Discussion Paper No. 916

(s,S) EQUILIBRIA IN STOCHASTIC GAMES WITH AN APPLICATION TO PRODUCT INNOVATIONS*

by

Prajit K. Dutta
Dept. of Economics, Columbia University
New York, NY 10027

and

Aldo Rustichini
Dept. of Economics, Northwestern University
Evanston, IL 60208

June 1990
Revised November 1990

---

* We have benefitted from the comments of seminar participants at Chicago, Columbia, McGill and Western Ontario, and especially Ariel Rubinstein, Bentley Mcleod, Joseph Greenberg, James Friedman, Jose Scheinkman, Ricardo Caballero, Andrew Caplin and Robert E. Lucas. A preliminary version of this paper was presented at the International Game Theory Conference at Ohio State University, July 1989. The current version was prepared when the first author was visiting the University of Rochester.

Abstract We study a class of two–player continuous time stochastic games in which agents can make (costly) discrete or discontinuous changes in the variables that affect their payoffs. It is shown that in these games there are Markov perfect equilibria of the two–sided (s,S) rule type. In such equilibria at a critical low state (resp. high state) player 1 (resp. 2) effects a discrete change in the environment. In some of these equilibria either or both players may be passive. On account of the presence of fixed costs (to discrete changes) the payoffs are non–convex and hence standard existence arguments fail. We prove that the best response map satisfies a surprisingly strong monotonicity condition and use this to establish the existence of Markov perfect equilibria. The first–best solution is also a two–sided (s,S) rule but the symmetric first–best solution has a wider s–S band than the symmetric Markovian equilibria. A further contribution of this paper is the development of a framework for continuous time games which allows players to react instantaneously to their opponent's moves. We mention various applications of the theory and discuss in detail an application to product innovations.

1. Introduction

In many economic applications, agents can make discrete or discontinuous changes in variables that affect their payoffs. Such changes may complement continuous changes: for example, a firm can decide to introduce an entirely new model of a product that it has on the market rather than make minor improvements or modifications in it. A similar problem is encountered in the decision between learning more about a currently used technology or discovering a new and totally unknown one, (Jovanovic–Rob (1990)). Similarly it is important to consider discontinuous adjustments through devaluations in foreign exchange markets as an alternative to open market interventions. Yet another example of discrete change is the entry and exit decisions of a firm (Dixit (1989b)). A final example is the choice faced by an agent between making discrete price changes and letting relative prices change continuously on account of inflation (Caplin–Sheshinski (1988)). Unfortunately the distinction between continuous and discrete changes is necessarily absent from discrete time models[1] while this difference can be seen immediately and crisply in a continuous time formulation. In the latter framework one can think of continuous changes as determining the rate of change of the economic environment whereas discrete changes determine "jumps" in such an environment. In this paper we study a continuous time stochastic model whose outstanding characteristic is that agents can make discontinuous changes in the environment.

In many applications in which a consideration of discrete change is important, such changes are made in a strategic setting. For instance each of the examples cited above is naturally embedded in models in which there are only a few economic agents who behave strategically[2]. This suggests the importance of a game theoretic formulation for "discontinuous action" models. Since the basic problem involves changes in the game environment, a purely repeated game model (in which the game environment is unchanging) obviously cannot be employed. The appropriate framework is that of a stochastic (or in other usage, dynamic or Markovian) game. We analyze a general stochastic game (see Partharasarathy (1973) or Fudenberg–Tirole (1989) for a description of such games) with the further specification that players' actions have discrete effects on the state variable of the game.

It is well known that in continuous time games in which players are allowed to make instantaneous moves, problems of consistency in the description of outcomes arise (see Stinchcombe (1989) for a number of examples). We analyze here a game in which discrete changes are costly (in a sense which will be made precise later). We then introduce a framework for a continuous time game which is flexible enough to allow players to react instantaneously to their opponents' moves. This framework may be of more general interest in the consistent analysis of continuous time games.

1.2 Description of the Game and Results

A real–valued state variable X(t) (the game environment) evolves according to some exogeneous stochastic process. Either of two players can at any time of their choosing change the state by a discrete and arbitrary amount. There is a fixed cost to making any such change. Players get per period returns which have two components: there is a direct payoff which depends on the state and from this is subtracted the costs of any discrete action undertaken. The critical assumption made on the direct payoffs is that they are monotone: player 1's payoff increases in the state, while that of player 2 decreases. A strategy of a player specifies a set of stopping times and associated jump sizes. We say that a strategy is Markovian if it conditions only on the payoff relevant variable X. Note incidentally that on account of the fixed cost the (gross of cost) return function of a player is necessarily non–concave (and this often has non–existence implications for Nash equilibria of course).

Our main theorems (Theorem 5.4 and 5.5) say: there is a class of simple Markov perfect equilibria in the stochastic game which are characterized by two–sided (s,S) rules. In such an equilibrium the strategies of the two players are completely described by four parameters: $L, L+\theta$ (for player 1) and $U, U-\mu$ (for player 2). If the state is ever at or below L, player 1 jumps it up to $L+\theta$, whereas if it ever gets upto U or above, player 2 jumps it down to $U-\mu$. Of course $L \leq L+\theta < U$ and $L < U-\mu \leq U$.

Figure 1

The parameters are determined by the primitives of the game (the fixed costs and period payoffs): by way of illustration we explicitly compute these parameters in some special cases (in which the flow payoffs are linear). We further show that it is not essential that both players be "active" in equilibrium. There are equilibria in which one or more players never take any discrete actions, i.e it may be the case for instance that $\theta=0$ and $L=\underline{E}$, where $\underline{E}$ is the minimum value of the state.

There are two results preliminary to the main theorems which are of independent interest. The first (Theorem 4.0) establishes that the best response to a one–sided (s,S) rule is a one–sided (s,S) rule. An equivalent way of stating that result is: the optimal policy of a single agent, whose payoffs are any monotone increasing function and who is faced with exogeneous two–sided uncertainty and an impulse barrier, is a (s,S) rule. This is a result new to the optimal control literature (the previous results were for the special case of linear flow payoffs (see Harrison et.al. (1983) and Constantanides–Richard (1978)). The second result is a monotonicity lemma. Recall that a player's best response problem is non–convex. Hence in games with discrete actions one cannot appeal to standard fixed point arguments. We define instead a partial order on the set of player strategies and the monotonicity lemma establishes that the best response respects this order. Equilibrium can then be derived by appealing to Tarski's fixed point theorem for monotone maps. The construction may be of more general interest.[3]

We further show that the first–best solution is also a two–sided (s,S) rule. We compare the symmetric first–best solution with the symmetric Markov–perfect equilibrium. A priori there is no reason to expect a systematic comparison between these two solutions. The objectives are different in the two cases (the first–best internalises all returns and costs) and the optimization problems are different too (the first–best is unconstrained whereas the equilibrium best–response is against an exogeneous impulse barrier). Consider the symmetric first–best solution (in which a discrete action is taken whenever the state gets down to L_c or up to $|L_c|$, $L_c < 0$, and the action is to move the state to 0). The corresponding symmetric Markov perfect equilibria is one in which the state is "jumped" whenever it gets below L_n or above $|L_n|$ and then it is moved to 0.[4] Theorem 6.2 establishes that $L_c < L_n$, i.e. that the (s,S) band is wider in the first–best than in equilibrium. This immediately implies that the environment of the game is more volatile under non–cooperative play than under first–best.

A final contribution of this paper is the introduction of a general framework for continuous time games with discrete instantaneous actions. Our formalisation involves the consideration of the play of an extensive form repeated game "within every instant". Such a device allows a consistent modelling of instantaneous reactions to other players' moves. Moreover the fact that moves are costly allows us to assign payoffs in a natural way to inconsistent outcomes. A detailed discussion of these issues is contained in Appendix 1.

1.3 An Application to Product Innovation

Our original interest in the general problem was sparked by a desire to understand the ongoing nature of product innovations. We discuss in some detail how such a problem can be analyzed within the game form described above.

Economists going back to Schumpeter have emphasized a view that innovations are once for all breaks with the past.[5] Consequently, much of the industrial organization literature has analyzed innovative activity by way of "the race for a single patent".[6] Historical and empirical evidence points, on the other hand, to the great importance of ongoing change. A large portion of productivity growth or quality improvement has often come from a cumulation of small changes to original technologies.[7] In fact the intensity of competition in product quality is a proof of the fact that the innovation process is an ongoing one. It is this product quality competition by which an initial technology is transformed into a spectrum of differentiated products. Such quality competition is probably the best explanation for the great imperfection of patent protection in practice.[8]

We shall now show why the game described above is a good model for this process of repeated innovations. Specifically consider a model in which each of two competing firms in a market innovate by introducing quality upgrades. The increase in product quality can be by an arbitrary amount and is achieved at some cost. The most important characteristic of this cost is that a fixed minimum amount has to be paid regardless of the magnitude of the quality improvement (on account of design changes, advertising etc.). Products earn current returns based on the two quality levels (for example, profits from either Cournot or Bertrand competition over the differentiated products). Let us suppose that current returns depend solely on *relative qualities*, i.e., that a firm's flow profits depend solely on how much better (or worse) its product is relative to that of the competition. This simplification implies that differences in quality alone is the payoff relevant variable. Further make the natural assumption that market demand has some random component. Interpreting differences in quality as the state and quality improvements as a discrete activity we are in the framework of the general stochastic game.

In the context of this application, the observable prediction of the two sided (s,S) rule equilibrium is that the payoff relevant quality difference will be seen to move continuously under the effect of random shocks until the relative quality of one of the products deteriorates so much that the laggard initiates a discrete, discontinuous improvement. Such activity may be interpreted as the emergence of "quality ladders". The interesting feature of such a ladder is that the gradualness of innovation dissemination is driven not by technological factors, but purely by strategic factors. Further there is, in such an equilibrium, no permanent leader or follower: the role of the market technological leader alternates. Of course it may be the case that one of the firms is passive. In this case we have a "perpetual leader" which makes all the innovations in the industry.

Section 2 sets up the model, Section 3 contains illustrative examples while Sections 4 and 5 analyze the best response and equilibrium issues. In Section 6 we consider the cartel's problem. Section 7 concludes.

2. The Model

2.1 A Description of the Game

Let $E \subseteq \mathbb{R}$. E will be the state space. We shall consider either of two specifications for E. In the first specification we take $E = \mathbb{R}$ and refer to this as the non-compact domain case. Alternatively we shall take $E = [\underline{E},\overline{E}]$, $-\infty < \underline{E} < \overline{E} < \infty$, and refer to this as the compact domain case.[9] The index i will refer to a generic player. In all statements pertaining to i, the index j will refer to the other player. Time is continuous and the horizon is infinite.

Assume until further notice that $E=\mathbb{R}$. A state variable, the environment of the game, evolves as a joint consequence of the actions of both players and some underlying Markov process. In particular, players can make discontinuous and instantaneous changes in the game environment (at some cost). This is modelled in the following manner; at any instant a player can "jump" the state by an (arbitrary) finite amount of his choosing. Let the jump by player i at instant s be denoted $\xi_i(s)$. Let A_t (respectively B_t) denote the (random) set of times before t at which player 1 (respectively 2) jumps. In the absence of any intervention by either player the environment at any instant changes according to the (infinitesimal) change in some exogeneous (real valued) process $[Z(t)\colon t\geq 0]$. Let (Ω,F) be a measurable space with $(F_t)_{t\geq 0}$, an increasing family of σ fields generated by $[Z(t)\colon t\geq 0]$, $F_t \subset F$, for every t. Call the environment of the game at time t, X(t). The evolution of this process can then be written as[10]

$$X(t) = Z(t) + [\ \Sigma_{s\in A_t}\xi_1(s) + \Sigma_{s\in B_t}\xi_2(s)] \tag{1}$$

We make the following **assumption on the exogeneous environment:**

(A0) The stochastic process $[Z(t): t \geq 0]$ is a Brownian motion with drift–variance (m, σ^2).

A convenient restriction to consider, although inessential for us, is that $m = 0$, i.e., exogenous factors on average leave the game environment unchanged.

