Technical Memorandum Number 785

Risk-Sensitive Markov-Perfect Equilibria

by

Madhvi Bhatt
Matthew Sobel

September 2004

Department of Operations
Weatherhead School of Management
Case Western Reserve University
330 Peter B Lewis Building
Cleveland, Ohio 44106

RISK-SENSITIVE MARKOV-PERFECT EQUILIBRIA

Madhvi S. Bhatt[1] and Matthew J. Sobel[2]

Abstract

Shubik (1959) initiates the literature on equilibrium point criteria in dynamic oligopoly models. Here, we investigate the existence and structure of equilibrium points when players are risk averse. We consider a sequential game in which each player maximizes the expected exponential utility of the present value of the time stream of rewards. Let X_t be the player's reward in period t, β the single-period discount factor, and $\lambda > 0$ parameterize risk sensitivity. We assume that the player wishes to maximize $E[-exp(-\lambda\Sigma_{t=1}^{\infty}\beta^{t-1}X_t)]$ and give sufficient conditions for the existence of a Markov-perfect equilibrium point. There are sets of sufficient conditions for a Markov-perfect equilibrium point to be myopic, namely a sequence of equilibrium points of static games. The myopia results are applied to a dynamic oligopoly model in which firms choose price and production quantities, encounter stochastic demand, and hold inventories.

Keywords

Sequential game, Markov perfect, myopic, risk, dynamic oligopoly, exponential utility.

[1] British Airways, London, U.K.
[2] Dept. of Operations, Weatherhead School of Management, Case Western Reserve University, 10900 Euclid Ave., Cleveland, OH 44106-7235.

August 2004

1. <u>Introduction</u>

Beginning with Shubik (1959), equilibrium point criteria in the literature on dynamic oligopoly, are risk-neutral; cf. Friedman (1977) and Fudenberg and Tirole (1986). More generally, the literature on sequential games focuses on risk-neutral criteria. Sequential games encompass models of many phenomena (Basar and Olsder, 1982). Starting with Shapley (1953), sequential games have been shown to possess equilibrium points under increasingly general sufficient conditions [Nowak (1992)] and there are numerous algorithms to compute them (Raghavan and Filar, 1991). Typically, each player is modeled as wishing to optimize the expected value of the sum of discounted rewards or the expected value of the limiting probability distribution of rewards.

Abundant data confirm that individuals are typically risk sensitive in diverse dynamic circumstances (Loewenstein et al., 2003). Also, risk-sensitive criteria are used in Shubik and Thompson (1959) and Sanghvi (1978) which study stochastic processes induced by competing decision makers. This paper establishes the existence and investigates the structure of equilibrium points of sequential games in which players are risk-sensitive.

Except for the risk-sensitive criterion, the model is the same as an N-person nonzero-sum sequential game with risk neutrality. The first half of the paper establishes the existence of Markov-perfect (*Markov* for short) equilibria when rewards are bounded and the sequential game is *finite*, i.e. there are finitely many states, actions, and players. See Pakes and McGuire (1994, 2001) and Maskin and Tirole (2001) for recent research on Markov equilibria. Equilibria exist among stationary strategies if a sequential game is risk-neutral (Fink, 1964), but one must seek equilibria of risk-sensitive sequential games in the larger set of Markov strategies.

Let $J = \{1, ..., N\}$ be the set of players and for player j let X_{tj} be the reward in period t ($t \in \mathcal{I}_+ = \{1, 2, ...\}$), $\beta_j \in$ (0,1) be the discount factor, and $B_j = \sum_{t=1}^{\infty} \beta_j^{t-1} X_{tj}$ be the present value of the rewards. We make two modeling choices that require jusification. First, player j

wishes to maximize $E[U_j(B_j)]$ in which $U_j(\cdot)$ is nonlinear. Second, $U_j(\cdot)$ is exponential, namely

$$E[U_j(B_j)] = E[-exp\,(-\lambda_j B_j)] = -E[\prod_{t=1}^{\infty} exp(-\lambda_j \beta_j^{t-1} X_{tj})] \qquad (1)$$

We call this the *risk-sensitive discounted criterion*.

An obvious alternative to representing preferences with an *inter*-period utility function $U_j(\cdot)$, as in (1), is to represent them with an *intra*-period utility function as in $E[\sum_{t=1}^{\infty} \Sigma_{t=1}^{\infty} \beta_j^{t-1} u_j(X_{tj})]$. However, weak sufficient conditions for the latter (Miyamoto and Wakker, 1996) essentially imply risk neutrality, *i.e.*, linearity of the intra-period utility function (Sobel, 2004). A nonlinear inter-period utility function is less restrictive than a nonlinear intra-period utility function.

An exponential utility function $U(x) = -e^{-\lambda x}$, $x \in \Re$, where $\lambda > 0$ parameterizes risk-sensitivity, induces a multiplicative reward structure (cf. (1)), and has desireable attributes. The coefficient of absolute risk aversion, $-U''/U'$, is constant if and only if U is linear or exponential. Among risk-averse increasing functions ($U'' < 0$ and $U' > 0$), only the exponential has a risk premium that is invariant with respect to wealth. (cf. Howard and Matheson (1972), Rothblum (1975), Jaquette (1976), Denardo and Rothblum (1979), Bell (1995), and Bell and Fishburn (2001)). Nowak (2004) analyzes the effects of an exponential utility function in static noncooperative games and two-stage stochastic games of resource extraction.

From a static perspective, it is apparent in (1) that the effective risk parameter in period t is $\lambda_j \beta_j^{t-1}$. Also, a Taylor series expansion of $exp(-\lambda_j \beta_j^{t-1} X_{tj})$ and $|\beta| < 1$ imply that the player is less risk sensitive as t grows. The time varying attitude towards risk can be made stationary, as usual, by enlarging the state space. In the original model, let S be the finite set of states with s_t the state in period t. In the model with an enlarged state space, the state in period t is $[s_t$; $(\beta_j^{t-1}\lambda_j, j \in J)]$.

The contraction mapping fixed point theorem plays an important role in theories of risk-neutral discounted dynamic programming and discounted sequential games (Bellman, 1957; Shapley, 1953; Blackwell, 1964; and Denardo, 1967). However, the mappings in risk-sensitive dynamic programs and sequential games are nonexpansive rather than contractions (Chung and Sobel, 1987 and Lemma 4). So the existence proof in §3 is more technical and relies on results in §2 and the Fan-Glicksberg extension of the Kakutani fixed point theorem (also employed by Nowak, 2004). Until §5 the sample spaces of X_{tj} (for all t and j) are embedded in a countable subset of a compact set. So without loss of generality, the rewards are discrete and bounded.

The second half of the paper presents conditions for risk-sensitive sequential games to have *myopic* equilibrium points. A particular dynamic oligopoly model is shown in §9 to satisfy the conditions. A myopic equilibrium point is a sequence of equilibrium points of static games. Myopia in a solution facilitates qualitative analysis and compution. Myopic solutions have been identified for many risk-neutral dynamic optimization problems and some dynamic oligopoly models (Arrow et al., 1958; Nerlove and Arrow, 1962; Veinott, 1965; Kirman and Sobel, 1974; and Sobel, 1990a).

