a viscosity supersolution of (14) in \mathcal{O} . The $u \leq v$ on $\bar{\mathcal{O}}$.*

Sketch of proof. The proof parallels arguments in [13], so we present only the main steps here. We start by constructing a strictly positive supersolution to the HJB equation (14). To this end, let the function $h: \bar{\mathcal{O}} \rightarrow \mathbb{R}_+$ be given by $h(x, y, \theta) = N(1 + x + y) - C\theta$ with N and C being positive

constants such that $N - C > M > 0$, where M is large enough so that for $x \geq 0, y \geq 0$,

$$U(x, y) \leq M(1 + x + y). \quad (27)$$

Let

$$H(Z, v(Z), Dv(Z), D^2v(Z)) = \min\{\beta v(Z) - \mathcal{L}v(Z) - U(Z), -v_\theta(Z)\}.$$

Then the Hamiltonian $H(Z, h(Z), Dh(Z), D^2h(Z)) > 0$. In fact,

$$\begin{aligned} H(Z, h(Z), Dh(Z), D^2h(Z)) &= \min\{\beta N(1 + x + y) - \beta C\theta - N\mu_1(x) - N\mu_2(y) - U((1 - \theta)x, \theta y), C\} \\ &\geq \min\{\beta N(1 + x + y) - \beta C - N(K_1^\mu x + K_2^\mu y) - M(1 + x + y), C\} \\ &= \min\{(\beta(N - C) - M) + (x + y)[(\beta - \max(K_1^\mu, K_2^\mu))N - M], C\}. \end{aligned} \quad (28)$$

Here, we used (27), together with the fact that $U((1 - \theta)x, \theta y) \leq U(x, y)$ and that K_1^μ (resp. K_2^μ) is the Lipschitz constant associated with X (resp. Y). Using (8), we get:

$$\begin{aligned} H(Z, h(Z), Dh(Z), D^2h(Z)) &\geq \min\{(\beta(N - C) - M) + (x + y)(N - M), C\} \\ &\geq \min\left\{\beta - \frac{M}{N - C}, C\right\} \\ &\geq \min\{\beta - 1, C\} = \bar{M} > 0. \end{aligned} \quad (29)$$

Next let $w^\lambda = \lambda v + (1 - \lambda)h$ for $0 < \lambda < 1$. Notice that w^λ is a viscosity supersolution of $H - (1 - \lambda)\bar{M} = 0$. In fact, let $\psi \in \mathcal{C}^2(\bar{\mathcal{O}})$ and assume that $w^\lambda - \psi$ has a minimum at Z_0 . Then if $\phi = \frac{1}{\lambda}(\psi - (1 - \lambda)h)$, then $v - \phi$ also has a minimum at Z_0 . Since v is a viscosity supersolution of $H = 0$ and inequality (29) holds, we have:

$$\begin{aligned} \lambda H(Z_0, v(Z_0), D\phi(Z_0), D^2\phi(Z_0)) \\ + (1 - \lambda)H(Z_0, h(Z_0), Dh(Z_0), D^2h(Z_0)) \geq (1 - \lambda)\bar{M}. \end{aligned} \quad (30)$$

Since the Hamiltonian $H(z, p, q, A)$ is jointly concave with respect to (p, q, A) , the above inequality yields:

$$H(Z_0, \psi(Z_0), D\psi(Z_0), D^2\psi(Z_0)) \geq (1 - \lambda)\bar{M}, \quad (31)$$

which in turn implies that w^λ is a viscosity supersolution of $H - (1 - \lambda) \bar{M} = 0$. Applying the comparison results of Theorem VI.5 of [13] to u and w^λ , we get

$$u \leq w^\lambda \text{ on } \bar{O}.$$

Letting $\lambda \rightarrow 1$ we obtain the result.

Remark 3.2. For completeness, we note that the above comparison results hold if we relax the continuity assumption and allow for u and v to be, respectively, upper- and lower-semicontinuous functions.

Moreover, one could also drop the concavity assumption on v by allowing v to be of sublinear growth⁹ and bounded from below. (We refer the technically oriented reader to [13].)

4. THE CASE OF AN ADDITIVE HARA UTILITY FUNCTION

In this section we assume that the dynamics of X_t and Y_t are linear i.e. that equations (1) and (2) are specialized to:

$$dX_t = \mu X_t dt + \sigma X_t dW_t^1 \tag{32}$$

and

$$dY_t = \mu_1 Y_t dt + \sigma_1 Y_t dW_t^2. \tag{33}$$

For definiteness we will assume that $\sigma > 0$ and $\sigma_1 > 0$. This will allow us to state that when $\gamma > 0$, ($\gamma < 0$) the innovations to X_t and Y_t are positively (resp. negatively) correlated. We also specialize our utility function¹⁰ to

$$U((1 - \theta) x, \theta y) = \frac{[(1 - \theta) x + \theta y]^p}{p}, \quad 0 < p < 1. \tag{32}$$

The homogeneity of the value function of degree p follows from the homogeneity of degree p of U and the linearity of the state equations. As we observed in Remark 2.4, the problem is trivial whenever $x = 0$. The homogeneity enables us to reduce the dimensionality of the original

⁹ A function $f(x, y)$ has sublinear growth if, for some constant N , $f(x, y) \leq N(1 + x + y)$, for x and y large.