The **direct flow payoffs** of each player depend on the state variable and will be denoted $r_i(x)$. Purely for notational purposes we make throughout a symmetry assumption on this payoff function. It should be understood (and the reader can easily check this) that none of the analysis that follows is contingent on this assumption:

(A1) **Symmetry**: r_i and r_j are symmetric, i.e. $r_i(x) = r_j(-x)$ for all $x \in \mathbb{R}$.

Let us denote the returns of players 1 and 2 as $r(x)$ and $r(-x)$ respectively. The only critical assumption on returns will be:

(A2) **Monotonicity**: r is strictly increasing in X and absolutely continuous.

When the domain of the state variable is $\mathbb{R}$ we shall employ the following **boundary conditions:**[11]

(A3) i) $\lim_{x \uparrow \infty} r(x) = \bar{r} < \infty$

(ii) $\lim_{x \downarrow -\infty} r(x) = -\infty$

A principal characteristic of discrete changes (in the game environment)is the fact that even small changes require a minimal outlay of resources. The most natural way to model such phenomenon would be to assume that there is a fixed cost to any discontinuous act. Let $c(\xi)$ denote the **cost to a discrete change** of amount $\xi \neq 0$. We assume for simplicity that the only cost is a fixed cost:[12]

(A4)
$$\begin{aligned} c(\xi) &= c_0 > 0 \text{ if } \xi \neq 0 \\ &= 0 \text{ if } \xi = 0 \end{aligned}$$

Note that in the pure fixed cost case,(A4),player 1 can make the state arbitrarily large (which he prefers) at no extra cost (and similarly 2 can make the state arbitrarily small at no extra cost). Indeed as single decision makers this is exactly what they would attempt to do.

2.2 Histories and Strategies

The continuous time formulation of the stochastic game, although natural and simplifying from several viewpoints, however presents substantial technical problems in specifying admissible strategies and outcomes. In Appendix 1 we present a formal development of these issues and restrict ourselves in this sub-section to a largely informal discussion. There are two principal complexities:

2.2.1 No "Next Instant." There is no "next instant after t." This is problematical when player i moves at instant t and j wishes to react to i's action "as soon as possible." To allow for such instantaneous reactions we will consider an extended definition of "time," which admits "reactions within an instant." At every instant t, *within the instant* firms can act countably often and in essence play an infinitely repeated game. "Time" then is a two variable index: the first component of it is "real" time and the second the number of the move in the extensive form game which takes place in that instant. The notation will be $\tau = (\tau_1, \tau_2)$. The "real" time component will be interchangeably denoted t. So, the change in the state at instant t is defined as

$$X(t) - X(t-) = \sum_{\tau_1 = t} \xi_1(\tau) + \sum_{\tau_1 = t} \xi_2(\tau) \tag{2}$$

Of course given that there is a fixed cost to be paid whenever $\xi \neq 0$, in a best response, only a finite number of potential action opportunities within the instant are actually used. In Appendix 1, we develop formally the interpretation of time and define extended σ-fields and stochastic processes for this time index. Notice incidentally, that if the infinite sum in (2) is not convergent we merely say the game terminates at t (with negative infinite payoffs to the player who made infinitely many moves).

2.2.2 Consistent Outcomes. A partial history at "time" τ is a description of activity for all s < τ_1, and moves within an instant τ_1 up to the current one (τ_2) (see Appendix 1 for details). An admissible strategy is just a prescription of action at all such nodes which only use available information (are progressively measurable). Let σ_i, i = 1,2 denote a pair of such strategies. An *outcome* is defined (as in (1)) by

$$X(\tau) = Z(\tau) + \sum_{\tau' \in A_\tau} \xi_1(\tau') + \sum_{\tau' \in B_\tau} \xi_2(\tau') \tag{3}$$

The series in (3) may of course not be convergent a.s., i.e. for some environments there may not be consistent outcomes. If X(t) is not well-defined for some environment because a countable number of actions are clustered around some instant $T \leq t$, we say that this is payoff equivalent to saying that the game terminates at T with lifetime payoffs, in this environment, of $-\infty$. (In Appendix 1 we formalize this.)

The lifetime payoffs for a pair of strategies σ_i, σ_j is given by

$$W_i(\sigma_i,\sigma_j)(x) = E_x \left\{ \delta \int_0^T r_i(x_s)e^{-\delta s} ds - \sum_{\tau} c(\xi_i(\tau))e^{-\delta\tau_1} \right\} \tag{4}$$

where T is the termination date (possibly ∞) and τ_i are the dates of player i's discrete actions; both of these random variables have distributions generated by the strategies. Further, x is the initial state.

2.3 Markovian Strategies and Markov Perfect Equilibrium

At any "time" τ, the payoff relevant variable is $X(\tau)$ alone. A player follows a *Markovian strategy* if at each moment he looks at the current value of X to determine whether he acts and by how much. Hence a Markovian strategy can be identified with a measurable function g: $E \to \mathbb{R}$. We shall understand $g(x) = x$ to mean "no action."

A *Markov Perfect Equilibrium* (MPE) is a pair of Markovian strategies g_i^* such that

$$W_i(g_i^*,g_j^*)(x) \geq W_i(g_i,g_j^*)(x), \text{ for all x, i \& Markovian } g_i \tag{5}$$

From hereon we restrict attention to Markovian strategies and analyze MPE. Of course it is easy to show that in our model MPE continue to be subgame perfect equilibria when all admissible strategies are considered.

2.4 An Application to Product Innovations

In this section we use the product innovation problem that was discussed in the introduction to motivate the assumptions (A1) – (A4). Quality or technology level or the attributes of firm i's product is given by the single dimensional payoff relevant variable $Y_i \in \mathbb{R}$. At any time, a firm can change this quality by an arbitrary discrete amount ξ_i[13] (after paying some cost). Additionally, the payoff characteristics of i's product may change randomly, through shifts in market demand. This exogeneous variable is denoted Z_i. Hence, i's payoff relevant quality variable is given by

$$Y_i(t) \;=\; Z_i(t) \;+\; \sum_{\tau_1 \le t} \xi_i(\tau) \tag{6}$$

Relative quality (from i's viewpoint) will be denoted by $Y_i - Y_j \equiv X_i$. Its evolution is described by:

$$X_i(t) \;=\; [Z_i(t) - Z_j(t)] \;+\; [\sum_{\tau_1 \le t} \xi_i(\tau) \;-\; \sum_{\tau_1 \le t} \xi_j(\tau)] \tag{7}$$

Given qualities or product attributes, Y_i, firms make production or pricing decisions. To concentrate attention on competition along the quality dimension, it will be convenient to employ a *reduced form flow payoff* which depends only on the two quality levels. One could imagine that the firms are Cournot or Bertrand competitors in the market and the flow payoffs correspond to the (one–period) Cournot or Bertrand profits. Consider the following payoff assumption that says the flow payoffs only depend on differences in quality:

(A5) the flow payoffs are $r_i(Y_i - Y_j)$ $\quad i = 1,2$

Although (A5) clearly precludes applications to certain products or industries where scale effects are important, there are many examples in the product differentiation literature where (A5) is satisfied.[14] The monotonicity assumption (A3) is now immediate since all it says is that firm i's (Cournot or Bertrand) profits are higher the better its product relative to j's. The upper bound on profits follows if we assume that the market size never gets unbounded. The unboundedness of losses is more difficult to justify in this example but would hold if the quality laggard produced an unbounded quantity in a Cournot equilibrium. If we assume that each firm's exogeneous demand is driven by a Brownian motion, then from (7) and (A5) it follows that relative quality, X_1 is also driven by such a process and serves as the "state variable" for this problem.[15]

3. Some Examples

Before we proceed to the general analysis of Sections 4 and 5 let us consider three simple examples. These examples will illustrate the theorems that follow. In each of these examples we make some special simplifying assumptions that allow us to compute the equilibria of the game. Throughout $E = \mathbb{R}$.

Example 3.1 $r(x) = x$ **and** $Z(t)$ **is a deterministic process with** $\dot{Z} = 0$

Example 3.1 analyzes a trivial deterministic law of motion which could be thought of as an extremely simple Brownian motion with m=0 and v=0. We claim that the following strategies constitute the **only** Markov–perfect equilibria in the game:

Take any pair (U^*, L^*) **satisfying** $U^* - L^* = c_0$. **Player 1 acts iff** $X < L^*$ **and then jumps the state to** U^*. **Player 2 acts iff** $X > U^*$ **and then jumps the state to** L^*.

Consider player 1's best response problem. He clearly does not want to jump the state beyond U^*. If he jumps at all his best option is to move the state to U^*. A discrete move at cost is worthwhile only if the current state is sufficiently payoff–poor. The point of indifference is precisely L^*. The result is a two–sided (s,S) rule equilibrium. A series of simple arguments should convince the reader that the above conditions are also necessary for an equilibrium. Note that the band width, $U^* - L^*$ is increasing in the fixed costs. Also note that the first–best (with flow payoffs $r(x) + r(-x)$) is never to make discrete changes.

Example 3.2 $r(x) = 1$ **for** $x \geq 0$, $r(x) = 0$ **when** $x<0$ **and** $Z(t)$ **is a deterministic process with** $\dot{Z} = 1$.

This example is also defined for a trivial Brownian motion, with v=0 and m=1. The flow payoffs are only weakly monotone; this simplifies the analysis but is not crucial to the example. Indeed, at some analytical cost, one could use here the linear returns specification of Examples 3.1 and 3.3. The **only** Markov perfect equilibria of this game are:

Case 1: $c_0 \leq 1$. **Player 1 acts iff** $X < L^*$ **and then jumps the state to** $X = 1$ **where** L^* **is defined by** $e^{\delta L^*} = 1 - c_0$. **Player 2 never acts.**

Case 2: $c_0 > 1$. **Neither player acts.**

That the strategies of case 1 form an equilibrium can be checked directly: suppose player 2 is a passive player. If player 1 ever jumps he obviously wants to jump the state to some positive number, say X=1. The question is: will he ever act and if yes, what states will he act at? The maximized returns (or value function) must be increasing in X and hence the jumping region must be of the form $(-\infty, L)$. It is clear that an immediate jump yields the player $1 - c_0$ whereas waiting for the state to improve yields lifetime returns of $e^{\delta x}$ if the current state is $x < 0$. A comparison of the two options yields the critical value of L^*. Let us now check that it is a best response for player 2 to be passive. If he acts he would clearly like to move the state to L^*. Pick any state from which player 2 acts. His returns are clearly cyclical, each cycle yielding $1 - e^{\delta L^*} - c_0$. But by the definition of L^* this is exactly zero. So player 2 is indifferent between acting and not acting at any state.

That this is the only equilibrium can be seen by noting that if there is an alternate equilibrium in which player 2 is active so must player 1 be. Else, given the upward drift of the state process, player 2 would like to jump the state to its minimum and with a non–compact domain this implies that he would have no best response. For the same reason an active player 2 in equilibrium must jump the state down to the minimum point at which 1 is inactive. Let this point be denoted L. As the state progresses up from L to 0, player 2 makes $1 - e^{\delta L}$ over this cycle and this return must be worthwhile, i.e. it must be that $1 - e^{\delta L} \geq c_0$. If the inequality is strict, i.e. 2 makes strictly positive returns over the cycle, then he initiates the cycle as soon as possible, i.e. at $x = 0$. But this is not possible since in that case the active player 1 makes negative returns in the game. So $1 - e^{\delta L} = c_0$, i.e. $L = L^*$. Finally, player 1 would initiate a jump at L^* only if 2 is passive; else 1 makes losses in equilibrium.

It is clear that if the flow payoffs over the infinite horizon do not cover fixed costs, then neither player acts, i.e. the strategies of case 2 are an equilibrium and the only ones under those cost specifications. Example 3.2 illustrates two–sided (s,S) equilibria with one or both players passive. Note that L^* is decreasing in the fixed cost. Further, the first–best policy is again "no jump."

Example 3.3 **$r(x) = x$ and Z is a Brownian motion.**

The previous two examples had easily computable equilibria but on account of the deterministic transitions, no interesting state dynamics. In this example we compute the equilibrium when the underlying state process is in fact a (non–trivial) Brownian motion. The calculations here are more labored and hence have been consigned to Appendix 2.

Lemma 3.1 (Best Response): Suppose that player 2's strategy is: jump iff $X \geq U$ and then jump down to $U-\mu$, for some $U-\mu < U < \infty$. Then, player 1's best response is given by: jump iff $X \leq L$ and then jump to $L+\theta$, where θ is a constant independent of

U,μ and the "band" U − L is determined solely by μ.