§2 develops properties of dynamic programs that are related to the decision problem faced by one risk-sensitive player when the remaining players' decision rules are held constant. The properties in §2 yield the existence of Markov equilibria in §3. §4 presents an over-arching general condition for a risk-sensitive sequential game to have a myopic equilibrium point. §5 shows that the condition is satisfied when players' rewards are additively separable with respect to state and action, and the dynamics depend only on players' actions. §6 shows that the over-arching condition is satisfied if players' rewards and the expected dynamics are affine functions of the state with constant coefficients. The dynamic oligopoly model in §7 satisfies the conditions in §§5 and 6.

2. Dynamic Program with Exponential Utility

The results in this section augment Chung and Sobel (1987), hereafter (CS), for dynamic programs of a single decision maker with exponential inter-period utility. See White (1989) and Sobel (1994) for references to the literature on risk-sensitive dynamic programming.

Suppressing the notational dependence on a player's identity, let $X_1,\ X_2,\ \ldots$ be the sequence of single period rewards in a Markov decision process with finite state space S and finite set of feasible actions $A(s)$ when the state is $s \in S$. The present value of the sequence of rewards is $B = \sum_{t=1}^{\infty} \beta^{t-1} X_t$ and the decision maker maximizes $E[U(B)] = E(-e^{-\lambda B})$.

As a consequence of the nonlinear reward structure it is insufficient to formulate the optimization problem with separate marginal distributions, each conditional on the current state and action, of the current reward and the subsequent state as is usually done with Markov decision processes (CS). Thus, the data here are the same as in the corresponding risk-neutral model augmented by the joint distributions and the value of the risk parameter λ.

Let s_t and a_t be the respective state and action in period $t \in \mathbb{I}_+$; let the conditional distribution of X_t, given $s_t = s$ and $a_t = a$, take values in a countable subset K of a compact set (for all t, s and a); and let

$$p_{sjk}^{a} = P\{s_{t+1} = j, X_t = k \mid s_t = s, a_t = a\}$$

Let $\lambda > 0$, $\zeta = S \times (0, \lambda]$, and let $s_1 = s$ be the initial state.

This sequential decision process leads to a dynamic programming-like functional equation with a value function on ζ defined by

$$\nu(s,\theta) = sup\,\{E(-e^{-\theta B} \mid s_1 = s)\} \qquad (s,\theta) \in \zeta \qquad (2)$$

The supremum is over all policies, i.e. nonanticipative decision rules for choosing the sequence of actions $a_1, a_2, \ldots$ From (CS),

$$\nu(s,\theta) = max\,\{-E\,[\,e^{-\theta X_1}\,\nu(s_2, \beta\theta) \mid s_1 = s, a_1 = a]\colon a \in A(s))\}\} \qquad (3)$$

$$= max\,\{\textstyle\sum_{j \varepsilon S} \sum_{k \varepsilon K} p_{sjk}^{a}\, e^{-\theta k}\, \nu(j, \beta\theta)\colon a \in A(s)\}$$

Let $B(t) = \sum_{j=1}^{t} \beta^{j-1} X_j$ with B continuing to denote $B(\infty)$. For $(s,\theta) \in \zeta$ let $v_0(\cdot,\cdot) \equiv -1$ and define

$$v_t(s,\theta) = sup\,\{E(-e^{-\theta B(t)} \mid s_1 = s)\} \qquad t \in \mathcal{I}_+$$

(CS) show that

$$v_t(s,\theta) = max\,\{E[e^{-\theta X_1} v_{t-1}(s_2,\beta\theta) \mid s_1 = s,\, a_1 = a a\text{: } a \in A(s)\}$$

$$= max\,\{\textstyle\sum_{j\varepsilon S}\sum_{k\varepsilon K}\; p^a_{sjk}\, e^{-\lambda k}\, v_{t-1}(j,\beta\theta)\text{: } a \in A(s)\} \qquad (4)$$

LEMMA 1. For each $(s,\theta) \in \zeta$, $\nu(s,\cdot)$ is nondecreasing on (0,λ] and $v_t(s,\theta) \le v_{t+1}(s,\theta)$ $(t \in \mathcal{I}_+)$.

PROOF. Fix $\theta > 0$. Since $B(t) \ge 0$ for all t, $-1 \le v_t(s,\theta) \le 0$ for all s, t, and θ. Therefore, $-\boldsymbol{e} \le v_0(s,\theta) \le v_1(s,\theta)$ where $\boldsymbol{e}$ denotes the S-vector with all components one. Induction yields $v_t(s,\theta) \le v_{t+1}(s,\theta)$. □

In order to make explicit the dependence of the value function on $\{p^a_{sjk}\}$, let $p = (p^a_{sjk}$, $s \in S$, $j \in S$, $a \in A(s)$, $k \in K$), Q = {p: $p \ge 0$ and $\sum_{j\varepsilon S}\sum_{k\varepsilon K} p^a_{sjk} = 1$, for all $s \in S$ and $a \in A(s)$ }, and write $\nu(s,\lambda,p)$ for the dependence of the value function on $p \in Q$.

Let ρ be the following metric on Q: $\rho(p,r) = \sum_{s,a,j,k} \mid p^a_{sjk}$ - $r^a_{sjk} \mid$ ($p \in Q$, $r \in Q$). Let d be the supremum metric on the set of bounded functions from ζ to $\Re$.

LEMMA 2. For each $(s,\theta) \in \zeta$, $\nu(s,\theta,\cdot)$ is continuous on Q with the metrics ρ and d.

PROOF. A proof that $v_t(s,\theta,\cdot)$ is continuous on Q (for each t) begins with $v_0 \equiv -1$. Take $\epsilon > 0$, fix t, let $\delta = \epsilon//t$, and let $b_k = d[v_k(\cdot,\cdot,p),v_k(\cdot,\cdot,p')]$ for $k = 1,...,t$. In order to show that $[v_k(s,\theta,p) - v_k(s,\theta,p')] < \epsilon k/t$ for $(s,\theta) \in \zeta$, and $k = 1, ..., t$,

$$v_k(s,\theta,p) = max\,\{E[e^{-\theta X_1}\, v_{k-1}(s_2,\beta\theta,q) \mid s_1 = s,\, a_1 = a]\text{: } a \in A(s)\}$$

$$= max\,\{\textstyle\sum_{j\varepsilon S}\sum_{k\varepsilon K}\; p^a_{sjk}\, e^{-\theta k}\; v_{k-1}(j,\beta\theta,p)\text{: } a \in A(s)\}$$

$$\le max\,\{\textstyle\sum_{j\varepsilon S}\sum_{k\varepsilon K}\; p^a_{sjk}\, e^{-\theta k}\, [v_{k-1}(j,\beta\theta,p') + b_{k-1}]\text{: } a \in A(s)\}$$

$$\le max\,\{\textstyle\sum_{j\varepsilon S}\sum_{k\varepsilon K}\; p^a_{sjk}\, e^{-\theta k}\, v_{k-1}(j,\beta\theta,p')\text{: } a \in A(s)\} + b_{k-1}$$

$$\leq \ max \ \{\sum_{j\varepsilon S}\sum_{k\varepsilon K} \ [p_{sjk}^{a}\ ' - | \ p_{sjk}^{a}\ ' - p_{sjk}^{a} \ |] \ e^{-\theta k}[v_{k-1}(j,\beta\theta,p')] :$$