¹⁰ The limit case $p = 0$, corresponding to logarithmic utility function can also be treated. However, since we assumed earlier that $U(0, 0)$ is finite, we would need to present further arguments.

problem from three to two whenever $x \neq 0$. If $x \neq 0$, for $z = y/x$, homogeneity yields:

$$v(x, y, \theta) = x^p v(1, z, \theta). \quad (33)$$

Define:

$$u(z, \theta) = v(1, z, \theta) = x^{-p} v(x, xz, \theta). \quad (34)$$

Then, at least formally, u solves:

$$\min\{\bar{\beta}u - \frac{1}{2}k_1^2 z^2 u_{zz} - k_2 z u_z - U(1 - \theta + \theta z), -u_\theta\} = 0, \quad (35)$$

where

$$\begin{aligned} \bar{\beta} &= \beta + p(1-p)\sigma^2/2 - \mu p \\ k_1^2 &= \sigma^2 - 2\gamma\sigma_1\sigma + \sigma_1^2 \\ k_2 &= \sigma^2(1-p) + \gamma\sigma_1\sigma(p-1) - \mu + \mu_1. \end{aligned} \quad (36)$$

We will assume that β is large enough, so that

$$\bar{\beta} > \max(k_1^2, k_2). \quad (37)$$

This growth condition will be subsequently used in the construction of the solution and the specification of the optimal exercise boundary.

We can show that in fact, u is the value function associated with the reduced problem:

$$u(z, \theta) = \sup_{\hat{\mathcal{A}}_\theta} E \int_0^\infty e^{-\bar{\beta}t} ((1 - \theta_t) + \theta_t z_t)^p dt, \quad (38)$$

subject to:

$$\begin{aligned} dz_t &= k_2 z_t dt + k_1 z_t dW_t^1, \\ d\theta_t &= dM_t, \end{aligned} \quad (39)$$

where W_t^1 is a Brownian motion and $\hat{\mathcal{A}}_\theta$ is defined along the same lines as \mathcal{A}_θ .

Remark 4.1. The drift in the stochastic differential equation (39) is not equal to the drift of the variable $z = \frac{y}{x}$ under the law of motion given by equations (30) and (31) (although the volatility is). This change in drift and of discount factor is necessary because the objective function in (38) does not account for the level of the variable X .

Remark 4.2. As pointed out in Remarks 2.3 and 2.5, u satisfies:

$$u(z, 1) = E \int_0^\infty e^{-\bar{\beta}t} z_t^p dt, \tag{40}$$

$$u(0, \theta) = \frac{(1 - \theta)^p}{\bar{\beta}}. \tag{41}$$

Also, $u(\cdot, \theta)$ is concave and strictly increasing and $u(z, \cdot)$ is concave and nonincreasing.

4.1. Optimal Exercise Boundary

Since the HJB equation is associated with a singular control problem, it is expected that the state space is partitioned in two regions, say region *I* and region *II*. In region *I*, $u_\theta < 0$ and hence, $\bar{\beta}u - (1/2)k_1^2 z^2 u_{zz} - k_2 z u_z - U(1 - \theta + \theta z) = 0$, while in region *II*, $u_\theta = 0$. If $(z, \theta) \in I$, then no action is taken. If $(z, \theta) \in II$ the optimal policy is singular. The next task is to characterize these two areas. We first assume that the common boundary of the two regions is described by a curve $z = \phi(\theta)$, and we provide an explicit description of this boundary. Later we show that, in fact, this boundary characterizes an optimal policy, in the sense that if $(z, \theta) \in II = \{(z, \theta) : z > \phi(\theta)\}$, an optimal policy is to jump instantaneously to the θ_0 that satisfies $z = \phi(\theta_0)$, while, if $(z, \theta) \in I = \{(z, \theta) : z \leq \phi(\theta)\}$ then no action should be taken.

Remark 4.3. The concavity of u implies that u_θ is monotone. Hence, if $(z, \theta) \in I$ and $\theta' > \theta$, then $(z, \theta') \in I$. Also from Remark 2.5, we know that $(0, \theta) \notin II$.