Lemma 3.2(Equilibrium): The symmetric equilibrium of the game is given by: player 1 (respectively 2) jumps iff $X \le L^*$ (respectively $X \ge |L|^*$). The jump size is θ^* where θ^* is given implicitly and uniquely by

$$-\left(\frac{1-e^{\theta}}{1+e^{\theta}} - \frac{1-e^{-\theta}}{1+e^{-\theta}}\right) + c_0 - \theta = 0$$

The band size can be computed from the following

$$e^{2L} = \theta\left(1 - \left(1 - \frac{4(e^{\theta}-1)(1-e^{-\theta})}{\theta^2(1+e^{-\theta})(1+e^{\theta})}\right)^{1/2}\right)\frac{1+e^{-\theta}}{2(e^{\theta}-1)}$$

where $\theta = \theta^*$. Again the first-best is "no jump."

4. A Best Response to (s,S) is an (s,S) Rule

Suppose player 2's strategy is in fact given by (U,μ) where $-\infty < U - \mu \le U < \infty$. Let V denote the value function for 1's best response problem:

$$V(x) = \sup_{g} W_1(g,(U,\mu))(x)$$

where (with some abuse of notation), $W_1(g,(U,\mu))(x)$ denotes the discounted expected return to a Markovian strategy g in response to a (U,μ) strategy, with initial state x. We now show that there is a unique Markovian best response strategy, say g^*. Further, it is also of the class (s,S) in that it is characterized by a single jumping point L and a single destination $L + \theta$.

Theorem 4.0 Consider a given Markovian policy of player 2, characterized by a pair (U,μ). Then:

i) There exists a unique optimal policy of player 1, say g^*, in the set of admissible strategies; this optimal policy is a Markovian policy, characterized by a pair (L,θ).

ii) The value function V is differentiable.

For future consideration, define $M \equiv \sup_{x \in E} V(x)$. It is easy to see that $V(x) \ge M -$

c_0 for all x. Further, $V(x) > M - c_0$ implies $g^*(x) = x$

The logic of the proof of Theorem 4.0i) is as follows: a straightforward modification of the argument in Bensoussan–Lions (1982, chapter 6 Theorem 1.1) establishes that there is an optimal policy which is characterized by a set of jump points E^* and associated destinations.[16] Whenever the state is at $x \in E^*$ it is jumped to a destination d(x). We next prove some properties of the value function which are then used to characterize optimal policies. We establish first that the value function to the player's best response problem has a unique maximum, denoted $L+\theta$. This suffices to show that whenever a discrete action is undertaken, it always involves a jump to $L+\theta$, i.e. for all $x \in E^*$, $d(x) = L+\theta$. We then show that the value function to the left of $L+\theta$ is strictly increasing. Hence there is a unique state, denoted L, such that if $x \leq L+\theta$, it is optimal to jump if and only if the state is less than L. Finally we demonstrate that it is never optimal to jump when $x > L+\theta$. The implication of these arguments is that for any optimal policy $E^* = (-\infty, L]$ and, consequently, that there is a unique optimal policy.

The following operator notation is used repeatedly in the sequel. Suppose we take an initial state $x \in (a,b)$. Let us consider the discounted returns to the policy: take no action as long as the state stays in the interval (a,b). The first time the process $[X_t : t \geq 0]$, hits either a (denoted T_a) or b (denoted T_b), the decision problem is terminated with terminal rewards V(a) and V(b) respectively. This policy defines an operator $\Gamma V(x;a,b)$ as follows:

$$\Gamma V(x;a,b) = E_x \left\{ \delta \int^{T_a \wedge T_b} e^{-\delta s} r(x_s) ds + e^{-\delta T_a} V(a) I(T_a \leq T_b) + e^{-\delta T_b} V(b) I(T_a > T_b) \right\}$$

where I(C) is the indicator function on a set C.

We first argue the straightforward point that the value function must be continuous. Pick any $x \leq U$ and suppose that it is a point of discontinuity. It cannot be that there is a sequence $[x_n : n \geq 0]$ such that $\lim_{n \to \infty} V(x_n) < V(x)$. Else there is an interval (a,b), $a < x < b$ such that no ϵ-optimal policy from initial state x "jumps" in (a,b). Clearly $V(x_n) \geq \Gamma V(x_n;a,b)$. Further $V(x) \leq \Gamma V(x;a,b) + \epsilon$ and the operator Γ is continuous. Combining all this yields $\lim_{n \to \infty} V(x_n) \geq V(x) - \epsilon$. This is true for all $\epsilon > 0$ and hence we conclude that if there is a discontinuity it must be the case that

$\lim_{n\to\infty} V(x_n) > V(x)$. In that case none of the states x_n can be a "jumping" state. In fact there must be an interval (a,b), $a < x < b$ such that no ϵ–optimal policy from any initial state x_n "jumps" in this interval. Hence, $\Gamma V(x_n;a,b) \geq V(x_n) - \epsilon$ and $V(x) \geq \Gamma V(x;a,b)$. The above arguments repeat to show that the discontinuity is impossible. Finally from (A3) it follows that this continuous function cannot attain its supremum for arbitrarily small x. Hence V attains maxima on a finite domain.

Lemma 4.1: Suppose $V(x^*) = M$. Consider $x_1 < x_2 < x^*$, with the property that $V(x_1) > M - c_0$. Then, $V(x_2) > V(x_1)$.

Proof: Let $\Delta \equiv x_2 - x_1$, and define $x_3 = sup\{x \in (-\infty,x_1)\colon V(x) = M - c_0\}$. By (A3) this set is non–empty and by the continuity of the value function, $x_3 < x_1$. Similarly, define $x_4 = inf\ \{x \in (x_1,x^* - \Delta)\colon V(x) = M - c_0\}$. We have two cases to consider:

Case 1 If the relevant set is empty, then define $x_4 = x^* - \Delta$. Let T_3 and T_4 denote the first hitting time of x_3 and x_4, respectively, when the initial state is x_1. Then,

$$V(x_1) = \Gamma V\ (x_1;\ x_3,x_4)$$

Given the continuity of sample paths of the stochastic process Z_t, the state remains within (x_3,x_4) prior to hitting either boundary. Hence, before $T_3 \wedge T_4$ there is no action by either player. Further, given the spatial homogeneity of the process, we know that

$$S_3 \equiv inf\ \{t \geq 0\colon X(t) = x_3 + \Delta;\ X(0) = x_2 = x_1 + \Delta\}$$

has the same distribution as T_3. Similarly,

$$S_4 \equiv inf\ \{t \geq 0\colon X(t) = x_4 + \Delta;\ X(0) = x_2 = x_1 + \Delta)$$

has the same distribution as T_4. Hence,

$$V(x_2) \geq \Gamma V\ (x_2;\ x_3+\Delta,\ x_4+\Delta) > \Gamma V\ (x_1;\ x_3,\ x_4) = V(x_1) \qquad (8)$$

The first inequality in (8) follows from the fact that the proposed policy for initial state x_2 may of course not be the optimal one. The strict inequality uses the following facts: the respective stopping times have the same distributions and further,

$V(x_3 + \Delta) = V(x^*) \geq V(x_3)$ and $V(x_4 + \Delta) \geq V(x_4) = M - c_0$. Add to this the fact that the flow payoffs r are strictly increasing in the state and the identity in distribution of S_3 and T_3 (and S_4 and T_4), i.e.

$$E_{x_2} \int_0^{S_3 \wedge S_4} e^{-\delta s} r(x_s)\, ds > E_{x_1} \int_0^{T_3 \wedge T_4} e^{-\delta s} r(x_s)\, ds.$$

Combining all this, (8) is proved.

Case 2. The set in the definition of x_4 is non-empty. Compare now the lifetime returns to starting at x_1 and eventually acting at either x_3 or x_4 against starting at x_2 and eventually acting at either $x_3 + \Delta$ or $x_4 + \Delta$. The above argument repeats. The proof of Lemma 4.1 is complete. ∎

Remark: (A3), the boundary condition, is really not required for the above proof. The only modification required, in the absence of this assumption, would come from the fact that the set defining x_3 might be empty. In such a case take $x_3 = -\infty$. The arguments then are identical.

It immediately follows that V has a unique global maximum:

Corollary 4.2. There does not exist $x^* \neq x^{**}$, *s.t.* $V(x^*) = V(x^{**}) = M$.

Let g^* denote an optimal policy. Since the cost of discrete actions are independent of the size of the action it follows that if a player is to act at any state he must necessarily move the state to the unique maximum of the value function, i.e. that $g^*(x) \neq x$ implies that $g^*(x) = x^*$. Of course, $x^* < U$. This follows immediately from Corollary 4.2. In the sequel we call x^*, $L + \theta$. Since $g^*(x) \neq x$ if $V(x) > M - c_0$ it further follows that if player 1 decides to act at some $\hat{x}$ then he acts at all $x < \hat{x}$. We prove a somewhat weaker version of this claim now. Notice that the policy of not acting on an entire interval $(\infty, \hat{x})$ is equivalent, from a payoff standpoint, to one in which the player acts at almost all $x < \hat{x}$. In other words the equivalence class for such a strategy is: player 1 acts at all but some isolated points in $(-\infty, \hat{x}]$. We now show that the optimal policy must be from this equivalence class:

Corollary 4.3. If $\hat{x} < x^*$ is an "action state," then so is almost every state smaller

than $\hat{x}$, i.e., if $g^*(\hat{x}) \neq \hat{x}$, then there does not exist an interval (a,b), $b \leq \hat{x}$ such that $g^*(x) = x$, for every $x \in (a,b)$.

Proof: By lemma 4.1 for all $x < \hat{x}$ it must be that $V(x) = M - c_0$. The proof of the lemma immediately implies that there cannot be an interval of states on which there is inaction and simultaneously a constant value. ∎

Now, let $L \equiv sup\ \{x < x^* : g^*(x) \neq x\}$. In the next lemma we show that $L > -\infty$.

Lemma 4.4 The "jumping point" is finite, i.e. $L > -\infty$.

Proof: Suppose to the contrary that $g^*(x) = x$, for every x. Fix $\epsilon > 0 > K$ and pick $\hat{x} < x^*$, such that $r(x) < K$ for all $x \leq \hat{x}$. This is possible by A3(ii). Let $\tilde{x} < \hat{x}$ further satisfy $1 - Ee^{-\delta\hat{T}} > 1 - \epsilon$, where $\hat{T} = inf\{t \geq 0\colon X(t) = \hat{x};\ X(0) = \tilde{x}\}$. By hypothesis,

$$V(\tilde{x}) = \Gamma V(\tilde{x}\ ;-\infty,\hat{x}) \tag{9}$$

$$< K\ (1 - Ee^{-\delta\hat{T}}) + \epsilon\ V(\hat{x}) \tag{10}$$

Note that by hypothesis it is inoptimal to ever jump, i.e.

$$V(\tilde{x}) \geq M - c_0 \tag{11}$$

Letting $K \to -\infty$ and $\epsilon \downarrow 0$, (10) and (11) yield $M = -\infty$ and this is clearly a contradiction. The lemma is proved. ∎

Remark: In the absence of the boundary condition (A3ii) one cannot rule out the case that "passive" play may be optimal, i.e. that $L = -\infty$.

What we have established so far is: in any optimal policy discrete actions are taken for any state below $L > -\infty$, with a "jump" up to $L + \theta$. Further, for any $x \in (L, L + \theta]$, player 1 refrains from acting. To complete the proof of our result we now show that no action is taken for $x \in (L + \theta, U)$.

Lemma 4.5 $g^*(x) = x$, for every $x \in (L + \theta, U)$.