$$a \in A(s)\} + b_{k-1}$$

$$\leq \ v_k(s,\theta,p') + b_{k-1} + \ max \ \{\sum_{j\varepsilon S}\sum_{k\varepsilon K}| \ p_{sjk}^{a} \ - \ p_{sjk}^{a}\ ' \ | \ e^{-\theta k}$$

$$\cdot \ [-v_{k-1}(j,\beta\theta,p')] : a \in A(s)\}$$

But $\sum_{j\varepsilon S}\sum_{k\varepsilon K} e^{-\theta k} \in [0,1]$, $|v_{k-1}(j,\beta\theta,p')| \leq 1$, and $b_0 = 0$ imply

$$v_k(s,\theta,p) - v_k(s,\theta,p') \ \leq \ b_{k-1} + \ max \ \{\sum_{j\varepsilon S}\sum_{k\varepsilon K}| \ p_{sjk}^{a} \ - \ p_{sjk}^{a}\ '|: a \in A(s)\}$$

$$< b_{k-1} + \epsilon/t, \ k = 1, ..., t.$$

So $v_k(s,\theta,p) - v_k(s,\theta,p') < k\epsilon/t$, $k = 1,...,t$, with the corresponding lower bound obtained similarly. Therefore, $d[v_t(s,\theta,p),v_t(s,\theta,p')] \ < \ \epsilon$.

In order to prove that $v_t(s,\theta)$ converges uniformly to $\nu(s,\theta)$, (2) and (3) imply

$$\nu(s,\theta) = sup_{\Pi}\{E_{\Pi}[-exp\ (-\theta\sum_{k=1}^{t}\beta^{k-1}X_k)$$

$$\cdot \ exp(-\theta\beta^t\sum_{k=1}^{\infty}\beta^{k-1}X_{t+k})|s_1 = s]\}$$

Let $X_k \in [0,u]$ and $\mathfrak{m} = exp\ [-\theta\beta^t u/(1-\beta)]$; so $exp\ (-\theta\beta^t\sum_{k=1}^{\infty}\beta^{k-1}X_{t+k}) \in [\mathfrak{m},1]$ and $v_t(s,\theta) \leq \nu(s,\theta) \leq \mathfrak{m}v_t(s,\theta)$. Therefore, $v_t(s,\theta) \in [-1,0]$ and $\mathfrak{m} \leq 1$ imply

$$0 \leq \nu(s,\theta) - v_t(s,\theta) \leq \ -(1-\mathfrak{m})v_t(s,\theta) \leq (1-\mathfrak{m})$$

So, $\nu(s,\theta) - v_t(s,\theta) < \epsilon$ if $1 - \mathfrak{m} \leq \epsilon$ and

$$1 - \mathfrak{m} \leq \epsilon \ \iff 1 - exp[\theta u\beta^t/(1-\beta)] \leq \ \epsilon$$

$$\iff \beta^t \leq \ -[(1-\beta)/\theta u) \ log(1-\epsilon)]$$

whose right side depends on ϵ but is invariant on Q. □

3. Existence of a Markov Equilibrium Point

For each $j \in J$ and $s \in S$, let $A^j(s)$ be the nonempty set of feasible actions that player j can take when the game is in state s. Let a_{tj} be the action taken by player j in period t, $a_t = (a_{tj}$, $j \in J)$, and

$$p^a_{sjk} = P\{s_{t+1} = j, (X_{t1},...,X_{tN}) = k \,|\, s_t = s, a_t = a\}$$

Let $M(A^j(s))$ denote the set of all probability measures on $A^j(s)$ and $\boldsymbol{\lambda} = (\lambda_j, j \in J)$. A (randomized) *policy* for player j is a sequence $\Pi_j = (\Pi_{1j}, \Pi_{2j}, ...)$ in which, for each $t \in \mathcal{I}_+$, $\Pi_{tj}(h_t)$ is a mapping that takes a value in $M(A^j(s_t))$, i.e. a probability distribution on $A^j(s_t)$, for each partial history $h_t = [(\lambda_j, j \in J), s_1,a_1,s_2,a_2,....,s_{t-1},a_{t-1},s_t]$. A *strategy* $\Pi = (\Pi_j, j \in J)$ specifies a policy for each player. Then $\Delta_j = \times_{t=1}^{\infty} \times_{s\varepsilon S} M(A^j(s))$ is the set of randomized stationary (with respect to ζ) or randomized Markov (with respect to S) policies for player j and $\Delta = \times_{j=1}^{N} \Delta_j$ is the set of stationary (with respect to ζ) strategies. As usual, (ψ_j, Π_{-j}) denotes $(\Pi_1,\Pi_2,...,\Pi_{j-1},\psi_j,\Pi_{j+1},...,\Pi_N)$ where $\Pi = (\Pi_1, \Pi_2, ..., \Pi_j,...., \Pi_N)$.

Let $V_j(s,\lambda,\Pi) = E_\Pi(-e^{-\lambda_j B_j} \,|\, s_1 = s)$. A strategy $\Pi^* = (\Pi^*_j \,;\, j \in J)$ is an *equilibrium point* if

$$V_j(s, \theta, \Pi^*) \geq V_j(s, \theta, \psi_j, \Pi^*_{-j}) \qquad \text{for all } (s,\theta) \in Z, \psi_j, j \in J \tag{5}$$

With this definition and the model's information structure, an equilibrium point is necessarily subgame perfect. An equilibrium point is said to be *Markov-perfect* if it is subgame perfect and consists of Markov strategies (Maskin and Tirole, 2001).

Theorem 1 in this section states that an equilibrium point exists that is stationary with respect to the expanded state space. That is, player j in period t uses a (possibly) randomized decision rule that may depend on s_t and $(\lambda_j \beta_j^{t-1}, j \in J)$ but does not depend otherwise on t or the elapsed history. Such a decision rule is Markov with respect to the original state space.

The set of player j's Markov policies is $A^j = \times_{t=1}^{\infty} \times_{s\in S} A^j(s)$. A randomized action for player j in state $(s,\theta) \in \zeta$ corresponds to an element of $\times_{s\varepsilon S} M(A(^j(s))$. Let $\#Z$ denote the cardinality of set Z, let $m = max\ \{\#A^j(s): s \in S, j \in J\}\}$, and let F be the set of sequences of m numbers, each bounded above by one. Since $M(A^j(s)) \subset F$ for all j and s, the argument below uses $\times_{j=1}^{N} \times_{t=1}^{\infty} \times_{s\varepsilon S} M(A^j(s)) \subset \times_{j=1}^{N} \times_{t=1}^{\infty} \times_{s\varepsilon S} F$.

With the topology of weak convergence, $\times_{j=1}^{N} \times_{t=1}^{\infty} \times_{s\varepsilon S} M(A^j(s))$ is a compact convex topological subspace of $\times_{j=1}^{N} \times_{t=1}^{\infty} \times_{s\varepsilon S} F$ which is a linear Hausdorff locally convex topological space. The Fan-Glicksberg extension of the Kakutani fixed point theorem [Glicksberg (1952)] is applied by constructing a point-to-set mapping on $\Delta \subset \times_{j=1}^{N} \times_{t=1}^{\infty} \times_{s\varepsilon S} F$. The proof is less technical than in Federgruen (1978) because the set of actions here is finite rather than a compact metric space.

THEOREM 1. A finite sequential game with rewards in a countable subset of a compact set and the risk-sensitive discounted criterion has an equilibrium point $\delta^* \in \Delta$.