Remark 4.4. The function ϕ must be monotone. In fact if $z < z'$, for any θ_0 ,

$$\theta_{0+}(z, \theta_0) \leq \theta_{0+}(z', \theta_0). \tag{42}$$

To see this, first suppose that $z' > z \geq 1$. Let T be the first time that $z_t(z', \theta_0) \leq z$. Notice that, by the dynamic programming principle, $\theta_{T+}(z', \theta_0) \geq \theta_{0+}(z, \theta_0)$. For each $0 < t \leq T$, let

$$\bar{\theta}_t(z', \theta_0) = \max\{\theta_t(z', \theta_0); \theta_{0+}(z, \theta_0)\},$$

and for $t \geq T$ let $\bar{\theta}_t = \theta_t$. If (42) does not hold, since $z \geq 1$, the $\bar{\theta}_t$ path dominates the candidate optimum. If $1 \geq z' > z$, an analogous reasoning can be made by choosing T to be the first time at which the path starting at z crosses the level z' and choosing for $t \leq T$,

$$\bar{\theta}_t(z, \theta_0) = \min\{\theta_t(z, \theta_0); \theta_{0+}(z', \theta_0)\}.$$

We proceed by assuming that the function ϕ has an inverse $\psi: [\phi(0), \phi(1)] \rightarrow [0, 1]$ and we construct the solution u of the Variational Inequality (35) by imposing the appropriate smooth pasting condition

$$u_{\theta z}(\phi(\theta), \theta) = 0 \quad (43)$$

across the optimal exercise boundary.

The solution $u: [0, +\infty) \times [0, 1] \rightarrow R^+$ is given by

$$u(z, \theta) = \begin{cases} \hat{u}(z, \theta) & z \in [0, \phi(\theta)) \\ \hat{u}(z, \psi(z)) & z \in [\phi(\theta), \phi(1)) \\ u(z, 1) & z \in [\phi(1), +\infty) \end{cases} \quad (44)$$

where $u(z, 1)$ is given in (40) and \hat{u} solves

$$\beta \hat{u} - \frac{1}{2} k_1^2 z^2 \hat{u}_{zz} - k_2 z \hat{u}_z - \frac{[1 - \theta + \theta z]^p}{p} = 0 \quad z \in [0, \phi(\theta)) \quad (45)$$

with the boundary condition

$$\hat{u}_\theta(\phi(\theta), \theta) = 0. \quad (46)$$

As it is explained in Appendix A, (43)–(46) imply that

$$\begin{aligned} \hat{u}(z, \theta) = A(\theta) z^{\rho_1} + \frac{2}{(\rho_1 - \rho_2) \sigma^2} & \left[z^{\rho_2} \int_0^z \frac{(1 - \theta + \theta x)^p}{p x^{\rho_2 + 1}} dx \right. \\ & \left. + z^{\rho_1} \int_z^{+\infty} \frac{(1 - \theta + \theta x)^p}{p x^{\rho_1 + 1}} dx \right] \end{aligned} \quad (47)$$

where

$$A(\theta) = \frac{2}{(\rho_1 - \rho_2) \sigma^2} \left[\int_\theta^1 \int_{\phi(\theta)}^{+\infty} \frac{(x - 1)(1 - \theta + \theta x)^{p-1}}{x^{\rho_1 + 1}} dx d\theta \right] \quad (48)$$

and $\rho_1 > 0 > \rho_2$ are the roots of the quadratic equation

$$\frac{1}{2} k_1^2 \rho^2 + (k_2 - k_1^2/2) \rho - \bar{\beta} = 0. \quad (49)$$

In order to guarantee that the solution u —together with the optimal exercise boundary—is well defined, we need the roots ρ_1, ρ_2 to satisfy $\rho_2 < p - 1$ and $\rho_1 > 1$ ¹¹. In Appendix A we show that these properties follow from the assumptions on the discount factor $\bar{\beta}$ in (37).

¹¹ See for example [15] (Section 6).

A simple computation using (43) and (46) yields that the optimal exercise boundary must satisfy:

$$\phi(\theta) = \sup \left\{ \ell : \int_0^\ell \frac{(x-1)[1-\theta+\theta x]^{p-1}}{x^{\rho_2+1}} dx < 0 \right\}. \tag{50}$$

This is easily seen since the integrand on (50) is monotone, negative for small values of ℓ , and the integral must vanish at $\phi(\theta)$.

4.2. *Qualitative Properties of the Optimal Exercise Boundary*

We present here some properties of the optimal exercise boundary:

PROPERTY 4.1.

$$\phi(0) = \frac{1 + |\rho_2|}{|\rho_2|} > 1.$$

Proof. From equation (50), we know that,

$$\int_0^{\phi(0)} x^{-\rho_2} dx = \int_0^{\phi(0)} x^{-(\rho_2+1)} dx.$$

The above condition holds for $\phi(0) = 0$, or for $\phi(0) = \frac{1+|\rho_2|}{|\rho_2|} > 1$. We claim that $\phi(0) \neq 0$. In fact, recall from Remark 4.3 that $\phi(\theta) > 0$, if $\theta > 0$. If $\phi(0) = 0$, let $z < 1$ be in the range of ϕ , say $z = \phi(\theta_1)$. Then:

$$\int_0^{\phi(\theta_1)} \frac{(x-1)[1-\theta+\theta x]^{p-1}}{x^{\rho_2+1}} dx < 0,$$

which gives a contradiction. Therefore, $\phi(0) = (1 + |\rho_2|)/|\rho_2|$.