Proof: Let $x_1 = inf\ \{ x \in (L+\theta,U)\colon g^*(x) \neq x\}$. If the set is empty the lemma is proved. Else let $T_2 = inf\ \{t \geq 0\colon X(t) = x_1;\ X(0) = L + \theta\}$ and $T_1 = inf\{t \geq 0$:

$X(t) = L;\ X(0) = L + \theta\}$. Then define $\hat{x} \in (L+\theta,\ U)$ in a way that $x_1 + \hat{x} - (L+\theta) < U$. Denote $\Delta = \hat{x} - (L+\theta)$. Then,

$$V(L + \theta) = M = \Gamma V(L + \theta;\ L, x_1)$$

$$V(\hat{x}) \geq \Gamma V(\hat{x}\ ;\ L+\Delta,\ x_1+\Delta) > \Gamma V(L+\theta;\ L,\ x_1) = M \geq V(\hat{x})$$

The reasoning which results in the above inequalities is identical to that of the proof of Lemma 4.1. The strict inequality clearly contradicts the definition of M.∎

We have proved that any optimal policy must be of the (s,S) type. Since L and $L+\theta$ are uniquely defined we have additionally established the uniqueness of best response. We complete the proof of the theorem by establishing the differentiability of the value function.

Let σ_2 denote the strategy of player 2, characterized by the pair (U,μ), and consider for any x the supremum of the payoff to player 1 over the set of admissible strategies. Straightforward modifications of the arguments in Bensoussan–Lions (1982), Chapter 6, give that this is equal to the supremum over the set of Markovian strategies, i.e., it is given by V(x) as defined at the beginning of section 4.1.

From Ito's Lemma it follows immediately that the payoff, for any initial point $x \in [L,U]$, of following the Markovian policy characterized by the pair $(L, L+\theta)$ is given by the solution of the following differential equation together with two boundary conditions. For notational simplicity we report the case m=0. The modifications to the expressions when $m \neq 0$ are well-known.

$$\frac{\sigma^2}{2} V''(x) - \delta V(x) + \delta r(X) = 0 \quad \text{for } x \in (L,U) \tag{12}$$

$$V(U) = V(U - \mu),$$

$$V(L + \theta) - c_o = V(L)$$

We recall that a solution to (12) is a function of the form

$$V(x) = a_1(L,\theta)\, e^{\alpha_1 x} + a_2(L,\theta)\, e^{\alpha_2 x} + R(x)$$

where $\alpha_1 = -\alpha_2 = (\sqrt{2\delta})/\sigma$ and R(x) is a special solution. For example, R can be taken to be:

$$R(x) = (\alpha_2/2) \int_0^x [e^{\alpha_1(x-y)} - e^{\alpha_2(x-y)}]\, r(y)\, dy$$

From the boundary conditions that follow the differential equation (12) one can straightforwardly calculate $a_1(L,\theta)$ and $a_2(L,\theta)$. It is easy to see that these depend on the special solution and c_0. Clearly the special solution is a continuously differentiable function and hence so are $a_1(L,\theta)$ and $a_2(L,\theta)$. The value function corresponds to the optimal choice of L and θ and in particular an optimal choice of $a_1(L,\theta)$ and $a_2(L,\theta)$. The argument in Dixit (1989a) now can be applied directly to show (that the first order conditions for such a choice imply) that V is a differentiable function. This completes the proof of Theorem 4.0. ∎

Remark: Bensoussan–Lions [1982], Constantanides–Richard (1978) and Harrison et al. [1983] have also analyzed this problem. The Bensoussan–Lions formulation, although very general, does not directly cover the case of a boundary with impulse reflection (as at U). Consequently their arguments on the sufficiency of Markovian policies and the existence of an optimal policy has to be appropriately modified. More importantly, in the general formulation of their problem it is not possible to establish the sharp (s,S) rule characterization for optimal policies . Harrison et al. (1983) and Constantanides–Richard (1978) do establish the optimality of (s,S) policies but restrict attention to the case of linear payoffs. (Further, they consider a somewhat different behavior at the boundary of the state space and their analyses admit fixed as well as proportional costs). Caplin–Sheshinski [1989] analyze a discrete time version of this problem, use a weaker restriction in which r need only be quasi–concave, but restrict attention to decreasing processes. Hence, it would appear that the optimality of (s,S) policies that we establish in Theorem 4.0 is a more general result, in some dimensions, than those hitherto available.

5. Monotonicity of Best Response and Existence of Equilibrium

In this section we show firstly that the best response map is "monotone" in a sense made precise shortly. Then, we use the Tarski fixed point theorem to establish that an equilibrium in two–sided (s,S) rules exists in the stochastic game. A few remarks on the two steps is in order.

The objective function in a player's optimization problem is fundamentally nonconcave. This is so since any action, howsoever small, involves a positive fixed cost whereas there is no cost for inaction. Non–concavities create well–known problems for equilibrium existence. For our game these problems may be summarized as follows: when the domain of X is compact (a case we detail in Section 5.3), i.e. when $E = [\underline{E},\overline{E}]$, then the best response is a non–convex valued correspondence. For some values of (U,μ) it can be shown that player 1 is indifferent between $L = \underline{E}$, $\theta = 0$ and $L = \underline{E}$ and $L + \theta = x^*$. Of course, given the nonconcavity in the objective function no convex combination of these two policies is as good. Further, for (U,μ) in one neighborhood, $L = \underline{E}$, $\theta = 0$ is the unique best response and in a second neighborhood, $L > \underline{E}$, $\theta > 0$ is the unique optimal choice. In other words, the best response correspondence is neither convex valued, nor does it admit a continuous selection. If the domain of X is the real line, then Theorem 4.0 establishes that, despite the nonconcavity, the best response is unique and indeed further arguments show that the best response function is continuous. In this case, standard fixed point arguments do not work because the domain is not compact. We now show that the best response map is however *monotone* once the right order is specified. This result, interesting in its own right, enables us to apply a fixed point theorem for monotone maps to guarantee existence of equilibrium.

For the next two subsections we continue to discuss the case where the domain of X is $\mathbb{R}$. Let $\Pi = \{(U,U-\mu) \in \mathbb{R}^2\colon\ -\infty < U-\mu \leq U < \infty\}$. Π can be identified with the space of (s,S) policies for player 2. Denote a generic element by π. Similarly, define Σ as the set of (s,S) policies and σ as a generic policy for player 1. We define a partial order on Π and Σ as follows:

Definition 5.1 (i) $\pi \succeq \hat{\pi}$ iff $U \leq \hat{U}$ and $U - \mu \leq \hat{U} - \hat{\mu}$.

(ii) $\sigma \succeq \hat{\sigma}$ iff $L \geq \hat{L}$ and $L + \theta \geq \hat{L} + \hat{\theta}$.

The partial order can be given the following interpretation. $\pi \succeq \hat{\pi}$ may be thought of as implying more "aggressive" behavior by player 2 under π rather than $\hat{\pi}$, in that player 2 acts sooner ($U \leq \hat{U}$) and lowers the state further ($U - \mu \leq \hat{U} - \hat{\mu}$) under the former regime. We now prove the rather surprising result that although the order on the strategy space is partial, the best response map is completely ordered. Let $B_1\pi$ denote a best response (of player 1) to π.

Proposition 5.1 **Consider** $\pi \neq \hat{\pi}$. **Then, either (a)** $B_1(\pi) \succeq B_1(\hat{\pi})$, **or (b)** $B_1(\hat{\pi}) \succeq B_1(\pi)$

or both.

Proof: For notational convenience, suppress temporarily the subscript on B_1. So suppose $B(\pi)$ and $B(\hat{\pi})$ are not ordered. Without loss of generality, suppose that $\hat{L} \leq L < L+\theta \leq \hat{L}+\hat{\theta}$, with either the first or third inequality strict. (This situation is pictured in Figure 2).

Figure 2

We use V (respectively $\hat{V}$) to refer to the value function in response to (U,μ) (respectively $(U,\hat{\mu})$). Then for any $\epsilon>0$, let $x \in (L,L+\theta-\epsilon)$. Define $T_1 = inf\,\{t \geq 0\colon X(t) = L\,; X(0)=x\}$, and $T_2 = inf\,\{t \geq 0\colon X(t) = L+\theta-\epsilon\,; X(0)=x\}$. Then,

$$V(x) \quad = \quad \Gamma V(x;\ L,L+\theta-\epsilon) \tag{13}$$

Similarly, let $T_3 = \inf\{t\geq 0\colon X(t) = L+\epsilon\ ;\ X(0) = x+\epsilon\}$ and $T_4 = \inf\{t\geq 0\colon X(t) = L+\theta\ ;\ X(0) = x\}$. By the spatial homogeneity of Brownian motion, the distribution of T_1 (respectively T_2) is identical to that of T_3 (respectively T_4). Note that

$$V(x+\epsilon) \quad = \quad \Gamma V(x+\epsilon;\ L+\epsilon,L+\theta) \tag{14}$$

Combining (13) and (14) and writing

$$\alpha(x) \ \equiv\ E_x\, \delta \int_0^{T_1 \wedge T_2} e^{-\delta s}\, r(x_s)\, ds\ ,$$

$$\alpha(x+\epsilon) \ \equiv\ E_{x+\epsilon}\, \delta \int_0^{T_3 \wedge T_4} e^{-\delta s}\, r(x_s)\, ds$$

from the identity in distribution of the stopping times we have

$$V(x+\epsilon) - V(x) = [\alpha(x+\epsilon) - \alpha(x)] + [V(L+\theta) - V(L+\theta-\epsilon)]\ E_x\, e^{-\delta T_1}\, I(T_1 \leq T_2)$$

$$+\ [V(L+\epsilon) - V(L)]\ E_x\, e^{-\delta T_2}\, I(T_1 > T_2) \tag{15}$$

From Theorem 4.0, we know that V is differentiable. Hence, from (15) it follows that,

$$V'(x) = \alpha'(x) + V'(L+\theta)\, E_x e^{-\delta T_1} I(T_1 \le T_2) + V'(L)\, E_x e^{-\delta T_2} I(T_1 > T_2)$$

$$= \alpha'(x)$$

$$< \alpha'(x) + \hat{V}'(L+\theta) E_x e^{-\delta T_1} I(T_1 \le T_2) + \hat{V}'(L) E_x e^{-\delta T_2} I(T_1 > T_2) \quad (16)$$

$$= \hat{V}'(x) \quad (17)$$

Note that (16) follows since, either $\hat{V}'(L+\theta) > V'(L+\theta) = 0$ (if $\hat{L}+\hat{\theta} > L+\theta$) or $\hat{V}'(L) > V'(L) = 0$ (if $\hat{L} < L$), or both (and all of this from Lemma 4.1). (17) follows from a derivation identical to (13)–(15). But if $V'(x) < \hat{V}'(x)$, for all $x \in (L, L+\theta)$, then $c_0 = V(L+\theta) - V(L) < \hat{V}(L+\theta) - \hat{V}(L) = c_0$, a contradiction. Hence, the proposition is proved. ∎

Proposition 5.2 Suppose $\pi \succeq \hat{\pi}$. Then, $B_1(\hat{\pi}) \succeq B_1(\pi)$.

Proof: By Proposition 5.1, a contradiction to this claim implies that $B(\pi) \succeq B(\hat{\pi})$ but not $B(\hat{\pi}) \succeq B(\pi)$. Hence, we have $\hat{L} \le L$ and $\hat{L} + \hat{\theta} \le L + \theta < U \le \hat{U}$, with at least one of the two weak inequalities strict. Now let

$$\hat{T} = \inf \{t \ge 0: X(t) = L + \theta;\ X(0) = L\}, \quad \text{and}$$

$$V(L) = -c_0 + V(L+\theta) \ge \hat{\beta}(L) + Ee^{-\delta \hat{T}} V(L+\theta) \quad (18)$$

$$\hat{V}(L) = \hat{\beta}(L) + Ee^{-\delta \hat{T}} \hat{V}(L+\theta) \ge -c_0 + \hat{V}(L+\theta) \quad (19)$$

where,

$$\hat{\beta}(L) = E_L \left[\delta \int_0^{\hat{T}} e^{-\delta s} r(x_s)\, ds - c_0 \sum_{\tau \le \hat{T}} e^{-\delta \tau} \right], \text{ for the policy } (\hat{L}, \hat{L} + \hat{\theta})$$

Combining (18) and (19), and noting that atleast one inequality is strict, we get

$$[V(L+\theta) - \hat{V}(L+\theta)] > Ee^{-\delta \hat{T}} [V(L+\theta) - \hat{V}(L+\theta)] \quad (20)$$

It is straightforward to show (and we do this in Lemma 5.3) that whenever $\pi \succeq \hat{\pi}$, $V(x) \le \hat{V}(x)$. (20) then yields a contraction. ∎

Lemma 5.3 **Suppose $\pi \succeq \hat{\pi}$. Then, $\hat{V}(x) > V(x)$.**

Proof: Suppose, to begin with, that $x \in (L,U)$. A candidate policy that could be employed by player 1 in response to $(\hat{U},\hat{\mu})$ is precisely $B(\pi)$. Since, $\hat{U} \geq U$ and $\hat{U} - \hat{\mu} \leq U - \mu$, with at least one strict inequality , the strict monotonicity of r and spatial homogeneity imply the lemma. But this in turn implies that the lemma holds for all x in $(-\infty,L]$ or $[U,\infty]$. ∎

A similar set of arguments establishes monotonicity in the best response of player 2: $\sigma \succeq \hat{\sigma}$ implies $B_2(\hat{\sigma}) \succeq B_2(\sigma)$. Define the composite map $B = B_2 \circ B_1$ Clearly, $\pi \succeq \hat{\pi}$ implies $B(\pi) \succeq B(\hat{\pi})$. We shall use this very useful property within the context of Tarski's fixed point theorem for complete lattices.