PROOF. For $\mu \in \times_{t=1}^{\infty} M(A^j(s))$ let $p_{sjk}(\mu) = E_\mu(p_{sjk}^{\boldsymbol{a}})$ be the “expected" joint conditional transition and reward probability when the players use the product measure μ to choose their actions in state s. For any $j \in J$ and $\delta \varepsilon \Delta$, let $\phi^j(\delta)$ be the set of player j's best responses. So the elements of $\phi^j(\delta) \subset \Delta_j$ satisfy

$$V_j(s,\theta,\delta) = \ sup\ \{\textstyle\sum_{i,k} e^{-\theta k_j}\ p_{sjk}(\delta_{-j},\mu) V_j(i, \beta \circ \theta, \delta) : \quad \mu \in \times_{t=1}^{\infty} M(A^j(s))\} \tag{6}$$

where $k = (k_j, i \in J)$, $\theta = (\theta_i, i \in \mathrm{J})$, and $\beta \circ \theta := (\beta_i \theta_i, i \in J)$.

The remainder of the theorem's proof consists of four observations.

LEMMA 3. ϕ: $\Delta \to 2^{\Delta}$ is a point to convex set mapping of a convex compact subset Δ of the linear Hausdorff locally convex topological space $\times_{j=1}^{\mathrm{N}} \times_{t=1}^{\infty} \times_{s\varepsilon S} F$ into itself.

PROOF. The space of functions that map states (s,θ) into N-tuples of measures is $\times_{j=1}^{N} \times_{t=1}^{\infty} \times_{s\varepsilon S} F$ which is a linear Hausdorff locally convex topological space when endowed with the product topology [Federgruen (1978) and Parthsarathy (1967)]. Since $A^j(s)$ is finite for each $s \in S$, $\times_{s\varepsilon S} M(A^j(s))$ is compact and convex. Similarly, Tychonoff's theorem yields compactness and convexity of $\times_{t=1}^{\infty} \times_{s\varepsilon S} M(A^j(s))$. Thus Δ is a metrizable subspace which is compact and convex. ☐

The second observation is that $\sum_{i,k} e^{-\theta_j k_j} p_{sik}(\,\cdot\,)$ takes values in $[0,1]$ and is continuous on $\times_{j \in J} M(A^j(s))$. Continuity is a consequence of uniformly bounded rewards and continuity of convex combinations of every bounded denumerable set of real numbers on the simplex of weights assigned to the numbers.

The third observation is that Lemma 2 implies that $V_j(s, \theta, \,\cdot\,)$ is continuous on Δ for $j \in J$ and $(s,\theta) \in Z$.

The next property completes the theorem's proof via the Fan-Glicksberg extension because there exists $\delta^* \in \Delta$ such that $\delta^* \in \phi(\delta^*)$.

LEMMA 4. ϕ is upper semi-continuous and nonexpansive.

PROOF. Fix $\{\delta_t\}_{t=1}^{\infty}$ with $\delta_t \in \Delta$ and $lim_{\,t \to \infty}\, \delta_t = \delta \in \Delta$. It follows that it is possible to select $\eta_t \in \phi(\delta_t)$ for each t such that there exists $\eta \in \Delta$ with $\eta = lim_{t \to \infty} \eta_t$. For each t, η_t^j is a best response for player j against $(\delta_t)_{-j}$. In the optimality equation (6), substitute δ_t for δ and η_t^j for η and let $t \to \infty$. It follows that η_j (where $\eta = (\eta_j, j \in J)$) satisfies the optimality equation for δ ($j \in J$, $s \in S$). It follows from (CS) that ϕ is nonexpansive. □

Existence of Equilibria under other Assumptions

If rewards take a continuum of values or take values in a countable set, then the results remain valid if rewards are uniformly bounded and measurable. Since Theorem 1 already assumes boundedness, the additional restriction is measurability.

A version of Theorem 1 can be established if S is countable, $A^j(s)$ is compact for each j and s, and rewards are uniformly bounded. Define topologies on the states and actions as in Federgruen (1978) and make corresponding continuity assumptions. It follows that there exists an equilibrium point in Δ.

4. <u>Myopic Equilibrium Points</u>

The remainder of the paper concerns the structure of Markov equilibria rather than their existence. So we no longer oblige the sequential game model to be finite. Let $A_s = \times_{j \in J} A^j(s)$, $C = \{(s, a)\colon a \in A_s,\, s \in S\}$, $A^j = \bigcup_{s \in S} A^j(s)$, $A = \times_{j \in J} A^j$, and $S(a) = \{s\colon (s, a) \in C\}$. The following definition weakens the definition of an equilibrium point (5) to a subset of initial states. A strategy $\Pi^* = (\Pi^*_j\colon j \in J)$ is an *equilibrium point with respect to* $T \subseteq S$ if

$$V_j(s, \theta, \Pi^*) \geq V_j(s, \theta, (\mu_j, \Pi^*_{-j})) \qquad \text{for all } \mu_j,\, j \in J,\, s \in T,\, 0 < \theta \leq \lambda \tag{8}$$

This section assumes that there are independent random variables ξ_1, ξ_2, ... such that ξ_t and $(s_1, a_1, ..., s_t, a_t)$ are independent. In §7, ξ_t is the exogeneous “raw" demand in period t. In other potential applications, ξ_t is the change in biomass in aquaculture (Sobel, 1990a) and the goodwill depreciation rate in an oligopoly model with advertising decisions (Sobel, 1990b). Let Θ_t denote the sample space of ξ_t.

Suppose there exist $\gamma_{tj}\colon A \to \Re$ ($t \in \mathbb{I}_+$, $j \in J$) and $m\colon Z \to \Re$ such that

$$V_j(s, \lambda, \delta) = m(s, \lambda) \bullet E_{\delta | s_1 = s}\{\textstyle\prod_{t=1}^{\infty} v_{tj}(a_t)\} \tag{9}$$

$$v_{tj}(a) = E\{exp\,[-\lambda\beta^{t-1}\gamma_{tj}(a, \xi_t)]\} \qquad a \in A \tag{10}$$

Note that $m(\,\cdot\,,\,\cdot\,)$ does not depend on δ. Let Γ_t denote the one-period (static) game among the players in J in which player j has available the set of moves $A^j = \bigcup_{s \in S} A^j(s)$ and incurs the *cost* $v_{tj}(a)$ when the players choose $a \in A$. First we suppose for each t that Γ_t has a pure strategy equilibrium point $a^*_t \in A$:

$$v_{tj}(a^*_t) \;\leq\; v_{tj}(k, a^*_{-j}) \qquad (k \in A^j,\; j \in J,\, t \in \mathbb{I}_+) \tag{11}$$

and that $\{a^*_t\}$ is *repeatable* in the following sense:

$$P\{\{s_{t+1} \in S(a^*_{t+1}) \mid s_t \in S(a^*_t),\, a_t = a^*_t\} = 1 \qquad (t \in \mathbb{I}_+) \tag{12}$$

If $a_t = a^*_t$ then $a_{t+1} = a^*_{t+1}$ is feasible (with probability one).

Let α be a strategy that specifies $a_t = a^*_t$ if $s_t \in S(a^*_t)$ [$\Leftrightarrow a^*_t \in \times_{j \in J} A^j(s)$] and specifies an arbitrary element of $\times_{i \in J} A^j(s)$ if $s_t \notin S(a^*_t)$. Let $\nu_j(s, \lambda_j, \delta) = E_{\delta | s_1 = s}[\prod_{t=1}^{\infty} v_{tj}(a_t)]$.

