Discussion Paper No. 230

A ZERO-SUM STOCHASTIC GAME MODEL OF DUOPOLY

by

S.D. Deshmukh* and Wayne Winston+

July 1976

*Graduate School of Management, Northwestern University, Evanston, Illinois 60201

+Graduate School of Business, Indiana University, Bloomington, Indiana 47401

A B S T R A C T

We consider a discrete time zero-sum stochastic game model of duopoly and give a partial characterization of each firm's optimal pricing strategy. An extension to a continuous time model is also discussed.

1. Introduction

Most previous game-theoretic models of oligopoly (see [6], [9], [11], and [12]) model oligopoly as a static process. The major exception to this rule is the work of Kirman and Sobel [8] which models the problem of determining an optimal inventory level as a non-zero sum stochastic game (see [13] and [15]).

Our paper is an attempt to develop a zero-sum stochastic game (see [1] and [14]) model of the problem of setting a price for a product in a two firm industry. The prices charged by both firms and their current market positions are assumed to influence (in a probabilistic fashion) the future market positions of each firm. This allows us to balance the immediate benefits gained by charging a high price against the future loss of customers caused by charging a high price.

In section 2, our model is introduced and the necessary notation is developed. Sections 3 and 4 give a partial characterization of the optimal policies for each firm. Finally, section 5 introduces a continuous time zero-sum stochastic game model of duopoly and extends the results of sections 3 and 4 to this model.

2. Model Formulation

We consider an industry consisting of two firms (referred to as firms 1 and 2) that produce the same product. The two firms are competing for N customers and at the beginning of any period $t (t = 1,2,...)$ firm 1 is assumed to "control" i customers while firm 2 "controls" N-i customers. (If firm i controls a customer, the customer's purchase, if any, must be made from firm i.) During a period each firm must choose the price at which they wish to sell their product. Firm 1 can choose from ℓ prices $(p_1 < p_2 < \dots < p_\ell)$ while firm 2 can choose from m prices $(p_1^* < p_2^* < \dots < p_m^*)$.

If firm 1 controls i customers at the beginning of period t, firm 1 chooses price p_r, and firm 2 chooses price p_s^*, the following sequence of events ensues.

1. Firm 1 earns a profit $ip_rD_1(p_r) = iR_1(p_r)$, where $D_1(\cdot)$ measures the dependence on price of the fraction of firm 1's customers who purchase the product during a period.
2. Firm 2 earns a profit $(N-i)\, p_s^* D_2(p_s^*) = (N-i)R_2(p_s^*)$, where $D_2(\cdot)$ measures the dependence on price of the fraction of firm 2's customers who purchase the product during a period.
3. With probability $q_{ij}(p_r,p_s^*)$ the number of customers controlled by firm 1 at the beginning of period t+1 changes to j.

To simplify the presentation, we assume $R_1(\cdot)$, $R_2(\cdot)$, and each $q_{ij}(\cdot\cdot)$ are differentiable functions defined on $[p_1,p_\ell]$, $[p_1^*,p_m^*]$ and $[p_1,p_\ell] \times [p_1^*,p_m^*]$, respectively.

For our purposes a two person zero-sum stochastic game (see Bewley and Kohlberg [1] and Shapley [14])is characterized by the following.

1. A finite state space S.
2. The finite set of actions A_i^v available to player v (v = 1,2) in state $i \in S$.
3. The reward or payoff (expected or actual) $r_i(k_1,k_2)$ which is paid from player 2 to player 1 during a period in which the state is i and player v chooses action k_v. Rewards are discounted by a factor β.
4. A set of transition probabilities $\{Q_{ij}(k_1,k_2),\ i,j \in S,\ k_v \in A_i^v\}$, where $Q_{ij}(k_1,k_2)$ is the probability that the state during period t+1 will be j given that the state during period t was i and player v chose action k_v.

If we assume that each firm wishes to maximize the difference between their β-discounted profit and their opponent's β-discounted profit (an assumption which was made in [2]), our duopoly model may be formulated as a two person zero-sum stochastic game with state space $S = \{0,1,\ldots,N\}$, action spaces $A_i^1 = \{1,2,\ldots,\ell\}$ and $A_i^2 = \{1,2,\ldots,m\}$, transition probabilities $Q_{ij}(k_1,k_2) = q_{ij}(k_1,k_2)$ and rewards $r_i(k_1,k_2) = iR_1(p_{k_1}) - (N-i)R_2(p^*_{k_2})$. Thus the actions of each firm correspond to the prices chosen during a period and "game" i is played during any period in which firm 1 controls i customers. The single firm version of this model has been analyzed by Deshmukh and Winston [5].