The principal obstacle to a direct application of the theorem to the stochastic game is that the space of strategies Π, (recall Definition 5.1) is a lattice but not complete in the order induced by $\succeq$ (see a statement of Tarski's theorem in Appendix 3). To complete the lattice we need to allow for the extended real variables $U = \pm \infty$, $U - \mu = \pm \infty$ (respectively $L = \pm \infty$, $L + \theta = \pm \infty$). Directly incorporating these into the space (and interpreting $U = \infty$, as "player 2 never acts," e.g.) does not work, since there is no best response to such a policy. Since player 1's payoffs are increasing in the state it is immediate that if he acts against a passive player he would like to increase the state arbitrarily. To get around this "problem at infinity," we artificially define a best response (to passive play) as the limit of best responses, as $U \uparrow \infty$. By Proposition 5.2, this limit is well–defined, and the "best response" map remains monotone. Hence Tarski's fixed point theorem can be applied. An additional argument is then required to show that the fixed point has not been created by the artificial extension of the map. We conclude by showing that the fixed point must be interior, i.e. that it is in fact an equilibrium in the stochastic game. The boundary arguments outlined above are stated precisely and proved in Appendix 3.

Theorem 5.4 **The stochastic game has an equilibrium, given by the two pairs $(L^*,\theta^*),(U^*,\mu^*)$ such that: $-\infty < L^* < L^* + \theta^* < U^* < \infty$, and $-\infty < L^* < U^* - \mu^* < U^* < \infty$. The equilibrium behavior is therefore described by a two–sided (s,S) rule: whenever $x \leq L^*$, player 1 acts and pushes the state upto $L^* + \theta^*$. Conversely, whenever $x \geq U^*$, player 2 lowers the state down to $U^* - \mu^*$.**

Proof: Define $\tilde{\Pi} \equiv \Pi \cup \{(U,U-\mu)\colon U = \pm \infty, U - \mu = \pm \infty\}$. $\tilde{\Pi}$ is clearly a complete lattice under $\succ$. Further, the best response is monotone on $\tilde{\Pi}$, by Proposition 5.2 and the extension of the map at the boundaries (see Appendix 3, equations (A.3.1)–(A.3.4) for details). Take $\hat{U}_0 = (\infty,\infty)$, $U_0 = (-\infty,-\infty)$ as the largest and smallest elements of the space in order to satisfy the boundary requirements of Tarski's Theorem (see Appendix 3 for a statement). So, there is a fixed point such that $\pi^* = B(\pi^*)$. We shall denote $\pi^* = (U^*,U^*-\mu^*)$ and $B_1(\pi^*) = L^*,L^*+\theta^*$ with the further convention that $B_{11}(\pi^*) = L^*$ and $B_{12}(\pi^*) = L^*+\theta^*$ (similarly, $B_{21}(B_1(\pi^*))$ and $B_{22}(B_1(\pi^*))$. Let us show that the fixed point is interior. It cannot be that $U^*=\infty$. Else, $B_{11}(\pi^*) \leq \bar{L} <\infty$ by Proposition A.3.1(i)–(ii) and hence $B_{21}(B_1(\pi^*)) \leq \bar{U}(\bar{L}) < \infty$, by the same proposition. Similarly, it cannot be that $U^* - \mu^* = -\infty$. Else, if $U^* > -\infty$, $B_{11}(\pi^*) \geq \underline{L}$ and $B_{12}(\pi^*) \in (\underline{L},U^*)$ and hence $B_{22}(B_1(\pi^*)) > -\infty$. If $U^* = -\infty$, then $B_1(\pi^*) = (-\infty,-\infty)$, but then $B_2(B_1(\pi^*)) \geq \bar{U} > -\infty$. So, the fixed point is interior. In turn, then π^*, $B_1(\pi^*)$ are best responses to each other, and hence form an equilibrium of the stochastic game. ■

5.3 The Case of a Compact Domain

The assumption that the state space of the stochastic process is the entire real line may be unsatisfactory for many readers. In many interesting economic examples there are natural restrictions on the range of values that the stochastic process can assume.[17] Here we prove that the above analysis extends (in fact, it becomes simpler) to the case in which $E = [\underline{E},\bar{E}]$, with $-\infty < \underline{E} < \bar{E} < +\infty$. In this formulation we have, of course, to specify the behavior of the process at the two boundary points. We shall consider both of the standard conditions:

(R) $(X_t)_{t\geq 0}$ is reflected at $\underline{E}$ and $\bar{E}$.

(A) $(X_t)_{t\geq 0}$ is absorbed at $\underline{E}$ and $\bar{E}$.

(See Karatzas and Shreve [1987] for a definition of reflected and absorbed processes.) We also have to specify what actions are feasible to ensure that players' jumps do not take the game environment outside the state space. We do this in the most obvious manner: we say that $X(\tau-) + \xi_1(\tau) + \xi_2(\tau) < \underline{E}$ (respectively $X(\tau-)+\xi_1(\tau)+ \xi_2(\tau) > \bar{E}$) implies $X(\tau) = \underline{E}$ (respectively $X(\tau) = \bar{E}$).

Lemma 4.1–4.4 and Corollaries 4.2–4.3 hold with no modifications so Theorem 4.0

holds. The best response is therefore well defined, and the best response to an (s,S) rule is an (s,S) rule. Similarly, the analysis of the monotonicity property of the best response is unchanged. In Proposition 5.1 if $L > \underline{E}$ the argument remains completely unchanged. On the other hand, when $L = \underline{E} = \hat{L}$ the conclusion is automatic. Finally, Proposition 5.2 holds without any modifications. We may conclude therefore:

Theorem 5.5 Let the state space be $[\underline{E},\overline{E}]$, and assume either (R) or (A). Then the stochastic game has an (s,S) equilibrium given by the two pairs (L^*,θ^*), (U^*,μ^*), with $\underline{E} \leq L^* \leq L^* + \theta^* \leq \overline{E}$, $\underline{E} \leq U^* \leq U^* - \mu^* \leq \overline{E}$.

Proof: Obviously Π is a complete lattice. Now Tarski's fixed point applies immediately. ∎

Remark The equilibrium of Theorems 5.4 and 5.5 need not of course be unique. We have investigated the following question: what is the range of possible behavior that can be found within the class of MPE. Suppose we restricted ourselves to Markovian policies which are described by a (closed) set of jumping points C and and a unique jumping destination x^* (in the strategies considered so far $C=(-\infty,L]$ etc.). It is our partially proven conjecture that any MPE then has the following property: C_i is a countable union of closed intervals which do not intersect and which alternate (i.e. an interval of player i follows every interval of j, except the terminal ones). The behavioral implication of such an equilibria is similar to that of the two–sided (s,S) equilibria with the modification that there are two relevant (L,U) intervals: one containing i's destination and the other j's. Eventually the state gets into either of these intervals, oscillates between them but never escapes from them.

5.4 An Application to Product Innovation

Recall from Section 2.4 the structure of the product innovation application. When firms use Markovian strategies the difference in qualities is the only relevant decision variable. Theorems 5.4 and 5.5 have the following observational implications: whenever the relative quality is too high or too low, one should observe an immediate innovation (to $L^* + \theta^*$ *or* $U^* - \mu^*$), and thereafter a process of alteration in technological leadership. The outcome contrasts sharply with the ϵ–preemption finding in some of of the patent race literature (see, for example, Fudenberg et al. [1983] or Harris–Vickers [1985]), in that no one firm can use an initial advantage to force the other out of the market. This difference reflects the fact that in our model there is no last period or final prize which the duopolists race towards. In a rough sense, one could think of our formulation as one with a succession of patent races each won by the firm which incurs the higher (fixed) costs. In particular, in this interpretation, technological laggards work harder in any given race.

As Theorem 5.5 makes clear, in some specifications it may be the case that only one of the players is active. So an initial technological leadership is maintained and improved over time. In this situation the only stimulus for innovations is disadvantageous changes in market demand for the single active firm.

6. The Cartel Solution

In this section we investigate the first-best or cartel problem for our stochastic game. The problem may be defined as follows: pick a set of (measurable) stopping times τ_i and an associated set of jumps to solve

$$\text{Max} \quad E_x \left\{ \delta \int_0^\infty [\alpha\, r(x) + (1-\alpha)\, r(-x)] e^{-\delta s} ds \; - c_0 \sum_{\tau_i} e^{-\delta \tau_i} \right\} \tag{14}$$

where $\alpha \in (0,1)$. To simplify the exposition we shall concentrate attention on the symmetric case, i.e. to the case $\alpha = {}^1/_2$. Without loss of generality consider the flow payoff to be $r(x) + r(-x)$ and denote this $\rho(x)$. So far all we have assumed about the flow payoffs r is that it is a monotone function. This does not of course have any implications for the sum ρ. We now make the further assumption:

(A6) r is a concave function.

Remark: It is immediate from (A6) that ρ is symmetric about $x=0$, concave and attains a unique maximum at $x=0$.

A natural conjecture for the optimal policy is that it will also be a two-sided (s,S) rule. Note that existing results that establish the optimality of (s,S) rules in single agent decision problems cannot be directly applied here. For instance, Caplin-Sheshinski (1988) only allows for decreasing stochastic processes whereas Harrison et.al (1983) impose an exogeneous one-sided barrier in the problem.

Even though the optimality of a two-sided (s,S) rule seems likely in the first-best problem, what is far from clear is the relationship that this rule will have to the similar rule that was obtained for the Markov Perfect equilibrium. For instance note that the cartel internalises all of the changes in returns due to a discrete change in the state. Part of any such change is the "cannibalization" effect, that the returns of one player increase if and only if the returns of the other player decrease. This might suggest that the cartel is less likely to use discrete actions. On the other hand, the cartel can better control all subsequent jumps and hence it may be profitable for it to use a discrete action when it is not for a single player who fears quick retaliation. We prove the following two results:

Theorem 6.0 **The first–best solution is given by a two–sided (s,S) rule (L_c, $L_c+\theta$, U_c, $U_c-\mu$) where $-\infty \leq L_c < L_c+\theta = 0 = U_c-\mu \leq \infty$ and is of the form: act iff $x \leq L_c$ or $x \geq U_c$ and then move the state to 0. Further, $|L_c| = U_c$.**

Note that since the game is symmetric there is always a symmetric equilibrium in the non–cooperative framework. It is clear that this symmetric equilibrium, denoted as (L_n ,$L_n+\theta$,U_n ,$U_n-\mu$) is the relevant point of comparison with the first–best solution. In such an equilibrium $|L_n| = U_n$ and so one point of comparison will involve the respective band sizes $2U_c$ versus $2U_n$. Unfortunately it is not true that symmetry of the equilibrium implies that the jump destinations are zero, i.e. that $L_n+\theta = U_n-\mu = 0$. But suppose there is in fact such a **strongly symmetric equilibrium**, i.e. an equilibrium in which such destinations are zero. Then we have:

Theorem 6.1 **Suppose that L_n defines a strongly symmetric MPE and L_c is the first–best solution. Then, $L_c < L_n$ (and so $U_c > U_n$).**