Property 4.1 guarantees that there exists a value $\underline{z}(p) > 1$ such that if $z \leq \underline{z}(p)$, the optimal policy is to stay put, even if $\theta = 0$.

PROPERTY 4.2. *Equation (50) defines an increasing function $\phi: [0, 1] \mapsto \mathbb{R}$.*

Proof. Define for each $0 \leq \theta \leq 1$,

$$F(k, \theta) = \int_0^k \frac{(x-1)[1-\theta+\theta x]^{p-1}}{x^{\rho_2+1}} dx. \tag{51}$$

The following properties are immediate:

(i) $F(0, \theta) = 0,$

(ii) $F_k(k, \theta) < 0,$ if $k < 1,$ $F_k(k, \theta) > 0,$ if $k > 1,$ and $\lim_{k \rightarrow \infty} F_k(k, \theta) = \infty.$

Hence, for each $0 \leq \theta \leq 1$ there exists exactly one $\phi(\theta) > 1$ such that $F(\phi(\theta), \theta) = 0,$ and furthermore, this $\phi(\theta)$ satisfies (50). Further, $F(1, \theta_1) < F(1, \theta_2)$ if, $\theta_1 < \theta_2.$ Hence:

$$\int_1^{\phi(\theta_1)} \frac{(x-1)[1-\theta+\theta x]^{p-1}}{x^{\rho_2+1}} dx - \int_1^{\phi(\theta_2)} \frac{(x-1)[1-\theta+\theta x]^{p-1}}{x^{\rho_2+1}} dx \geq 0,$$

that is, $\phi(\theta_1) < \phi(\theta_2).$

PROPERTY 4.3. *The function $\phi(\theta)$ satisfies:*

$$\phi(1) = \frac{p + |\rho_2|}{p + |\rho_2| - 1}.$$

Proof. Since $p + |\rho_2| - 1 > 0$ (see Appendix A) we know that

$$\int_0^{\phi(1)} x^{p+|\rho_2|-1} dx = \int_0^{\phi(1)} x^{p+|\rho_2|-2} dx.$$

Hence:

$$\phi(1) = \frac{p + |\rho_2|}{p + |\rho_2| - 1}.$$

Remark 4.5. As stated above, the growth conditions on the roots of the characteristic quadratic (49), namely $\rho_2 < p - 1$ and $\rho_1 > 1$ are needed in order to guarantee that the solution in (47) is well posed. In addition it is worth observing at this point that if the condition $\rho_2 < p - 1$ is violated, then (50) will yield $\phi(1) = +\infty,$ i.e. for this range of parameters, full conversion is never optimal. We do not examine these cases herein.

Remark 4.6. If $z \geq \bar{z}(p) \equiv \frac{p+|\rho_2|}{p+|\rho_2|-1},$ then the optimal policy leads to full conversion. Observe that as $p \rightarrow 1$ (linear utility) then $\phi(0) \rightarrow \phi(1) \rightarrow (1 + |\rho_2|)/|\rho_2|,$ which together with the monotonicity of ϕ implies that the optimal boundary converges as $(p \rightarrow 1)$ to the vertical line $\phi(\theta) = \frac{1+|\rho_2|}{|\rho_2|}.$ That is, the optimal policy converges to: “do nothing if $z < \phi(\theta)$ and convert everything if $z \geq \phi(\theta).$ ”

The next Proposition studies the behavior of ϕ when the parameters of the problem change. Notice that if a change in a parameter value makes ϕ increase (for every value of θ), this change in parameter value implies a more conservative policy. To help analyze the effect of parameters on the solution to the problem, we will slightly abuse the notation and write the optimal exercise boundary curve as $\phi(\beta, \mu, \mu_1, \gamma, \sigma, \sigma_1, \theta)$.

PROPERTY 4.4. *The function ϕ satisfies:*

- (i) $\frac{\partial \phi}{\partial \beta} < 0$,
- (ii) $\frac{\partial \phi}{\partial \mu} > 0$,
- (iii) $\frac{\partial \phi}{\partial \mu_1} < 0$,
- (iv) $\frac{\partial \phi}{\partial \gamma} < 0$,
- (v) If $\gamma \leq 0$, $(\frac{\partial \phi}{\partial \sigma}) > 0$,
- (vi) If $\gamma \leq 0$, $(\frac{\partial \phi}{\partial \sigma_1}) > 0$.