Let $\Delta(\Pi)$ be the set of <u>stationary strategies</u> for firm 1 (2). Then $\delta \in \Delta$ is a set of N+1 probability vectors

$$\begin{aligned}\vec{\delta}_0 &= (\delta_0(1),\ldots,\delta_0(\ell))\\ &\vdots\\ \vec{\delta}_N &= (\delta_N(1),\ldots,\delta_N(\ell)),\end{aligned}$$

where $\delta_i(k)$ is the probability that firm 1 will charge p_k during any period in which firm 1 controls i customers. Similarly $\pi \in \Pi$ is also a set of N+1 probability vectors

$$\begin{aligned}\vec{\pi}_0 &= (\pi_0(1),\ldots,\pi_0(m))\\ &\vdots\\ \vec{\pi}_N &= (\pi_N(1),\ldots,\pi_N(m)),\end{aligned}$$

where $\pi_i(k)$ is the probability that firm 2 will charge p^*_k during any period in which firm 1 controls i customers. Thus a stationary strategy describes a pattern of behavior that is independent of time.

A strategy $\delta[\pi]$ is a pure strategy if for each $i = 0,1,\ldots,N$ there exist $f(i)$ $[\bar{f}(i)]$ such that $\delta_i(f(i)) = 1[\pi_i(\bar{f}(i)) = 1]$. Let $V(i,\delta,\pi)$ be the expected discounted payoff received by firm 1 during a horizon of infinite length when firm 1 uses δ, firm 2 uses π and at the beginning of period 1 firm 1 controls i customers. If the game is played for an infinite number of periods we seek stationary strategies $\underset{\sim}{\delta} \in \underset{\sim}{\Delta}$ and $\underset{\sim}{\pi} \in \underset{\sim}{\Pi}$ (termed infinite horizon optimal) which satisfy

$$V(i,\underset{\sim}{\delta},\pi) \geq V(i,\underset{\sim}{\delta},\underset{\sim}{\pi}) \geq V(i,\delta,\underset{\sim}{\pi}) \quad \pi \in \Pi,\ \delta \in \Delta,\ i = 0,1,\ldots,N.$$

The existence of infinite horizon optimal strategies was demonstrated by Shapley [14]. Since the state and action spaces in our problem are finite the results of Denardo [3] show that if the game is played for an infinite number of periods restricting our attention to stationary policies is without loss of generality. For infinite horizon problems, we therefore restrict our attention to stationary policies.

Suppose the game is to be played for T periods. For any $\delta = (\delta^{T1},\ldots\delta^{TT})$ $\varepsilon\ X_{t=1}^{t=T}\Delta$ and $\pi\varepsilon(\pi^{T1},\ldots,\pi^{TT})\varepsilon\ X_{t=1}^{t=T}\Pi$ let $V_t^T(i,\delta,\pi)$ be the discounted expected payoff accruing to firm 1 during periods $t,t+1,\ldots,T$ when firm 1 follows δ^{Tk} during period k, firm 2 follows π^{Tk} during period k, and firm 1 controls i customers at the beginning of period t. We seek $\underset{\sim}{\delta}^T \varepsilon\ \Delta^T$ and $\underset{\sim}{\pi}^T \varepsilon\ \Pi^T$ (termed T-period optimal) satisfying

$$V_t^T(i,\underset{\sim}{\delta}^T,\pi^T) \geq V_t^T(i,\underset{\sim}{\delta}^T,\underset{\sim}{\pi}^T) \geq V_t^T(i,\delta^T,\underset{\sim}{\pi}^T)$$

$$\delta^T \ \varepsilon\ X_{t=1}^{t=T}\Delta,\ \pi^T \ \varepsilon\ X_{t=1}^{t=T}\Pi,\ t = 1,2,\ldots,T,\ i = 0,.,\ldots,N.$$

It is well know that T-period optimal policies exist and may be computed by backwards induction.