THEOREM 2. Under assumptions (9), (11), and (12), α is an equilibrium point with respect to $S(a_1^*)$.

PROOF. If $s_1 = s \in S(a_1^*)$, the players except j use their components of α, and j uses μ_j, then by (9)

$$\begin{aligned}
\nu_j[s,\lambda_j,(\mu_j,\alpha_{-j})] &= E\{\textstyle\prod_{t=1}^{\infty} exp\,[-\lambda_j\beta_j^{t-1}\gamma_{tj}(a_t,\xi_t)]\} \\
&= E_{(\mu_j,\alpha_{-j})}E_{\xi|\mu_j,\alpha_{-j}}\{\textstyle\prod_{t=1}^{\infty} exp\,[-\lambda_j\beta_j^{t-1}\gamma_{tj}(a_t,\xi_t)]\} \\
&= E_{(\mu_j,\alpha_{-j})}\textstyle\prod_{t=1}^{\infty} E_{\xi}\,\{exp\,[-\lambda_j\beta_j^{t-1}\gamma_{tj}(a_t,\xi_t)]\} \\
&= E_{(\mu_j,\alpha_{-j})}\textstyle\prod_{t=1}^{\infty} v_{tj}\,(a_t^j,a_{-j}^*) \\
&\geq E_{(\mu_j,\alpha_{-j})}\textstyle\prod_{t=1}^{\infty} v_{tj}\,(a_t^*) = \nu_j(s,\lambda_j,\alpha)
\end{aligned}$$

If $s_1 \in S(a_1^*)$, then (12) implies $a_t \in A_{s_t}$ for all t. □

Now we abandon the assumption that Γ_t has a pure strategy equilibrium point $a_t^* \in A$ for each t. If A is finite, from Nash (1950) for each t there is a randomized equilibrium point (possibly with a degenerate probability distribution) $\mathbf{a}_t^*$ of Γ_t, $t \in \mathcal{I}_+$. Since finiteness is no longer assumed, define $A(t) = \{a \in A\colon\ P\{\ \boldsymbol{a}_t^* = a\} > 0\}$ and generalize (11) and (12):

$$\Gamma_t \text{ has a (possibly randomized) equilibrium point } \boldsymbol{a}_t^* \qquad (t \in \mathcal{I}_+) \tag{11'}$$

$$\mathrm{P}\{s_{t+1} \in \textstyle\bigcap_{a\in A(t+1)} S(a) \mid \boldsymbol{a}_t^* = a_t\} = 1 \qquad t \in \mathcal{I}_+ \tag{12'}$$

Let η be a strategy that specifies $\mathrm{P}\{a_t = a\} = P\{\boldsymbol{a}_t^* = a\}$ if $s_t \in \bigcap_{a\in A(t)} S(a)$ and η specifies an arbitrary element of A_{s_t} if $s_t \notin \bigcap_{a\in A(t)} S(a)$.

COROLLARY 1: Assumptions (9), (11'), and (12') imply that η is an equilibrium point with respect to $\bigcap_{a\in A_{s_1}} S(a)$.

5. Separable Rewards and Dynamics Depending only on Actions

This section shows that condition (9) is satisfied if the single period reward function is an additively separable function of the state and action, and the dynamics depend only on the actions. These assumptions yield myopic equilibria in a risk-neutral sequential game (Sobel, 1981). For each $t \in \mathcal{I}_+$ and $j \in J$ suppose there are $K_{tj} : A \times \Theta_t \ \to \mathbb{R}$, $L_{tj} : S \to \mathbb{R}$ and $w_n : A \times \Theta_t \ \to \mathbb{R}$ such that

$$X_{tj} = K_{tj}(a_t, \xi_t) + L_{tj}(s_t) \tag{13}$$

$$s_{t+1} = w_t(a_t, \xi_t) \tag{14}$$

Let

$$\gamma_{tj}(a, x) = K_{tj}(a, x) + \beta L_{t+1,j}[w_t(a, x)[\qquad a \in A, x \in \Theta \tag{15}$$

An expression for $V_j(s, \lambda, \delta)$ analogous to (9) results from (13), (14), and (15).

LEMMA 5. The structure (13) and (14) implies

$$V_j(s, \lambda, \delta) = \ - exp[- \lambda L_{1j}(s)] \cdot E_{\delta|\mathrm{s}_1 = s}[\textstyle\prod_{\mathrm{t}=t=1}^{\infty} v_{tj}(a_n)] \tag{16}$$

PROOF. The substitution of (13) and (14) in

$$B_j = \textstyle\sum_{t=1}^{\infty} \beta^{t-1} X_{tj}$$

yields

$$\begin{aligned} \textstyle\sum_{t=1}^{\infty} \beta^{t-1} X_{tj} &= \textstyle\sum_{t=1}^{\infty} \beta^{t-1}[K_{tj}(a_t, \xi_t) + \ L_{tj}(s_t)] \\ &= L_{1j}(s) + \textstyle\sum_{t=1}^{\infty} \beta^{t-1}[K_{tj}((a_t, \xi_t) + \beta L_{t-1,j}[w_t(a_t, \xi_t)] \\ &= L_{1j}(s) + \ \textstyle\sum_{t=1}^{\infty} \beta^{t-1} \gamma_{tj}(a_t, \xi_t) \end{aligned}$$

Then

$$\begin{aligned} V_j(s, \lambda, \delta) &= E[- exp[\ - \lambda \textstyle\sum_{t=1}^{\infty} \beta^{t-1} X_{tj})] \\ &= E\{ - exp[- \lambda L_{1j}(s)] \cdot exp[- \lambda \textstyle\sum_{t=1}^{\infty} \beta^{t-1} \gamma_{tj}(a_t, \xi_t)]\} \\ &= \ - exp[- \lambda L_{1j}(s)] \cdot E\{exp[- \lambda \textstyle\sum_{t=1}^{\infty} \beta^{t-1} \gamma_{tj}(a_t, \xi_t)]\} \\ &= \ - exp\ [- \lambda L_{1j}(s)] \cdot E\{\textstyle\prod_{t=1}^{\infty} exp[- \lambda \beta^{t-1} \gamma_{tj})(a_t, \xi_t)] \end{aligned}$$

Let $$G_{tj}(a, x) = exp\ [[- \lambda \beta^{t-1} \gamma_{tj}(a, x)] \qquad a \in A, x \in \Theta \tag{17}$$

Independence of a_t and ξ_t yields

$$\begin{aligned} E\{\textstyle\prod_{t=1}^{\infty} G_{tj}(a_t, \xi_t)\} &= E_{\{a_j\}} E_{\{\xi_j | a_j\}} \{\textstyle\prod_{t=1}^{\infty} G_{tj}(a_t, \xi_t)\} \\ &= E_{\{a_i\}} E_{\{\xi_j\}} \{\textstyle\prod_{t=1}^{\infty} G_{tj}(a_t, \xi_t)\} \end{aligned}$$

Independence of $\xi_1, \xi_2, ...$ implies

$$E[\prod_{t=1}^{\infty} G_{tj}(a_t,\xi_t)] = E_{\{a_j\}}\{\prod_{t=1}^{\infty}[\, E_{\{\xi_j\}} G_{tj}(a_t,\xi_t)]\} = E_{\{a_j\}}[\prod_{t=1}^{\infty} v_{tj}(a_t)] \qquad \square$$

The first term in (16) is invariant with respect to strategy δ.