Proof of Theorem 6.0: The proof will follow by establishing that the value function for the first–best problem, denoted W, has the following property: W is continuous, symmetric and attains a unique maximum at x = 0. That W is continuous can be established by arguments identical to those employed in the proof of Theorem 4.0. Symmetry follows straightforwardly from the symmetry of r. Finally in order to show that W is in fact strictly increasing on $\mathbb{R}_-$ and strictly decreasing on $\mathbb{R}_+$, we appeal to an argument identical to that in Lemma 4.1. ∎

Proof of Theorem 6.1: Exactly as in Section 4, let us define the operator $\Gamma_\rho W$ as:

$$\Gamma_\rho W\ (x;a,b) = E_x\{\delta\int^{T_a\wedge T_b} e^{-\delta s}\rho(x_s)ds + e^{-\delta T_a}W(a)I(T_a\leq T_b) + e^{-\delta T_b}W(b)I(T_a>T_b)$$

Suppose per absurdum that in fact $L_c > L_n$. Then, for $x \in (\ L_c,0)$,

$$W(x) = \Gamma_\rho W\ (x;L_c+\epsilon,\epsilon) \quad \text{and} \quad W(x-\epsilon) = \Gamma_\rho W\ (x-\epsilon;L_c,0).$$

But from the spatial homogenity of Brownian motion it follows that the distribution of T_ϵ, with starting state x is identical to the distribution of T_0 with starting state $x-\epsilon$ (respectively the distribution of $T_{L_c+\epsilon}$ with starting state x is the same as that of T_{L_c} with initial state $x-\epsilon$.) So,

$$W(x) - W(x-\epsilon) = \alpha(x) - \alpha(x-\epsilon) + E_x e^{-\delta T_\epsilon}\,[W(\epsilon) - W(0)] + E_x e^{-\delta T_{L+\epsilon}}\,[W(L_c+\epsilon) - W(L_c)]$$

where $\alpha_\rho(x) = E_x\ \delta\int^{T_\epsilon\wedge T_{L+\epsilon}} e^{-\delta s}\rho(x_s)ds$ and $\alpha_\rho(x-\epsilon) = E_x\delta\int^{T_0\wedge T_L} e^{-\delta s}\rho(x_s)ds.$

Dividing by ϵ and taking limits it then follows from the differentiability of W and $W'(L_c) = W'(0) = 0$ that

$$W'(x) = \alpha_\rho'(x) \equiv E_x\delta\int e^{-\delta s}\,\rho'(x_s)ds$$

Similarly,

$$V'(x) = \alpha'_r(x) + E_x e^{-\delta T_L}\,V'(L_c) > \alpha'_r(x) \equiv E_x\ \delta\int e^{-\delta s}\,r'(x_s)ds$$

Since $\rho'(x) = r'(x) - r'(-x)$ it then follows that $\alpha'_r(x) > \alpha'_\rho(x)$. The above arguments have hence established that for all $x \in (L_c,0)$, $V'(x) > W'(x)$. But then,

$$c_0 \geq V(0) - V(L_c) > W(0) - W(L_c) = c_0$$

a contradiction. Hence the theorem is proved. ∎

7. Conclusions and Extensions

In this paper we analyzed a class of continuous time stochastic games in which players can make discontinuous changes in the variables that affect their payoffs. We placed three principal restrictions on the games studied. The first restriction was that only single–dimensional state formulations were studied. Although many economic applications do in fact analyze this case it is an important restriction for our study in that it appears difficult to dispense with. With a many dimensional state space even the decision theoretic problem is not fully understood. The challenge is two–fold: to find a sharp generalization of the concept of an (s,S) rule which generates a similar response and then to show that the monotonicity arguments employed in this paper can be appropriately extended.

A second restriction that was critical for the general analysis was the pure fixed costs assumption. In some special cases, as with linear payoffs, we can admit fixed plus proportional costs. Again the decision theoretic problem itself is not known to have a (s,S) characterization in this specification of costs.

The final restriction that facilitated the analysis is that we allowed players to only make discontinuous changes. A natural generalization of the framework of this paper would be to allow players to make complementary discrete and continuous changes. We have some preliminary results on this formulation in the linear payoffs case and hope to develop a more general theory.

Although we analyzed a model in which the exogeneous stochastic process is a Brownian motion, we feel that this specification is largely inessential. The examples reported in Section 3 seem to suggest that two–sided (s,S) rule equilibria are true more generally than in this specific case. We have not however explored this in any detail.

Given the above restrictions we showed the stochastic game has a Markov perfect equilibrium in two–sided (s,S) rules. We established this result by way of proving the optimality of (s,S) rules in a general and unsolved stochastic control problem and a monotonicity lemma on the best response map. We compared the equilibrium with the first–best solution which was shown to be also of a two–sided (s,S) rule type but with narrower bands. Finally, we introduced a framework for continuous time games which is flexible enough to accomodate instantaneous reactions by a player to his opponent's moves. Since standard formulations are often unable to do this, and hence result in inconsistencies, our framework may be of more general interest.

Footnotes

[1]One could attempt to make this distinction in discrete time by introducing gestation lags to the realization of some actions and calling such actions "discontinuous". The shortcomings of this approach are that i) the gestation lags are necessarily exogeneous and arbitrary, ii) the sharp distinction between the two types of actions are lost and iii) the analysis is necessarily considerably more inelegant than it is in a continuous time formulation of the problem.

[2]Notice that all the references in the previous paragraph study single–agent or decision theoretic problems only.

[3]See also Milgrom–Roberts (1989), for applications of fixed point theorems for monotone maps in supermodular games.

[4] Symmetry does not imply that the jump destination in equilibrium is 0. If that is the case we call the equilibrium strongly symmetric and the statement in the text refers to such an equilibrium.

[5]"The historic and irreversible change in the way of doing things we call 'innovation' and we define: innovations are changes in production functions which cannot be decomposed into infinitesimal steps. Add as many mail coaches as you please, you will never get a railroad by so doing" (Schumpeter [1935], p.7).

[6]A few representative contributions are Dasgupta–Stiglitz (1989), Lee–Wilde (1980), Fudenberg et al. (1983), Harris–Vickers (1985); see Reinganum (1989) for a survey of this area.

[7]By way of example, consider Schumpeter's railroads. Fishlow [1966] found in his study of the American railroads that at a time of significant cost reductions, the years between 1870 and 1910, the largest cost saving, by far, was due to a succession of improvements in the design of locomotives and freight trains. The process included no single major break with the past and yet "(its) lack of a single impressive innovation should not obscure its rapidity" (ibid, p. 35). Rosenberg (1982) makes a compelling case for the historical importance of continuous, ongoing change and also contains numerous other examples..

[8]Mansfield et al. [1981], found from their survey of 48 product innovations, of which 70% were patented, that about 60% were competed against within four years of the patent.

[9]Most of the discussion will in fact be for the non–compact domain case which is analytically the somewhat cleaner case. In many applications it may be more appropriate to have a compact domain. For this reason we develop the compact domain case fully as well.

[10]There are some subtle issues associated with the question: when is the sum in equation (1) well–defined? We discuss them in detail in Section 2.2.

[11]A3(i) is crucial solely in the proof of Proposition A.3. It is our belief that this assumption could be weakened to $\lim_{x\uparrow\infty} r'(x) = 0$. Of course, here and in A3(ii) we make use of the fact that the payoff function being monotonic and absolutely

continuous is differentiable a.e.. $r'(x)$ should be understood to refer to this a.e. derivative.

[12]In some examples, for instance the linear returns case (Example 3.3) we can allow the more general specification of fixed and proportional costs:
(A4)′ $c(\xi) = c_0 + c_1|\xi|, \quad c_1 > 0$

[13]For technical simplicity in the arguments that follow, we allow a firm to both improve as well as worsen its product quality, i.e., $\xi_i \in \mathbb{R}$. However, from hereon, we shall carry on the discussion as if $\xi > 0$.

[14]Dutta–Lach–Rustichini (1990) contains two examples, drawn from Kreps (1990)) and Shaked–Sutton (1983), in which the Cournot and Bertrand equilibrium profits of duopolists with differentiated products are solely a function of relative quality.

[15]The literature on repeat innovations and ongoing technological change in a strategic setting is very small. Ericson–Pakes [1989] contain some of the ideas that we discuss here, but in a perfectly competitive setting. To the extent that much innovative activity contains explicit incorporation of rivals' behavior, this generalization seems both economically important and of course technically much more complex.

[16]The proof of Bensoussan–Lions consists of two main steps. In the first, they formulate a Quasi–Variational Inequality (QVI) which is a general version of the system of equations A.2.1 – A.2.6 that we formulate in Appendix 2 for the special case of linear payoffs. This QVI is defined in (1.7) of chapter 4, p. 344. The connection of this QVI with the optimal control problem is explained in Section 1, Chapter 6, p. 615. The QVI has a regular unique solution as proved in the existence result of Theorem 1.1,chapter 4, p. 345.

The second step proves that the value function is equal to the solution of the QVI and that an optimal Markovian policy exists. This is done in Theorem 1.1 chapter 6, p. 619. The main difference between the problem they analyze and ours is in the boundary conditions. They assume a zero boundary condition on the value whereas in our case this is replaced with the condition $V(U) = V(U-\mu)$. Therefore in the first step one needs to replace the space of functions satisfing the zero boundary condition with the space of functions satisfying $V(U) = V(U-\mu)$. The arguments for the second step remain unchanged.

[17]For example, in the innovations context there is an alternative interpretation of the relative variable X. In this interpretation one could think of X as market share (and hence of course $x \in [0,1]$). By innovating, firms directly alter market share and derive profits from market share alone. The market share of firm 1 evolves through exogeneous changes in the market environment, the changes in Z, and through the innovation–induced changes ξ_1.

References

Bensoussan, A. and Lions, J. L., 1982, *Impulse Control and Quasi-Variational Inequalities*, Gauthier-Villars.

Caplin, A. and E. Sheskinski, 1988, "The Optimality of (s,S) Rules," mimeo, Princeton University.

Constantanides, G. and S.Richard, 1978,"Existence of Optimal Simple Policies for Discounted- Cost Inventory and Cash Management in Continuous Time," *Operations Research*, 26, 620-636.,

Dasgupta, P. and Stiglitz, J., 1980, "Uncertainty, Industrial Structure and the Speed of R&D," *Bell Journal of Economics*, 11, 1-28.

Dixit, A., 1989a, "A Simplified Expostion of the Theory of Optimal Control of Brownian Motion Control," mimeo, Princeton University.

Dixit, A., 1989b, " Entry and Exit Decisions Under Uncertainty," *Journal of Political Economy*," 97, 620-638.

Dutta, P. K. and Rustichini, A., 1990, "A Theory of Games of Entry with Applications to Product Innovations and Asset Sales," mimeo, Northwestern University.

Dutta, P. K., Lach, S. and Rustichini, A., 1990, "Innovation and Product Differentiation," mimeo, Northwestern University.

Ericson, R. and Pakes, A., 1989, "An Alternative Theory of Firm and Industry Dynamics," Columbia University Discussion Paper.

Fishlow, A., 1966, "Productivity and Technological Change in the Railroad Sector, 1840-1910,"in *Output, Employment and Productivity in the U.S. After 1800*, Studies in Income and Wealth, National Bureau of Economic Research, New York.

Fudenberg, D., Gilbert, R., Stiglitz, J., and J. Tirole, 1983, "Preemption, Leapfrogging and Competition in a Patent Race," *European Economic Review*, 24, 3-31.

Fudenberg, D. and J.Tirole, *A Course in Game Theory*, 1989, MIT Press, forthcoming.

Harris, C. and Vickers, J., 1985, "Perfect Equilibrium in a Model of Race," *Review of Economic Studies*, 52, 193-209.

Harrison, J. M., Sellke, T. M., and Taylor, A. J., 1983, "Impulse Control of a Brownian Motion," *Mathematics of Operations Research*, 8, 454-466.

Jovanovic ,B. and R.Rob, 1990, "Long Waves and Short Waves: Growth Through Intensive and Extensive Search," *Econometica*, 58, 1391-1410.

Karatzas, I. and S.Shreve, 1988, *Brownian Motion and Stochastic Calculus*, Springer-Verlag, Berlin.

Kreps, D., 1990, *A Course in Microeconomic Theory*, Princeton University Press, Princeton University.

Lee, T. and Wilde, L., 1980, "Market Structure and Innovation: A Reformulation," *Quarterly Journal of Economics*, 94, 429–436.