For T-period optimal strategies $\underset{\sim}{\delta}^T$ and $\underset{\sim}{\pi}^T$

define $\qquad \vec{\underset{\sim}{\delta}}_i^{Tt} = (\underset{\sim}{\delta}_i^{Tt}(1),\ldots,\underset{\sim}{\delta}_i^{Tt}(\ell))$

and $\qquad \vec{\underset{\sim}{\pi}}_i^{Tt} = (\underset{\sim}{\pi}_i^{Tt}(1),\ldots,\underset{\sim}{\pi}_i^{Tt}(m))$

where $\underset{\sim}{\delta}_i^{Tt}(k)(\underset{\sim}{\pi}_i^{Tt}(k))$ is the probability that player 1 (2) will choose price $p_k(p_k^*)$ during period t given that player 1 (2) is following $\underset{\sim}{\delta}^T(\underset{\sim}{\pi}^T)$ and firm 1 controls i customers. Also let $V_t^T(i)$ be the discounted expected payoff accruing to firm 1 during periods $t,t+1,\ldots,T$ when firm 1 follows $\underset{\sim}{\delta}^T$, firm 2 follows $\underset{\sim}{\pi}^T$, and firm 1 controls i customers at the beginning of period t. Finally, let $\Delta V_t^T(i) = V_t^T(i+1) - V_t^T(i)$ and define for $\vec{\delta} = \delta(1),\ldots,\delta(\ell))$ and $\vec{\pi} = (\pi(1),\ldots,\pi(m))$

$$J_t^T(i,\vec{\delta},\vec{\pi}) = \sum_{r=1}^{r=\ell}\sum_{s=1}^{s=m}\sum_{j=0}^{j=N} \delta(r)\ \pi(s)\ \{iR_1(p_r) - (N-i)R_2(p_s^*)$$

$$+ \beta q_{ij}(p_r,p_s^*)\ V_{t+1}^T(j)\},$$

where $V_{T+1}^{T}(j) = 0$.

Then $V_t^T(i) = J_t^T(i, \vec{\delta}_i^{Tt}, \vec{\pi}_i^{Tt})$.

3. Characterization of Optimal Pricing Strategies

Our characterization of optimal pricing strategies will require the following assumptions.

(1) $R_1(p)$ $(R_2(p^*))$ is a concave function attaining its maximum at $p = p_\ell$ $(p^* = p_m^*)$

(2) $q_{i,i+k}(p,p^*) \geq q_{i+1,i+k+1}(p,p^*)$, $k > 0$

(3) $q_{i,i+k}(p,p^*) \leq q_{i+1,i+k+1}(p,p^*)$, $k < 0$

(4) $\sum_{k=0}^{k=r} q_{ik}(p,p^*) \geq \sum_{k=0}^{k=r} q_{i+1,k}(p,p^*)$, $0 \leq r \leq N$

(5) $q_{ik}(p,p^*) = 0$, $k > i+h_1$, $k < i-h_2$

(6) $$\frac{\partial q_{ik}(p,p^*)}{\partial p} \leq 0 \text{ and } \inf_{\Omega_1} \frac{\partial q_{ik}(p,p^*)}{\partial p} = c_1 < 0,$$

(7) $$\frac{\partial q_{ik}(p,p^*)}{\partial p} \geq 0 \text{ and } \sup_{\Omega_2} \frac{\partial q_{ik}(p,p^*)}{\partial p} = c_2 > 0$$

(8) $$\frac{\partial q_{ik}(p,p^*)}{\partial p^*} \geq 0 \text{ and } \sup_{\Omega_1} \frac{\partial q_{ik}(p,p^*)}{\partial p^*} = c_1^* > 0$$

(9) $$\frac{\partial q_{ik}(p,p^*)}{\partial p^*} \leq 0 \text{ and } \inf_{\Omega_2} \frac{\partial q_{ik}(p,p^*)}{\partial p^*} = c_2^* < 0,$$

where $\Omega_1 = \{(i,k,p,p^*): i = 0,1,\ldots,N,\ i < k \leq N,$
$p \in [p_1,p_\ell], p^* \in [p_1^*,p_m^*]\}$

and $\Omega_2 = \{(i,k,p,p^*): i = 0,1,\ldots,N,\ 0 < k < i,$
$p \in [p_1,p_\ell],\ p^* \in [p_1^*,p_m^*]\}$.

Assumption (1) implies that restricting consideration to prices no higher than p_ℓ and p_m^* is not a serious drawback. Assumptions (2) and (3) imply that the more customers a firm controls the more likely they are to lose customers. Assumption (4) implies that the number of customers under firm 1's control during period t+1 is stochastically increasing in the number of customers controlled by firm 1 during period t. Assumptions (2)-(4) would be satisfied, for instance, by a birth death model in which "births" are more likely in lower states and "deaths" are more likely in higher states. Assumption (5) limits the size of single period transitions. Assumptions (6)-(9) restrict the sensitivity of $q_{ik}(p,p^*)$ to changes in p and p^*, and imply that an increase in price has an adverse effect on the firm's next period market share.

The following lemma is necessary to our development.