COROLLARY 2. A sequential game with (13) and (14) satisfies (9). So

(a) Assumptions (11), (12), (13), and (14) imply that α is an equilibrium point with respect to $S(a_1^*)$.

(b) Assumptions (11'), (12'), (13), and (14) imply that η is an equilibrium point with respect to $\bigcap_{a \in A(s_1)} S(a)$.

6. <u>Affine Structure</u>

This section shows that condition (9) is satisfied if the single period reward function and the expected dynamics are affine functions of the state with coefficients that are constant with respect to the action $a \in A$. The same assumptions yield myopic equilibria in a risk-neutral sequential game (Sobel, 1990a) and are similar to those in Denardo and Rothblum (1983). Monahan and Sobel (1995) apply the results in this section to a risk-sensitive oligopoly model with advertising decisions.

Let $\Re^{k \times n}$ denote the set of $k \times n$ matrices of real numbers. We assume that there are $L_j \in \mathbb{R}^{N\times 1}$, $M \in \mathbb{R}^{N\times N}$, and for each $a \in A$ there are $K_j(a,\xi_t) \in \mathbb{R}$ and $w(a,\xi_t) \in \mathbb{R}^{N\times 1}$ such that

$$X_{tj} = K_j(a_t,\xi_t) + L_j \cdot s_t \qquad (t \in \mathcal{I}_+,\, j \in J) \qquad (18)$$

$$s_{t+1} = w(a_t,\xi_t) + M \cdot s_t \qquad (t \in \mathcal{I}_+) \qquad (19)$$

For simplicity, assume that K_j, L_j, w, and M are stationary with respect to t.

Let $\tau(M)$ denote the spectral radius of matrix M and define

$$\mathcal{L}_j = [I + \beta_j M (I - \beta_j M)^{-1}] L_j$$

$$\gamma_j(a,x) = K_j(a,x) + \beta_j \mathcal{L}_j (I - \beta_j M)^{-1} w(a,x) \qquad (a \in A,\, x \in \Theta)$$

LEMMA 6. Assumptions (18), (19), and $\tau(M) < \beta_j^{-1}$ imply

$$\sum_{t=1}^{\infty} \beta_j^{t-1} X_{tj} = \mathcal{L}_j s_1 + \sum_{t=1}^{\infty} \beta_j^{t-1} \gamma_j(a_t, \xi_t) \quad (20)$$

$$V_j(s, \lambda_j, \delta) = -exp\,[-\lambda \mathcal{L}_j s] \cdot E_{\delta|s_1=s}\{\prod_{t=1}^{\infty} exp[-\lambda_j \beta_j^{t-1} \gamma_j(a_t, \xi_t)]\} \quad (21)$$

PROOF. $\sum_{t=1}^{\infty} \beta_j^{t-1} X_{tj} = \sum_{t=1}^{\infty} \beta_j^{t-1} [K_j(a_t, \xi_t) + L_j(s_t)]$

$$= L_j s_1 + \sum_{t=1}^{\infty} \beta_j^{t-1} K_j(a_t, \xi_t) + \beta_j \sum_{t=1}^{\infty} \beta_j^{t-1} L_j s_{t+1} \quad (22)$$

Let w_t denote $w(a_t, \xi_t)$. Then

$$s_{t+1} = w_t + M s_t$$

$$= w_t + M(w_{t-1} + M s_{t-1})$$

$$= w_t + M w_{t-1} + M^2 s_{t-1}$$

$$= M^t s_1 + \sum_{i=1}^{t} M^{t-i} w_i$$

Substituting in (22) confirms (20):

$$\sum_{t=1}^{\infty} \beta_j^{t-1} X_{tj} = L_j s_1 + \sum_{t=1}^{\infty} \beta_j^{t-1} K_j(a_t, \xi_t)$$

$$+ \beta_j \sum_{t=1}^{\infty} \beta_j^{t-1} L_j (M^t s_1 + \sum_{i=1}^{t} M^{t-i} w_i)$$

$$= [L_j + \beta_j L_j \sum_{t=1}^{\infty} \beta_j^{t-1} M^t] s_1$$

$$+ \sum_{t=1}^{\infty} \beta_j^{t-1} [K_j(a_t, \xi_t) + \beta_j L_j \sum_{i=1}^{t} M^{t-i} w_i]$$

$$= [L_j + \beta L_j M (I - \beta \mathrm{M} M)^{-1}] s_1$$

$$+ \sum_{t=1}^{\infty} \beta_j^{t-1} [K_j(a_t, \xi_t) + \beta L_j (I - \beta M)^{-1} w(a_t, \xi_t)]$$

$$V_j(s, \lambda, \delta) = E\{-exp[-\lambda_j \sum_{t=1}^{\infty} \beta_j^{t-1} X_{tj}]\}$$

$$= E\{-exp\Big(-\lambda_j [L_j s + \beta_j L_j M (I - \beta M)^{-1} s]\Big)$$

$$\cdot\, exp[-\lambda_j \sum_{t=1}^{\infty} \beta_j^{y-1} [K_i(a_t, \xi_t) + \beta_j L_j (I - \beta_j M)^{-1} w(a_t, \xi_t)]\}$$

$$= -exp\,(-\lambda_j \mathcal{L}_j s) \cdot E[exp - \lambda_j \sum_{t=1}^{\infty} \beta_j^{t-1} \gamma_j(a_t, \xi_t)] \qquad \square$$

<u>Higher-Order Models</u>

In some applications it is natural to formulate models in which dynamics and rewards depend directly on states and actions in earlier periods. Instead of (18) and (19), as in Sobel (1990b), we assume that there are $r \in \mathcal{I}_+$, functions K_{ij} and w_i, vectors L_{ij}, and matrices M_i such that

$$X_{tj} = \sum_{i=1}^{r}[K_{ij}(a_{t-i+1}, \xi_{t-i+1}) + L_{ij} \cdot s_{t-i+1}] \tag{18'}$$

$$s_{t+1} = \sum_{i=1}^{r}[w_i(a_{t-i+1},\ \xi_{t-i+1}) + M_i \cdot s_{t-i+1}] \tag{19'}$$

It is surprising that a higher-order model can be analyzed with essentially the same effort as when $r = 1$. Let prime denote transpose, $\omega_i(a,x)' = (w_i(a,x)', 0, ..., 0)$, and

$$\mathfrak{M} = \begin{bmatrix} M_1 & M_2 & \cdot & \cdot & M_{r-1} & M_r \\ I & 0 & \cdot & \cdot & 0 & 0 \\ 0 & I & & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & & \cdot \\ 0 & 0 & \cdot & \cdot & I & 0 \end{bmatrix}$$

LEMMA 7. Assumptions (18'), (19'), and $\tau(\sum_{i=1}^{r}\beta_j^{i-1}M_i) < 1$ imply

$$\sum_{t=1}^{\infty}\beta_j^{t-1}X_{tj} = \mathfrak{L}_j + \sum_{t=1}^{\infty}\beta_j^{t-1}\gamma_j(a_t, \xi_t) \tag{20'}$$

where $\mathfrak{L}_j$ is a function of s_{-r+2}, ..., s_1, a_{-r+2}, ..., a_0, ξ_{-r+2}, ..., ξ_0 and

$$\gamma_j(a,x) = \sum_{i=1}^{r}\beta_j^{i-1}[K_{ij}(a,x) + \beta_j L_j(I - \beta_j\mathfrak{M})^{-1}\omega_I(A,X)] \qquad (A \in a,\ X \in \Theta)$$