Mansfield, E., Schwartz, M., and Wagner, S., 1981, "Imitation Costs and Patents: An Empirical Study," *Economic Journal*, 91, 907–918.

Milgrom, P. and J.Roberts, 1990, "Rationalizability, Learning and Equilibrium in Games with Strategic Complementarities, *Econometrica*, 58, 1255 – 1278.

Parthasarathy, T., 1973, Discounted, Positive and Non–Cooperative Stochastic Games," *International Journal of Game Theory*.

Reinganum, J., 1989, The Timing of Innovation, in Schmalansee, R. and R.Willig eds. *Handbook of Industrial Organization*, North–Holland, New York.

Rosenberg, N.,1982, *Inside the Black Box: Technology and Economics*, Cambridge University Press, Cambridge, England.

Schumpeter, J., 1935, The Analysis of Economic Change, *Review of Economic Statistics*.

Stinchcombe, M., 1989, Maximal Strategy Sets for Continuous Time Game Theory, mimeo, University of California, San Diego.

Shaked, A. and Sutton, J., 1983, "Natural Oligopolies," *Econometrica*, 51, 1469–1484.

Tarski, A., 1955, A Lattice Theoretical Fixed Point Theorem and its Applications, *Pacific Journal of Mathematics*, 5, 285–309.

Appendix 1

In this section we provide a complete and detailed description of the extensive form of the game analyzed in the paper. We begin with a definition of time appropriate for our framework.

Our *time index set* T will have a special form. A time τ will be a pair $(\tau_1,\tau_2) \in \mathbb{R}_+ \times \mathbb{N}$. We give the index set the lexicographic order: for every pair τ and τ' in T, we say $\tau < \tau'$ if $\tau_1 < \tau_1'$ or $\tau_1 = \tau_1'$ and $\tau_2 < \tau_2'$. We shall always assume that T has such an order. T_τ will be the order interval $[0,\tau] = \{\theta \in T: \ 0 \leq \theta \leq \tau\}$.

A *history* at time τ is a map $h:[0,\tau] \to \mathbb{R}^3$, where for any $\theta \in [0,\tau]$, $h(\theta) = (x(\theta_1,\theta_2), \xi_1(\theta_1,\theta_2), \xi_2(\theta_1,\theta_2))$. Here $x(\theta_1,\theta_2)$ is the value of the state variable at (θ_1,θ_2) and $\xi_i(\theta_1,\theta_2)$ is the displacement of the state variable at the θ_2-stage of the instant game at θ_1, by the action of player i. The set of histories in $[0,\tau]$ is denoted by H_τ; and by H_∞ for the interval $[0,+\infty]$. Given the structure of the stochastic process we consider it natural to impose the following conditions on the set of histories:

1. $\theta_1 \to X(\theta_1)$ is a cadlag function (continuous from the right and with limit from the left).
2. $X(\tau_1) - X(\tau_1-) \; = \; \sum_i \xi_i(\theta) \qquad \theta_1 = \tau_1$, for every $\tau_1 \in \mathbb{R}_+$

(It is part of the definition of history (the assumption) that the series in 2 is convergent).

At any time τ the *action set* for every player is the real line; an action is the instantaneous displacement of the state variable by the specified amount. A *strategy* is a set of maps, indexed by time, from the set of histories to the action set; that is: σ: $T \times H_\infty \to \mathbb{R}$ where $\sigma(\tau,h)$ is the action at time τ for the history h; $\sigma(\tau,h) = 0$ is to be interpreted as "no action is taken". The *null strategy* (denoted by $\sigma^0(\tau,h)$) is $\sigma^0(\tau,h) \equiv 0$ for every τ and h. Of course, where facing a null strategy of the opponent, a player is in fact considering a simple control problem.

We now proceed to describe the restrictions on the set of strategies imposed by the information available to the players and by the stochastic process itself. Each player knows at every time τ the exact value of the state variable and the moves of the other player in the instant game, upto the stage being currently played.

To each set of histories H_τ , $\tau \leq +\infty$, we give a measurable space structure as follows. For any integer m, an m– cylinder $C = C(A,(\tau^1, ..., \tau^m))$, where A is a Borel subset of $\mathbb{R}^{3m}$, $\tau^i \in [0,\tau]$ is defined by:

$$C \equiv \{h \in H_\tau:\ (h(\tau^1), ..., h(\tau^m)) \in A\}$$

B_τ is defined as the smallest σ–field containing all the cylinders, of all finite dimensions. Notice that the family $\{B_\tau\}_{\tau\geq 0}$ is increasing (that is, $\tau_1 < \tau_2$ implies $B_{\tau_1} \subset B_{\tau_2}$). The index set T is given the topology induced by the lexicographic order. (A subbase for such topology is the set of intervals $[0,\tau]$ and $(\tau,+\infty)$). B(T) and $B(T_\tau)$ denote now the Borel σ–algebra on T and on T_τ respectively. (They are the same as the Borel σ–algebra generated by the product topology on $\mathbb{R}_+ \times N$). We can at this point define a notion of measurability for the strategies.

Definition: A map σ: T $\times$ H $\rightarrow \mathbb{R}$ is said to be progressively measurable relative to $\{B_\tau, \tau \in T\}$ *if and only if* the mapping $(\tau,h) \rightarrow \sigma\ (\tau,h)$ is $B(T_\tau) \times B(H_\tau)$ measurable for each $\tau > 0$.

This measurability condition on strategies requires the action taken at time τ to depend only on the history up to that time. It is therefore natural to assume the strategies to be progressively measurable. We need however, an additional notion before we can give a formal definition of admissible strategies.

For a pair of strategies (σ_1,σ_2) we denote A_τ (resp. B_τ) as the set of jumps of player 1 (resp. 2) at periods prior to $\tau = (t,0)$. Define

$$X(t,\sigma_1,\sigma_2) = Z(t) + \sum_{A_\tau} \xi_1(\tau) + \sum_{B_\tau} \xi_2(\tau) \qquad \text{(A.1.1)}$$

whenever the summations on the right are well defined.

A second natural condition will be imposed on strategies, in order to avoid trivial equilibria. We require (in intuitive terms) that each player uses strategies from a set which if he were alone would yield, almost surely, a well defined stochastic process. Formally we say

(S) An admissible strategy is a progressively measurable function σ: T x H → ℝ such that for every $t \geq 0$, the stochastic process $X(t,\sigma,\sigma^0)$ is well defined.

(A.1.1) above describes the outcome of the stochastic process. We now turn to payoffs: we let the payoff for player i, for strategy pair (σ_1,σ_2), be:

$$W_i\ (x,\sigma_1,\sigma_2) \ = \ \cdot E_x \left\{ \delta \int_0^{+\infty} r_i(x(t))\ e^{-\delta t}\ dt - \sum_\tau c(\xi_i(\tau))\ e^{-\delta\tau_1} \right\}$$

whenever the stochastic process t is well defined.

In order to complete the description of the game, we need to specify an outcome in the case in which the interaction of the two strategies does not yield a solution of the stochastic process. If the pair of (admissible) strategies does not give, with probability one, a well defined stochastic process, then

$$X(x,\sigma_1,\sigma_2) \ \equiv \ 0 \qquad \text{and}$$

$$W_i(x,\sigma_1,\sigma_2) \ = \ -\infty \ \textit{for}\ i = 1,2.$$

Remark It should be noticed that a pair of strategies satisfying the above condition (S) will not necessarily produce a well defined stochastic process, as is the case in the following example.

$$\begin{aligned} \sigma_1(\tau,h) \ &= \ 0 \text{ if } \xi_2(\tau_1,\tau_2) = 0, \\ &= 1 \text{ for every rational } t > \tau_1, \text{ if } \xi_2(\tau_1,\tau_2) \neq 0 \text{ for some } \tau_2 \\ \sigma_2(\rho,h) \ &= \ 1 \text{ at } \tau = (1,0), \text{ and } 0 \text{ otherwise.} \end{aligned}$$

On the other hand, the admissibility restriction rules out trivial equilibria, in which each player forces the other to a payoff of $-\infty$ because the strategy adopted by a player gives a non–well defined stochastic process. Indeed in the compact domain case, any equilibrium must result in a well–defined stochastic process with finite payoffs for both players.

Appendix 2

Proof of Lemma 3.1: Fix a strategy for player 2 given by a pair (U,μ). We construct a function V and a pair L and θ which satisfy:

$$\frac{\sigma^2}{2} V''(x) - \delta V(x) + \delta x = 0 \quad x \in (L,U) \tag{A.2.1}$$

$$V'(L) = 0 \tag{A.2.2}$$

$$V'(L+\theta) = 0 \tag{A.2.3}$$

$$V(U) = V(U-\mu) \text{ if } L \leq U-\mu \tag{A.2.4}$$

$$V(U) = V(L) \text{ if } L \geq U-\mu \tag{A.2.5}$$

$$V(x) = V(L+\theta) - c_0 \text{ for every } x \leq L \tag{A.2.6}$$

Pick a candidate strategy of player 1, say (L,θ). (A.2.1) is the differential equation which must be satisfied by the lifetime returns to any such strategy in the region in which no player moves (see e.g. Bensoussan–Lions (1982)). The boundary conditions (A.2.4) – (A.2.5) are similarly satisfied by such returns. Now (A.2.3) is the first order necessary condition which determines the optimal θ by determining the landing position $L + \theta$. Since the jumping position L is also chosen optimally we have the additional condition (A.2.2). Equation (A.2.6) gives the general form of the value function in the region where player 1 moves. Let $\lambda_\pm$ be the two roots of $\frac{\sigma^2}{2}x^2 - \delta = 0$

$$\lambda_\pm = \pm \frac{\sqrt{2\delta}}{\sigma} \tag{A.2.7}$$

For simplicity of exposition assume that $2\delta = \sigma^2$ so that $\lambda_+ = 1$. Further suppose $\delta = 1$. The solutions of equation (A.2.1) have the form:

$$g(x) = a_1 e^{-x} + a_2 e^{x} + x \tag{A.2.8}$$

where a_1 and a_2 are parameters to be determined. The four unknown parameters (a_1,a_2,L,θ) will now be determined by the equations (A.2.2) to (A.2.5). So the value function will in fact described by:

$$V(x) = \begin{cases} g(L+\theta) - c_0 & \text{if } x \leq L; \\ g(x) & \text{if } L \leq x \leq U; \\ g(U) & \text{if } x \geq U. \end{cases}$$

Equations (A.2.2) and (A.2.3) are solved for a_1 and a_2 to give

$$a_1 = \frac{e^L}{1 + e^{-\theta}}, \quad a_2 = -\frac{e^{-L}}{1+ e^{\theta}} \tag{A.2.9}$$

In equation (A.2.6) set x = L and substitute (A.2.9) above to get

$$-\left(\frac{1 - e^{\theta}}{1 + e^{\theta}} - \frac{1 - e^{-\theta}}{1 + e^{-\theta}} \right) + c_o - \theta = 0 \tag{A.2.10}$$

(A.2.10) determines a unique optimal θ, denoted $\hat{\theta}$, which depends only on c_o and not on (U,μ). Now we proceed to determine L.

We consider first the case $L \leq U - \mu$. Equations (A.2.4) and (A.2.9) now give

$$\left(\frac{1 - e^{-\mu}}{1 + e^{\theta}} e^{-(L-U)} + \frac{e^{\mu} - 1}{1 + e^{-\theta}} e^{L-U} \right) - \mu = 0$$

which is a simple quadratic equation in $e^{L-U} \equiv y$, with two roots $y_\pm(\mu)$:

$$y_\pm(\mu) = \{1 \pm [1 - 4 \frac{(1-e^{-\mu})(e^{\mu}-1)}{(1+e^{\theta})(1+e^{-\theta})}]^{1/2}\} \frac{1 - e^{-\theta}}{2(e^{\mu}-1)} \tag{A.2.11}$$

Since $L \leq U$, the root $y_+(\mu)$ is ruled out, and we find

$$L = U + \ln y_-(\mu), \text{ for } L \leq U - \mu. \tag{A.2.12}$$

Notice $\ln y_-(\mu) < 0$ for $\mu > 0$.