Proof. Recall that ϕ satisfies:

$$F(\phi, \theta, \beta, \mu, \mu_1, \gamma, \sigma, \sigma_1) = 0, \tag{52}$$

where F is the obvious analogue to the function defined in (51) above, obtained by making the dependence on the parameters $\beta, \mu, \mu_1, \gamma, \sigma$, and σ_1 explicit. Hence, if $v \in \{\beta, \mu, \mu_1, \gamma, \sigma, \sigma_1\}$, we know from the implicit function theorem that:

$$\frac{\partial \phi}{\partial v} = - \frac{F_v(\phi(\beta, \mu, \mu_1, \gamma, \sigma, \sigma_1, \theta), \theta, \beta, \mu, \mu_1, \gamma, \sigma, \sigma_1)}{F_k(\phi(\beta, \mu, \mu_1, \gamma, \sigma, \sigma_1, \theta), \theta, \beta, \mu, \mu_1, \gamma, \sigma, \sigma_1)}. \tag{53}$$

Further:

$$F_v = - \frac{\partial \rho_2}{\partial v} \int_0^{\phi(\beta, \mu, \mu_1, \gamma, \sigma, \sigma_1, \theta)} \frac{(x-1)[1-\theta+\theta x]^{p-1}}{x^{p_2+1}} \ln x \, dx. \tag{54}$$

Notice that the integrand in equation (54) is always positive and thus

$$\text{sign}(F_v) = - \text{sign} \left(\frac{\partial \rho_2}{\partial v} \right).$$

Further, from Properties 4.1 and 4.2 we know that $\phi(\beta, \mu, \mu_1, \gamma, \sigma, \sigma_1, \theta) > 1$ and, as a consequence, that $F_k(\phi(\beta, \mu, \mu_1, \gamma, \sigma, \sigma_1, \theta), \theta, \beta, \mu, \mu_1, \gamma, \sigma, \sigma_1) > 0$. Hence

$$\text{sign} \left(\frac{\partial \phi}{\partial v} \right) = \text{sign} \left(\frac{\partial \rho_2}{\partial v} \right).$$

Hence it suffices to establish the sign of $\frac{\partial \rho_2}{\partial v}$. In Appendix A, we show that ρ_2 is the negative (in fact less than $p-1$) root of the quadratic equation (49), whose coefficients, $\bar{\beta}$, k_1^2 and k_2 are given by (36) in terms of the vector of parameters $(\beta, \mu, \mu_1, \gamma, \sigma, \sigma_1)$. Hence, if we write (49) as follows:

$$G(\beta, \mu, \mu_1, \gamma, \sigma, \sigma_1, \rho) \equiv \frac{1}{2}k_1^2\rho^2 + (k_2 - k_1^2/2)\rho - \bar{\beta}, \quad (55)$$

then implicit differentiation again yields, for each $v \in \{\beta, \mu, \mu_1, \gamma, \sigma, \sigma_1\}$,

$$\frac{\partial \rho_2}{\partial v} = -\left(\frac{\partial G}{\partial \rho}\right)^{-1} \frac{\partial G}{\partial v}. \quad (56)$$

Since G is convex in ρ and ρ_2 is the smallest of the two roots of equation (55), $\frac{\partial G}{\partial \rho}(\beta, \mu, \mu_1, \gamma, \sigma, \sigma_1, \rho_2) < 0$. Hence

$$\text{sign}\left(\frac{\partial \phi}{\partial v}\right) = \text{sign}\left(\frac{\partial G}{\partial v}\right),$$

and by using (36) we can easily establish (i). To show that (ii) holds, observe that:

$$\frac{\partial G}{\partial \mu}(\beta, \mu, \mu_1, \gamma, \sigma, \sigma_1, \rho_2) = p - \rho_2 > 0. \quad (57)$$

Similarly:

$$\frac{\partial G}{\partial \mu_1}(\beta, \mu, \mu_1, \gamma, \sigma, \sigma_1, \rho_2) = \rho_2 < 0, \quad (58)$$

and therefore (iii) holds. Also:

$$\begin{aligned} \frac{\partial G}{\partial \gamma}(\beta, \mu, \mu_1, \gamma, \sigma, \sigma_1, \rho_2) &= -\rho^2\sigma_1\sigma + \rho\sigma_1\sigma + \rho\sigma_1\sigma(p-1) \\ &= \sigma_1\sigma(-\rho_2^2 + \rho_2p) < 0. \end{aligned} \quad (59)$$

Hence (iv) holds. Further,

$$\begin{aligned} \frac{\partial G}{\partial \sigma}(\beta, \mu, \mu_1, \gamma, \sigma, \sigma_1, \rho_2) &= \sigma(\rho_2^2 + \rho_2 - 2\rho_2p - p(1-p)) + \gamma\sigma_1\rho_2(p - \rho_2) \\ &= \sigma(\rho_2 - p)(\rho_2 - (p-1)) + \gamma\sigma_1\rho_2(p - \rho_2). \end{aligned} \quad (60)$$

Since $\rho_2 < p-1$, the first term of the right hand side of (62) is always positive. The second term is non-negative if $\gamma \leq 0$. Hence (v) is established.

Finally:

$$\frac{\partial G}{\partial \sigma_1}(\beta, \mu, \mu_1, \gamma, \sigma, \sigma_1, \rho_2) = \sigma_1(\rho_2^2 - \rho_2) + \gamma \sigma \rho_2(p - \rho_2).$$

Again, the first term in the right hand side is positive and the second is non-negative if $\gamma \leq 0$, and this establishes (vi).