Also, $V_j(s, \lambda_j, \delta) = -exp[-\lambda_j\mathfrak{L}_j] \cdot E_{\delta|s_1=s}\{\prod_{t=1}^{\infty}exp[-\lambda_j\beta_j^{t-1}\gamma_j(a_t, \xi_t)]\}$ (21')

PROOF. Use (18') and (19') in a sample path version of the proof of Lemma 1 in Sobel (1990b). ☐

The first terms in (21) and (21'), respectively, are constant with respect to strategy δ and need not be considered in an optimization. Also, (21) and (21') satisfy (9) with $m(s,\lambda_j) = -exp(-\lambda_j\mathcal{L}_j s)$ and $m(s,\lambda_j) = -exp(-\lambda_j\mathfrak{L}_j)$, respectively, and (10) with γ_{tjj} invariant with respect to t.

COROLLARY 3. [(18) and (19)] and [(18') and (19')] each satisfy (9). So:

(a) Assumptions (11), (12), (18), and (19) [respectively (18') and (19')] imply α is an equilibrium point with respect to $S(a_1^*)$.

(b) Assumptions (11'), (12'), (18), and (19) [respectively (18') and (19')] imply that η is an equilibrium point with respect to $\bigcap_{a \in A(s_1)} S(a)$.

7. Dynamic Oligopoly with Inventories

The following oligopoly model is taken from Kirman and Sobel (1974). If $N = 1$ and dependence on price is suppressed, it is a dynamic version of the classical "newsvendor" model [Porteus (1990)]. At the beginning of each period t, N firms make simultaneous pricing and production decisions. Let z_{tj} and $\rho_{tj} \geq 0$ be firm j's production quantity and announced price, respectively, in period t. Let s_{tj} be firm j's inventory level at the onset of period t. Assume that production occurs in less than one period; so the total quantity of goods that is available to satisfy demand in period t is $y_{tj} = s_{tj} + z_{tj}$. Let $a_{tj} = (y_{tj}, \rho_{tj})$, let D_{tj} be the nonnegative demand for firm j's goods during period t, and let $s_t = (s_{t1}, ..., s_{tN})$, $z_t = (z_{t1}, ..., z_{tN})$, $\rho_t = (\rho_{t1}, ..., \rho_{tN})$, $y_t = (y_{t1}, ..., y_{tN})$, $a_t = (a_{t1}, ..., a_{tN})$, and $D_t = (D_{t1}, ..., D_{tN})$. So $y_t = s_t + z_t$ and $z_t \geq 0$ corresponds to $y_t \geq s_t$. Under the assumption that excess demand is backlogged,

$$s_{t+1} = s_t + z_t - D_t = y_t - D_t \tag{23}$$

If $s_{tj} < 0$ then firm j owes $|s_{tj}|$ to its consumers. The vector of "realized" demands $D_t = p(a_t, ,\xi_t)$ can depend on a_t and a nonnegative random vector ξ_t of "raw" demand. For example, $D_t = (d_{0j} - \xi_{tj}\rho_{tj})^+$. Assume that ξ_t and $(s_1, a_1, D_1,..., s_t, a_t)$ are mutually independent and that $\xi_1, \xi_2, ...$ are independent random variables (or vectors). Therefore, for any realization of a_1, a_2 , ..., it follows that D_1, D_2, are conditionally independent.

Let firm j's profit in period t be

$$X_{tj} = g_j(y_t, \rho_t, \xi_t) - c_j z_{tj} \tag{24}$$

in which c_j is the constant unit production cost and g_j is the revenue net of inventory-related costs. For example,

$$g_j(y_t,\ \rho_t,\ \xi_t) = \rho_{tj} \cdot max\{y_{tj}, p(y_t, \rho_t, \xi_t)\}$$
$$- c_{1j} \cdot [y_{tj} - p(y_t, \rho_t, \xi_t)]^+ - c_{2j} \cdot [p(y_t, \rho_t, \xi_t) - y_{tj}]^+$$

Substituting $a_t = (y_{tj}, \rho_{tj})$ and $z_{tj} = y_{tj} - s_{tj}$ shows that (24) satisfies (13) with $K_{tj}(a, x) = g_j[(y,\rho),p(a, x)] - c_j y_j$ and $L_{tj}(s) = c_j s_j$. Also, (23) satisfies (14) with $w[(y,\rho),s] = y_{tj} - p[(y,\rho),s]$.

The decision variables and inventory levels are constrained to lie in compact sets with the following assumption: there are scalars b_1, b_2, b_3, and b_4 such that $b_1 \leq y_{tj} \leq b_2$, $0 \leq \rho_{tj} \leq b_3$, and $P\{D_{tj} \leq b_4\} = 1$ for all t and j.

From Lemma 1, this structure satisfies (9) because it satisfies (13) and (14). If assumptions (11) and (12) are satisfied too, then Corollary 2 asserts the existence of a myopic deterministic equilibrium point $\{a_t^*\}$.

PROPOSITION 1. If $g_j(\,\cdot\,, x)$ and $p_j(\,\cdot\,, x)$ are continuous on A and $g_j(\,\cdot\,, \rho_j, a_{-j}, x)$ is concave on A for each $x \in \Theta$, ρ_j, and a_{-j}, then Γ_t has an equilibrium point in which y_j is unrandomized ($j \in J,\ t \in \mathcal{I}_+$).

PROOF. Debreu (1952). □

Under the assumptions of Proposition 1, let $y_t^* = (y_{tj}^*, j \in J)$ be the unrandomized order-up-to levels included in an equilibrium point for Γ_t. Repeatability (12) is ensured by the following condition on demand bounds in Veinott (1965).

PROPOSITION 2. Under the assumptions of Proposition 1, for each t let m_t satisfy $P\{D_t \geq m_t\} = 1$. If $s_1 \leq y_1^*$ and $y_{t+1}^* \geq y_t^* - m_t$ for all t, then $y_t = y_t^*$ is feasible for all t

PROOF. $P\{D_t \geq m_t\} = 1$ and $y_{t+1}^* \geq y_t^* - m_t$ imply

$$P\{D_t \geq y_t^* - y_{t-1}^*\} = P\{y_t^* - D_t \leq y_{t+1}^*\} = 1 \qquad \square$$

Propositions 1 and 2 greatly simplify Bouakiz and Sobel (1992) by yielding unrestrictive sufficient conditions for that model to have a myopic optimum. They use an inductive argument to establish the existence of an optimal base-stock policy in a dynamic "newsvendor" model with an exponential inter-period utility criterion.

Let $\Gamma(\lambda\beta^{t-1})$ denote Γ_t and let $\{a_t^*\}$ be the static equilibrium points (possibly randomizing ρ_{jt}) whose existence is asserted by Proposition 1. The feasibility of $a_t = \boldsymbol{a}_t^*$ for all t might be studied via the behavior of the sets of equilibrium points of $\Gamma(\theta)$ for $\theta \in (0,\lambda]$. This is an issue of the dependence of the sets of equilibrium points on the values of a parameter in the game. The results in Topkis (1998) (also, cf. Lippman et al., 1987 and Milgrom and Roberts, 1990) may be useful to show for each n that $\Gamma(\lambda\beta^{t-1})$ has least and greatest equilibrium points and that each is monotone with respect to n.