We now consider the case $L \geq U - \mu$. Equation (A.2.5) gives, setting $x = L - U$, the equation

$$\frac{e^x - 1}{1 + e^{-\theta}} - \frac{e^{-x} - 1}{1 + e^{\theta}} - x = 0 \tag{A.2.13}$$

which determines a unique negative x as a function of θ, with $x = 0$ if $\theta = 0$.

One can now check directly that as $L \to U+$, θ in equation (A.2.13) tends to 0; and as $\theta \to +\infty$, $y_-(\mu)$ in equation (A.2.11) tends to $+\infty$. With $\hat{\theta}$ fixed from (A.2.9), this determines a best $\hat{L}$. So the optimal jump is determined from (A.2.9) and the jumping point from either (A.2.12) or (A.2.13). So Lemma 3.1 is proved. ∎

Proof of Lemma 3.2: The best response map,constructed above, $(\hat{L},\hat{\theta})$ is monotonic (componentwise); that is, if $U \leq U'$, $\mu' \leq \mu$ then $L(U,\mu) \leq L(U',\mu')$ and $\theta(U,\mu) \leq \theta(U',\mu')$. The statement is trivial for $\hat{\theta}$, and for L as a function of U, for both the case $L \geq U - \mu$ and the case $L \leq U - \mu$. Also, L is a non constant function of μ only where $L \leq U - \mu$. We therefore need to prove our statement only in this case; equivalently, we need to prove that the function $x(\mu)$ is decreasing in μ. This follows immediately from the fact that the root y_- is decreasing in μ (because $\frac{\mu}{e^{\mu}-1}$ is decreasing, and $\frac{(e^{\mu}-1)(1-e^{-\mu})}{\mu^2}$ is increasing in μ) ∎.

Remark It may be interesting to point out that the optimal θ is increasing in c_0 (the fixed cost of every jump). The intuition is, of course, that once it becomes optimal to jump, then a higher fixed cost makes a larger size of the jump more convenient (because it makes jumps less frequent).

From Theorem 5.4 it follows that there is an equilibrium in this case. The band for the symmetric equilibrium can then be computed from (A.2.11).

Appendix 3

Recall that $B_1(U,U-\mu)$ is the best response of player 1 to a strategy pair (U,μ) for U and μ finite. We now extend this function to infinite values of U and μ. So we define:

$$B_1(\infty,\ U-\mu) \ = \lim_{U_\alpha\uparrow\infty} B_1(U_\alpha, U_\alpha - \mu_\alpha) \quad -\infty < U-\mu < \infty \tag{A.3.1}$$

$$B_1(\infty,\ \infty) \ = \lim_{U_\alpha-\mu_\alpha\uparrow\infty} B_1(\infty, U_\alpha - \mu_\alpha) \tag{A.3.2}$$

$$B_1(U\ ,-\infty) \ = \lim_{U_\alpha-\mu_\alpha\downarrow-\infty} B_1(U, U_\alpha - \mu_\alpha) \qquad -\infty < U < \infty \tag{A.3.3}$$

$$B_1(-\infty,-\infty) \ = \lim_{U_\alpha\downarrow-\infty} B_1(U_\alpha,-\infty)\ . \tag{A.3.4}$$

Given the monotonicity result of Proposition 5.2, all of the above limits are well-defined. Let us now establish a couple of useful properties of these limits. The proposition that follows establishes that the constructed best response of a boundary point, is interior. For $B_1(U,\mu) = (L,L+\theta)$, we write $B_{11} = L$, $B_{12} = L + \theta$.

Proposition A.3.1 (i) Suppose $U - \mu < \infty$. Then, there is $\bar{L} < \infty$, independent of $U - \mu$, such that $B_{11}\ (\infty,\ U - \mu) \leq \bar{L}$;

(ii) $B_{11}\ (\infty,\infty) \leq \bar{L}$;

(iii) Fix $U > -\infty$.. Then, there is $L\ (U) > -\infty$ such that
$B_{11}\ (U,-\infty) \geq L\ (U)$.

Proof: (i) Let $U_\alpha\uparrow\infty$, with $U - \mu$ fixed and finite. Let L_α,θ_α denote the best response, and suppose, per absurdum that L_α (and hence $L_\alpha + \theta_\alpha)\uparrow\infty$. Fix $\epsilon > c_0$ and $\psi < 1$ such that $\psi c_0 + \epsilon < c_0$. Pick $\hat{x} < \infty$ such that $r(\hat{x}) > \bar{r} - \epsilon$, where $\bar{r} = \lim_{x\uparrow\infty} r(x)$ which by A4 (i) is finite. Further, let L_α,θ_α be such that $\hat{x} < L_\alpha + \theta_\alpha$ and $E_{\hat{x}} e^{-\delta\hat{T}} < \psi$, where $\hat{T} = \inf\ \{t{:}X(t) = \hat{x};X(0) = \tilde{x} \in (\hat{x},L_\alpha + \theta_\alpha)\}$, under the policy:

"no jump till the first time $\hat{x}$ is hit, then jump to $L_\alpha + \theta_\alpha$ and follow the optimal policy thereafter." We have

$$V(\hat{x}) \geq E_{\hat{x}} \left\{ \delta \int_0^{\hat{T}} e^{-\delta S} r(x_s)\, ds + e^{-\delta \hat{T}} [V(L_\alpha + \theta_\alpha) - c_0] \right\}$$

$$> (1 - Ee^{-\delta \hat{T}})(\bar{r} - \epsilon) + Ee^{-\delta \hat{T}} V(L_\alpha + \theta_\alpha) - \psi c_0 \qquad \text{(A.3.5)}$$

$$> V(L_\alpha + \theta_\alpha) - \epsilon - \psi c_0$$

$$> V(L_\alpha + \theta_\alpha) - c_0$$

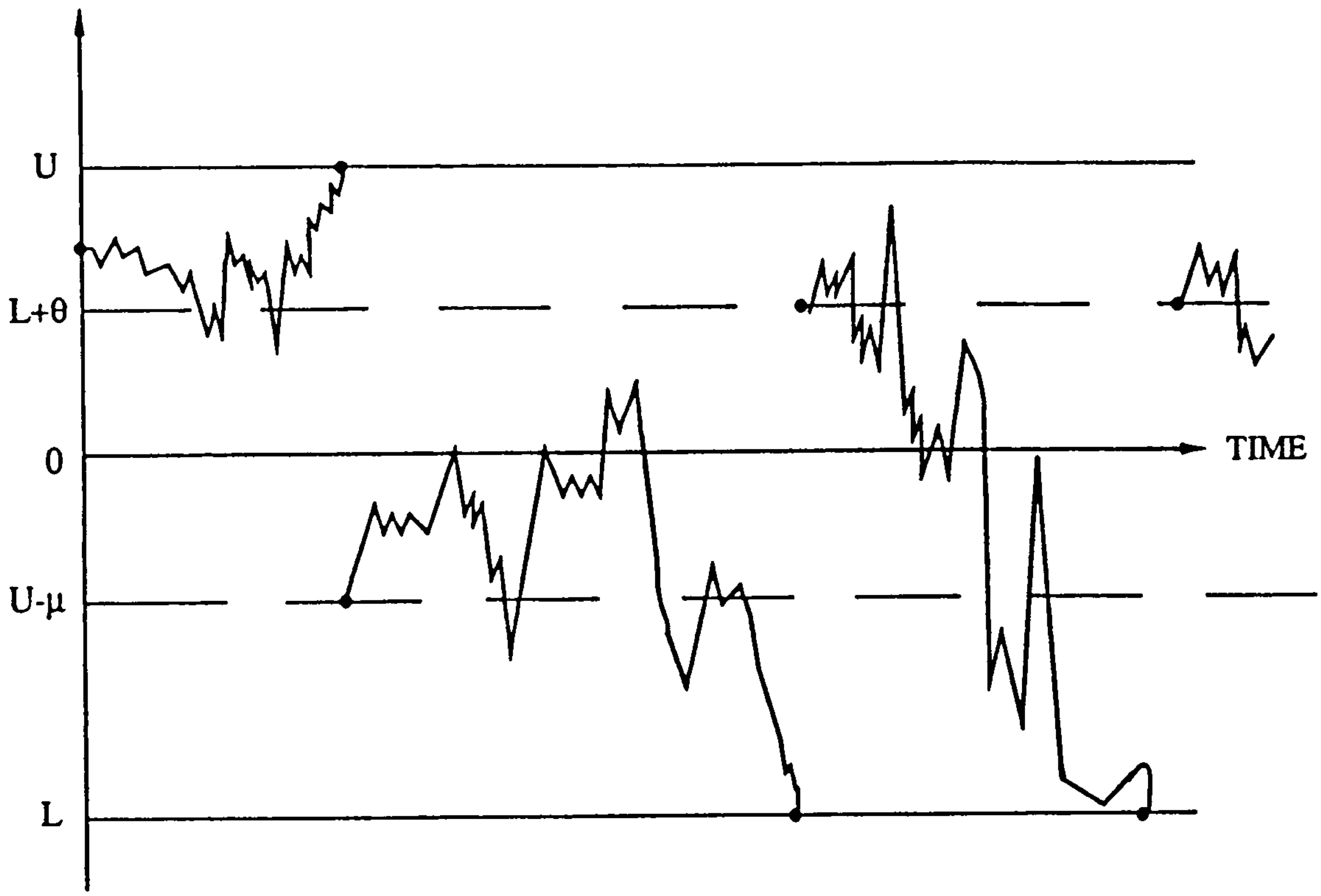
But (A.3.5) implies that not jumping is optimal at $\tilde{x}$, i.e., that $L_\alpha \leq \tilde{x}$. In turn, the argument was completely independent of $U - \mu$. So, from (A.3.1)–(A.3.2), (i) and (ii) follow.

Lemma 4.4 established that for fixed $(U, U-\mu)$, $L > -\infty$. In fact, the *same* argument establishes that this jumping point lower bound could be chosen, independently of $U - \mu$. That establishes (iii). ∎

Finally, we report here a statement of Tarski's fixed point theorem (Tarski (1955)). We say X is a lattice if it is an ordered set with the property that inf $\{x,y\}$ and sup $\{x,y\}$ exist for every x,y in X. We say that X is complete if sup (Y) and inf (Y) exist for every non–empty subset Y of X. A mapping $f{:}X \rightarrow X$ is said to be monotone increasing if $x \succeq y$ implies that $f(x) \succeq f(y)$. Then,

Tarski's Fixed Point Theorem : Every monotone increasing mapping on a complete lattice has a smallest and largest fixed point.

Figure 1



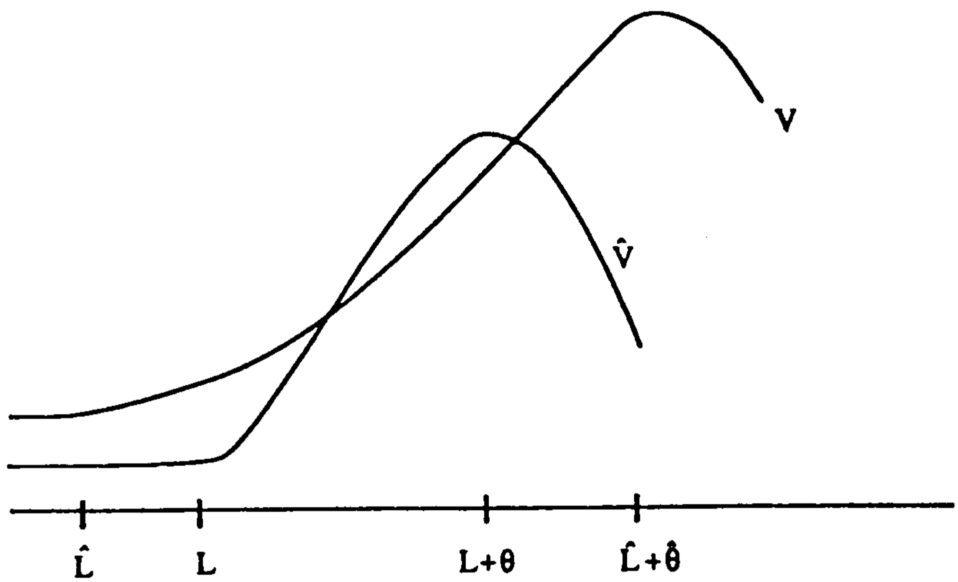


Figure 2