The last Proposition states that an increase in the expected rate of return of X leads to a more conservative policy (an increase in ϕ). An increase in the discount factor or in the expected rate of return of Y in turn, leads to a less conservative policy. An increase in the correlation between innovations in X and in Y , lowers the option value of remaining at a given θ and hence leads to a less conservative policy. The comparative statics with respect to the volatilities σ or σ_1 are more complicated. Unless the correlation coefficient γ is negative, an increase in the log volatility of X , can have an arbitrary effect on the variance of $\log X - \log Y$, and the variance of this difference is what determines the effect on the map ϕ . In fact, if we increase k_1^2 , the variance of the innovations of $\log X - \log Y$, and change the other parameters in $\{\beta, \mu, \mu_1, \gamma, \sigma, \sigma_1\}$ in order to keep $\bar{\beta}$ and k_2 unchanged, then one can easily verify that ϕ goes up.

Remark 4.7. The comparative statics with respect to the risk-aversion parameter p are more complicated. A benchmark can be obtained by examining the case where the two flow processes are identical and independent martingales i.e. $\sigma = \sigma_1$, $\gamma = 0$ and $\mu = \mu_1 = 0$. Since ρ_2 is the negative root of equation (49) we have, using the formulas for the coefficients given in (36) above, that:

$$\frac{\partial \rho_2}{\partial p} = \frac{(p - \rho_2 - 1/2)}{p - 2\rho_2}.$$

Since $\rho_2 < p - 1$, we have that:

$$0 < \frac{\partial \rho_2}{\partial p} < 1.$$

Using Property 4.1 we have that

$$\frac{\partial \phi(0)}{\partial p} > 0,$$

and, similarly, using Property 4.3,

$$\frac{\partial \phi(1)}{\partial p} < 0.$$

That is, if we assume more risk aversion (p goes down), the optimal policy is *less* conservative for low values of θ and *more* conservative for high values of θ . To understand the economic intuition behind the first part of this result, imagine that X_t and Y_t satisfy the same diffusion, that $\gamma = 0$, that $\theta_0 = 0$ and that $X_0 = Y_0$. In this case, if we ignored the irreversibility, the usual diversification result would lead to an optimal choice of $\theta = 1/2$. The cost of departing from the policy of $\theta = 1/2$ is larger if the utility function exhibits more risk aversion (a smaller p) and hence risk-aversion tends to decrease the threshold $\underline{z}(p)$. As risk-aversion increases, the same tendency towards diversification increases the threshold $\bar{z}(p)$ at which full conversion occurs.

Write $\phi_p(\theta)$ for the exercise boundary when the risk-aversion parameter is p and consider the decomposition:

$$\phi_p(\theta) = \phi_1(\theta) + (\phi_p(\theta) - \phi_1(\theta)). \quad (61)$$

The first term in the right hand side of (61) is the *quasi-option effect*. The second term results from risk-aversion and is negative for low θ 's and positive for high θ 's.

5. A VERIFICATION RESULT AND CONCLUSION

We first state the main theorem which characterizes the value function of the original problem (7) when the utility function is of the form given by (4). The proof of the theorem is presented in Appendix B.

THEOREM 5.1. *Let $a(x, y, \theta) = x^p u(\frac{y}{x}, \theta)$ where u is the value function of the reduced problem given by (44). Moreover, consider the curve given by (50). Then:*

- (i) $a \equiv v$ is the value function of (7) and,
- (ii) the optimal policy M_t^* is such that if $\theta_t^* = \theta + M_t^*$ and $\psi = \phi^{-1}$ then

$$\theta_t^* = \min\{\theta, \sup_{s \leq t} \psi(Y_s/X_s)\}.$$

In this paper we considered a problem of developing portions of an original environment when the development is irreversible and the benefits of preservation and of the development project are uncertain. In the general case, we were able to establish that the value function associated with the problem is the unique viscosity solution of a Hamilton–Jacobi–Bellman equation. For the case where the utility function exhibits constant relative risk aversion between zero and one, and the benefits satisfy a log-linear diffusion, we were able to characterize the solution much more explicitly. In this case, we showed that the optimal policy was given by an *exercise boundary*. Given the ratio of the benefits at any time t , z_t , if the fraction converted up to that point in time is below that given by a function $\psi(z_t)$, then the optimal policy prescribes that the fraction converted be brought to the level $\psi(z_t)$. On the other hand, since conversion is irreversible, if the fraction converted exceeds $\psi(z_t)$, then no action should be taken.

This exercise boundary was shown to have several properties. We showed that there exists a $\underline{z}(p) > 1$ such that if $z \leq \underline{z}(p)$, then even if nothing has been converted, no action should be taken. This means that conversion should not start until the benefits from conversion exceed the benefits from preservation by a certain margin. This, of course, reflects the fact that development is irreversible. The function ψ is monotone, i.e. larger current relative benefits to conversion lead to conversion of a higher fraction of the environment. We also established several comparison results relating changes in parameters of the problem to changes in the optimal policy.