Acknowledgement

This research was partially supported by NSF Grant ECS-8705698.

REFERENCES

Arrow, K. J., Beckmann, M. J., Karlin, S., 1958. The optimal expansion of the capacity of a firm. In: Arrow, K.J., Karlin, S., Scarf, H. (Eds.), Studies in the Mathematical Theory of Inventory and Production, Stanford University Press, Stanford, CA.

Basar, T., Olsder, G. J., 1982. Dynamic Noncooperative Game Theory, Academic Press, New York.

Bell, D.E., 1995. Risk, return, and utility, Management Sci. 41, 23-30.

Bell, D.E., Fishburn, P.C., 2001. Strong one-switch utility, Management Sci. 47, 601-604.

Bellman, R., 1957. Dynamic Programming, Princeton University Press, Princeton, NJ.

Blackwell, D., 1965. Discounted dynamic programming, Annals Math. Stat. 36, 226-235.

Bouakiz, M., Sobel, M. J., 1992. Inventory control with an exponential utility criterion, Oper. Res. 40, 603-608.

Chung, K-J., Sobel, M. J., 1987. Discounted MDP's: distribution functions and exponential utility maximization, SIAM J. Control & Optimization 25, 49-62.

Debreu, G., 1952. A social equilibrium existence theorem, Proc. Nat. Acad. Sci., USA 38, 886-893.

Denardo, E. V., 1967. Contraction mappings in the theory underlying dynamic programming, SIAM Rev. 9, 165-177.

Denardo, E. V., U. G. Rothblum, 1979. Optimal stopping, exponential utility, and linear programming," Math. Programming 1**6**, 228-244.

Denardo, E. V., U. G. Rothblum, 1983. Affine structure and invariant policies for dynamic programs," Math. Oper. Res. 8, 342-365.

Federgruen, A., 1978. On N-person stochastic games with denumerable state space, Adv. Appl. Prob. 10**,** 452-471.

Fink, A. M., 1964. Equilibrium in a stochastic n-person game, J. Sci. Hiroshima Univ. Series A-I 28, 89-93.

Friedman, J.W., 1977. Oligopoly and the Theory of Games. North Holland, Amsterdam.

Fudenberg, D., Tirole, J., 1986. Dynamic Models of Oligopoly. Harwood, New York.

Glicksberg, I., 1952. A further generalization of the Kakutani fixed point theorem with application to Nash equilibrium points,' Proc. Amer. Math. Soc. 3, 170-174.

Howard, R. S., Matheson, J. E., 1972. Risk-sensitive Markov decision processes, Manag. Sci. 18, 356-369.

Jaquette, S. C., 1973. Markov decision processes with a new optimality criterion: discrete time, Annals Stat. 1, 496-505.

Jaquette, S. C., 1976. A utility criterion for Markov decision processes, Manag. Sci. 23, 43-49.

Kirman, A. P., Sobel, M. J., 1974. Dynamic oligopoly with inventories,' Econometrica 42, 279-87.

Lippman, S. A., Mamer, J. W., McCardle, K. F., 1987. Comparative statics in non-cooperative games via transfinitely iterated play," J. Econ. Theory 41, 288-303.

Loewenstein, G., Read, D., Baumeister, R.F. (Eds.), 2003. Time and Decision, Russell Sage Foundation, New York.

Maskin, E., Tirole, J., 2001. Markov perfect equilibrium I. Observable actions, J. Econ. Theory 100 191-219.

Milgrom, P., Roberts, J., 1990. Rationalizability, learning and equilibrium in games with strategic complementarities, Econometrica 58, 1255-1277.

Monahan, G. E., Sobel, M. J., 1997. Risk-sensitive dynamic market share attraction games, Games Econ. Behav. 20 149-160.

Miyamoto, J.M., Wakker, P., 1996. Multiattribute utility theory without expected utility foundations, Oper. Res. 44, 313-326.

Nash, J., 1950. Equilibrium points in n-person games, Proc. Nat. Acad. Sci., USA 36, 48-49.

Nerlove, M., Arrow, K. J., 1962. Optimal advertising capacity under dynamic conditions, Economica 22, 129-142.

Nowak, A. S., 1992. Existence of equilibrium stationary strategies in discounted noncooperative stochastic games with uncountable state space, J. Optimization Theory Appl. 45, 591-602.

Nowak, A. S., 2004. Notes on Risk-sensitive Nash equilibria, Annals Int. Soc. Dynamic Games 7, forthcoming.

Pakes, A., McGuire, P., 1994. Computing Markov-perfect equilibria: numerical implications of a dynamic differentiated product model," RAND J. Econ. 25, 555-589.

Pakes, A., McGuire, P., 2001. Stochastic algorithms for dynamic models: Markov perfect equilibrium, and the 'curse' of dimensionality, Econometrica 69, 1261-1281.

Parthasarthy, K., 1967. Probability Measures on Metric Spaces, Academic Press, New York.

Porteus, E. L., 1990. Stochastic inventory theory, In: Heyman, D. P., Sobel, M. J. (Eds.), Handbooks in Operations Research and Management Science, Vol. 2: Stochastic Models, North-Holland, Amsterdam.

Raghavan, T. E. S., Filar, J. A., 1991. Algorithms for stochastic games - a survey, ZOR - Methods and Models of Oper. Res. 35, 437-472.

Rothblum, U.G., 1975. Multivariate constant risk posture, J. Econ. Theory 10, 309-332.

Sanghvi, A. P., 1978. Sequential games as stochastic processes, Stochastic Proc. Applic. 6, 323-336.

Shapley, L. S., 1953. Stochastic games,' Proc. Nat. Acad. Sci., USA 39, 1095-1100.

Shubik, M., 1959. Strategy and Market Structure. John Wiley, New York.

Shubik, M., Thompson, G. L., 1959). Games of economic survival, Naval Res. Logistics Quart. 6, 111-123.

Sobel, M. J., 1981, Myopic solutions of Markov decision processes and stochastic games," Oper. Res. 29, 995-1009.

Sobel, M. J., 1990a. Myopic solutions of affine dynamic models, Oper. Res. 38, 847-853.

Sobel, M. J., 1990b. Higher order and average reward myopic affine dynamic models, Math. Oper. Res. 15, 299-310.

Sobel, M. J., 1994. Mean-variance tradeoffs in an undiscounted MDP, Oper. Res. 42, 175-183.

Sobel, M. J., 2004. Discounting and risk neutrality, In: Geanakoplis, J., Dubey, P. (Eds.), Feitrschrift Volume for Martin Shubik, Elsevier Science, Amsterdam/New York, to appear.

Topkis, D.M., 1998. Supermodularity and Complementarity. Princeton University Press, Princeton, NJ.

Veinott, A. F. Jr., 1965. Optimal policy for a multi-product, dynamic non-stationary inventory problem, Manag.. Sci. 12, 206-222.

White, D. J., 1989. Mean, variance and probabilistic criteria in finite Markov decision processes: a survey, J. Optimization Theory Applic. 56, 1-30.