Lemma 1. For $t = 1,2,\ldots,T$ and $i = 0,1,\ldots,N-1$

(10) $\Delta V_t^T(i) \geq 0$

(11) $\Delta V_t^T(i) \leq (R_1(p_\ell) + R_2(p_m^*)) \sum_{k=0}^{k=T-t} \beta^k = C_t^T$

Proof. The proof is by backwards induction on t. For $t = T$ firm 1 always charges p_ℓ and firm 2 always charges p_m^* so the lemma holds trivially. We therefore assume the validity of the Lemma for t+1 and verify it for t.

Proof of (10)

By the definitions of $\underset{\sim}{\delta}^T$ and $\underset{\sim}{\pi}^T$

$$\Delta V_t^T(i) \geq J_t^T\,(i+1, \vec{\underset{\sim}{\delta}}_i^{Tt}, \vec{\underset{\sim}{\pi}}_{i+1}^{Tt}) - J_t^T(i, \vec{\underset{\sim}{\delta}}_i^{Tt}, \vec{\underset{\sim}{\pi}}_{i+1}^{Tt})$$

$$= \sum_{r=1}^{r=\ell} \sum_{s=1}^{s=m} \sum_{j=0}^{j=N} \underset{\sim}{\delta}_i^{Tt}(r)\; \underset{\sim}{\pi}_{i+1}^{Tt}(s)\; \{K_1(p_r) + R_2(p_s^*)$$

$$+ \beta[q_{i+1,j}(p_r,p_s^*) - q_{ij}(p_r,p_s^*)]\; V_{t+1}^T(j)\}.$$

The last expression is non-negative by the induction hypothesis along with (4) and a lemma of Derman's [4].

Proof of (11)

By the definitions of $\underset{\sim}{\delta}^T$ and $\underset{\sim}{\pi}^T$

$$\Delta V_t^T(i) \leq J_t^T(i+1, \vec{\underset{\sim}{\delta}}_{i+1}^{Tt}, \vec{\underset{\sim}{\pi}}_i^{Tt}) - J_t^T(i, \vec{\underset{\sim}{\delta}}_{i+1}^{Tt}, \vec{\underset{\sim}{\pi}}_i^{Tt})$$

Using (2) and (3) this reduces to

$$\Delta V_t^T(i) \leq \sum_{r=1}^{r=\ell} \sum_{s=1}^{s=m} \{R_1(p_r) + R_2(p_s^*)$$

$$+ \beta \sum_{j=0}^{j=N-1} q_{ij}(p_r,p_s^*) \; \Delta V_{t+1}^T(k)$$

$$+ \beta q_{iN} \; [V_{t+1}^T(i+1) - V_{t+1}^T(N)]$$

$$\leq [R_1(p_\ell) + R_2(p_m^*)] \sum_{k=0}^{k=T-t} \beta^k,$$

where the last inequality is a consequence of (1) and the induction hypothesis. Q.E.D.

Define $$\bar{i}_t^T(k,1) = \frac{\beta \; C_{t+1}^T \; [c_2h_2(h_2+1) - c_1h_1(h_1+1)]}{R_1'(p_k)}$$

and $$\bar{i}_t^T(k,2) = \frac{\beta \; C_{t+1}^T \; [c_1^*h_1(h_1+1) - c_2^*h_2(h_2+1)]}{R_2'(p_k^*)}$$

Given any k_0 we can prove

Theorem 1. If the duopoly game is played for T periods then for $i > \bar{i}_t^T(k_0,1)$ $[N-i > \bar{i}_t^T(k_0,2)]$ and $k < k_0$, $\underset{\sim}{\delta}_i^{Tt}(k) = 0 \quad [\underset{\sim}{\pi}_i^{Tt}(k) = 0]$.

Proof. To prove the result for player 1 define for $p \varepsilon [p_1,p_\ell]$ and $p^* \varepsilon \; [p_1^*,p_m^*]$

$$G_t^T(i,p,p^*) = iR_1(p) - (N-i)R_2(p^*) + \beta \sum_{j=0}^{j=N} q_{ij}(p,p^*) \; V_{t+1}^T(j).$$

It suffices to prove that for $i > \bar{i}_t^T(k_0,1)$, $p \; \varepsilon \; [p_1,p_{k_0}]$ and $p^* \; \varepsilon \; [p_1^*,p_m^*]$ $\frac{\partial G_t^T(i,p,p^*)}{\partial p} \geq 0$, for then (independent of firm 2's strategy) charging p_{k_0} benefits firm 1 more than charging p_k $(k < k_0)$. By (1), (5), (6), and (7),

$$\frac{\partial G_t^T(i,p,p^*)}{\partial p} \geq iR_1'(p_{k_0})$$

$$+ \beta \sum_{j=i+1}^{j=i+h_1} c_1[V_{t+1}^T(j) - V_{t+1}^T(i)]$$

$$+ \beta \sum_{j=i-h_2}^{j=i-1} c_2[V_{t+1}^T(i) - V_{t+1}^T(j)] \geq 0,$$

where the last inequality follows from (11) and the definition of $\bar{i}_t^T(k_0,1)$. A similar proof establishes the analogous result for player 2. Q.E.D.