The problem we discussed is of course similar to that of converting one asset into another. In complete markets set-ups such as that of exercising an American option in the classical Black–Scholes framework, the optimal policy leads to full conversion at a point in time. However, in the presence of incomplete markets, though it is usually assumed that agents choose a point in time to convert the full amount, there is no reason why this would be the case in an optimal policy, when we permit fractional conversion. The framework of this paper should prove useful to investigate this type of question.

APPENDIX A

Derivation Concerning the Optimal Exercise Boundary

In this Appendix, we prove two claims made in Subsection 4.1 where we established the properties of the optimal exercise boundary.

Since $\rho_2 < 0$ any bounded solution of (45) is given by

$$\hat{u}(z, \theta) = A(\theta) z^{\rho_1} + H(z, \theta) \quad (\text{A.1})$$

where

$$H(z, \theta) = \frac{2}{(\rho_1 - \rho_2) \sigma^2} \left[z^{\rho_2} \int_0^z \frac{(1 - \theta + \theta x)^p}{p x^{\rho_2 + 1}} dx + z^{\rho_1} \int_z^{+\infty} \frac{(1 - \theta + \theta x)^p}{p x^{\rho_1 + 1}} dx \right]. \quad (\text{A.2})$$

The definition of the value function yields $u(z, 1) = H(z, 1)$ and hence

$$A(1) = 0.$$

Imposing (43) and (46) implies

$$A'(\theta) \phi(\theta)^{\rho_1} + H_{\theta}(\phi(\theta), \theta) = 0 \quad (\text{A.3})$$

and

$$\rho_1 A'(\theta) \phi(\theta)^{\rho_1 - 1} + H_{\theta z}(\phi(\theta), \theta) = 0. \quad (\text{A.4})$$

Finally the last two relations give (48) and (50).

We now show that (48) and the restrictions on the roots $\rho_2 < p - 1$ and $\rho_1 > 1$ hold. In fact, evaluating (49) at the point $p - 1$ we get

$$\begin{aligned} C_1 &= \frac{1}{2} k_1^2 (p - 1)^2 + \left(k_2 - \frac{k_1^2}{2} \right) (p - 1) - \bar{\beta} \\ &= \frac{k_1^2}{2} (p - 1)(p - 2) + k_2(p - 1) - \bar{\beta} \end{aligned} \quad (\text{A.5})$$

and, at the point 1,

$$C_2 = k_2 - \bar{\beta}. \quad (\text{A.6})$$

To establish that $\rho_2 < p - 1$ and $\rho_1 > 1$, it suffices to show that $C_1 < 0$ and $C_2 < 0$. Notice that for $0 < p < 1$, $C_1 < k_1^2 - \bar{\beta}$. Therefore, we need to have $\max\{k_1^2 - \bar{\beta}, k_2 - \bar{\beta}\} \leq 0$ which holds given (37).

APPENDIX B

Proof of Verification Theorem

Proof of (i). First we observe that the function $a(x, y, \theta)$ is a classical solution in the interior of the domain $\bar{\mathcal{C}}$. Moreover, by using the definition of constrained viscosity solutions, we can show that a is a constrained viscosity solution on $\bar{\mathcal{C}}$. Moreover, a is concave in x and y and nondecreasing on θ (the latter argument can be proven directly using the form of u). Using the fact that the value function is the unique (constrained) viscosity solution in the above class (see Theorem 3.2), we easily conclude that a coincides with the value function.

Proof of (ii). For ease of presentation, we present the optimal policy for the reduced one-dimensional problem. Once we establish the optimality, the verification of the original two dimensional problem is routine (e.g. [9]). Although the verification arguments follow along the lines of [15], we present them here for completeness.

A straightforward computation using (44), (47), (48) and (50) yields

$$u_\theta(z, \theta) \leq 0. \quad (\text{B.1})$$

Applying Ito's formula to $e^{-\bar{\beta}t}u(z_t, \theta_t)$, where t is arbitrary, yields:

$$\begin{aligned} e^{-\bar{\beta}t}u(z_t, \theta_t) &= u(z, \theta) + \int_0^t e^{-\bar{\beta}s} (-\beta u(z_s, \theta_s) + \hat{\mathcal{L}}u(z_s, \theta_s)) ds + \\ &+ \int_0^t e^{-\bar{\beta}s} u_\theta(z_s, \theta_s) \tilde{m}_s ds \\ &+ e^{-\bar{\beta}t} \sum_i [u(z_{t_i}, \theta_{t_i}^+) - u(z_{t_i}, \theta_{t_i})], \end{aligned} \quad (\text{B.2})$$

where, \tilde{m}_t represents the density of the absolutely continuous part of the control process M_t , and the last summation is over all times $t_i < t$ at which M_t jumps. Taking expectations in (B.2) and using (44) and (B.1) yields:

$$u(z, \theta) \geq E \left[\int_0^t e^{-\bar{\beta}s} \frac{[1 - \theta_s + \theta_s z_s]^p}{p} ds + e^{-\bar{\beta}t} u(z_t, \theta_t) \right].$$

Letting $t \rightarrow \infty$ and using the monotone convergence theorem as well as the fact that $\lim_{t \rightarrow \infty} e^{-\bar{\beta}t} u(z_t, \theta_t) = 0$ (this follows from the form of u and the growth conditions) we get:

$$u(z, \theta) \geq E \int_0^\infty e^{-\bar{\beta}s} \frac{[1 - \theta_s + \theta_s z_s]^p}{p} ds.$$

Next assume that we use the policy M_t^* such that:

$$\theta_t^* = \min\{\theta, \sup_{s \leq t} \psi(z_s)\}. \quad (\text{B.3})$$

In other words, M_t^* represents a vertical reflection on the curve $\theta = \psi(z)$. Then the optimality of (B.3) follows from the fact that θ_t never enters the interior of the region $\{\theta: \theta \geq \psi(z)\}$ for $t \geq 0$, and that $u_\theta = 0$ whenever θ_t increases.

REFERENCES

1. K. J. Arrow, Optimal capital policy with irreversible investment, in "Value, Capital and Growth, Essays in Honor of Sir John Hicks" (J. N. Wolfe, Ed.), Edinburgh Academic Press, Edinburgh, Scotland.
2. K. J. Arrow and A. C. Fisher, Environmental preservation, uncertainty and irreversibility, *Quarterly J. Econ.* **88** (1974), 312–319.
3. G. Bertola, Irreversible investment, Princeton University, 1989, unpublished.
4. F. Black and M. Scholes, The pricing of options and corporate liabilities, *J. Polit. Econ.* **81** (1973), 637–659.
5. I. Capuzzo-Dolcetta and P.-L. Lions, Viscosity solutions of Hamilton–Jacobi equations, *Trans. Amer. Math. Soc.*, to appear.
6. M. Crandall, H. Ishii, and P.-L. Lions, User's guide to viscosity solutions of second order partial differential equations, *Bull. Amer. Math. Soc.* **27** (1992), 1–67.
7. M. Crandall and P.-L. Lions, Viscosity solutions of Hamilton–Jacobi equations, *Trans. Amer. Math. Soc.* **277** (1983), 1–42.
8. A. Dixit, Irreversible investment with uncertainty and scale economies, *J. Econ. Dyn. Control* **19** (1995) 327–350.
9. D. Duffie, W. Fleming, H. M. Soner, and T. Zariphopoulou, Hedging in incomplete markets with HARA utility, *J. Econ. Dyn. Control*, **21** (1997), 753–782.
10. W. Fleming and H. M. Soner, "Controlled Markov Processes and Viscosity Solutions," Springer-Verlag, 1993.
11. X. Freixas and J.-J. Laffont, The irreversibility effect, in "Bayesian Models in Economic Theory" (M. Boyer and R. Khilstrom, Eds.), North Holland, 1984.
12. C. Henry, Investment decisions under uncertainty: the irreversibility effect, *Amer. Econ. Rev.* **64** (1974), 1006–1012.
13. H. Ishii and P.-L. Lions, Viscosity solutions of fully nonlinear second-order elliptic partial differential equations, *J. Diff. Eq.* **83** (1990), 26–78.
14. I. Karatzas and S. Shreve, "Brownian Motion and Stochastic Calculus," Springer-Verlag, 1996.
15. T. O. Kobila, A class of solvable stochastic investment problems involving singular controls, *Stochastics Stochastics Rep.* **43** (1993), 29–63.

16. C. D. Kolstad, Fundamental irreversibilities in stock externalities, *J. Public Econ.* **60** (1996), 221–233.
17. T. Kurtz, A control formulation for constrained Markov processes, mathematics of random media, in “Lectures in Applied Mathematics,” Vol. 27, American Mathematical Society, Providence (1984), 139–150.
18. P. L. Lions, Optimal control of diffusion processes and Hamilton–Jacobi–Bellman equations, Part 1: The Dynamic Programming Principle and Application and, Part 2: Viscosity Solutions and Uniqueness, *Commun. PDE's* **8** (1983), 1101–1174 and 1229–1276.
19. F. R. McDonald and D. R. Siegal, The value of waiting to invest, *Quarterly J. Econ.* **101** (1986), 707–728.
20. R. S. Pyndick, Irreversible investments, capacity choice and the value of the firm, *Amer. Econ. Rev.* **78** (1988), 969–985.
21. H. Soner, Optimal control with state space constraints, *SIAM J. Control Optim.* **26** (1986), 552–562 (Part I) and 1110–1122 (Part II).
22. T. Zariphopoulou, Investment-consumption models with constraints, *SIAM J. Control Optim.* **32** (1994), 59–84.
23. H. Zhu, “Characterization of Variational Inequalities in Singular Control,” Ph.D. thesis, Brown University, 1986.