Thus the optimal pricing policy has a kind of monotonicity property, in the sense that for any price p_{k_0} it is never optimal for a firm to charge less than p_{k_0} if it controls a sufficiently high share of the market. We note that if both the bounds in Theorem 1 exceed N the theorem provides no information.

To prove the infinite horizon analog of Theorem 1 let $(\underset{\sim}{\delta},\underset{\sim}{\pi})$ be any cluster point of the sequence $\{(\underset{\sim}{\delta}^{T1},\underset{\sim}{\pi}^{T1}),\ T = 1,2,\ldots\}$.

A result of Sobel's [16] shows that $\underset{\sim}{\delta}$ and $\underset{\sim}{\pi}$ are infinite horizon optimal strategies for player 1 and player 2, respectively. After defining

$$C = \frac{[R_1(p_\ell) + R_2(p_m^*)]}{1 - \beta}$$

$$\bar{i}(k,1) = \frac{\beta C[c_2h_2(h_2+1) - c_1h_1(h_1+1)]}{R_1'(p_k)},$$

and

$$\bar{i}(k,2) = \frac{\beta C[c_2^*h_1(h_1+1) - c_1^*h_2(h_2+1)]}{R_2'(p_k^*)}$$

the above observation and Theorem 1 immediately yield

<u>Theorem 2</u>. Given any k_0, $i > \bar{i}(k_0,1)[N-i > \bar{i}(k_0,2)]$ and $k < k_0$ imply $\underset{\sim}{\delta}_i(k)=0[\underset{\sim}{\pi}_i(k)=0]$.

4. Separable Stochastic Games

We now consider a type of two person zero-sum stochastic game for which the existence of optimal pure strategies can be demonstrated.

Definition. A two person zero-sum stochastic game is separable if for all states and actions

(12) $r_i(k_1,k_2) = r_i^1(k_1) + r_i^2(k_2)$

and

(13) $Q_{ij}(k_1,k_2) = Q_{ij}^1(k_1) + Q_{ij}^2(k_2)$.

It is easy to prove

Theorem 3. If a separable two person zero-sum stochastic game is played for a horizon of finite or infinite length there exist pure optimal stationary strategies for each player.

Proof. We first assume the game is to be played for T periods. Then (12) and (13) imply there exists at least one pure strategy, call it $\vec{\delta}_i^{Tt}$ which maximizes (for any $\vec{\pi}$) $J_t^T(i,\vec{\delta},\vec{\pi})$. The infinite horizon result now follows by an argument identical to that used to prove Theorem 2. Q.E.D.

For an example of our duopoly game which satisfies (12) and (13) consider a birth and death model specified by

$$q_{i,i+1}(p_r,p_s^*) = \lambda_i^1(p_r) + \mu_{N-i}^2(p_s^*)$$

$$q_{i,i-1}(p_r,p_s^*) = \mu_i^1(p_r) + \lambda_{N-i}^2(p_s^*)$$

$$q_{ii}(p_r,p_s^*) = 1 - (\lambda_i^1(p_r) + \mu_i^1(p_r)) - (\lambda_{N-i}^2(p_s^*) + \mu_{N-i}^2(p_s^*))$$

In Section 5 a model with this structure will be described in greater detail.

5. A Continuous Time Stochastic Game

We now extend our previous results to a continuous time zero-sum stochastic game model of duopoly.

A continuous time zero-sum stochastic game is characterized by the following:

1. A finite state space S.

2. A finite set of actions A_i^v available to player $v(v = 1,2)$ whenever

the state is i. At every instant during which the state is i player v must choose an action $k_v \in A_i^v$.

3. The rate, $\bar{r}_i(k_1,k_2)$, at which a reward is paid from player 2 to player 1 when the state is i and player v chooses action k_v. Rewards are assumed to be continuously discounted at an interest rate r.

4. A set of transition rates $a_{ij}(k_1,k_2)$ where for $j \neq i$, $a_{ij}(k_1,k_2)$ is the instantaneous rate at which a transition occurs from state i to state j when the state is i and player v chooses action k_v.

We define $a_{ii}(k_1,k_2) = -\sum_{j \neq i} a_{ij}(k_1,k_2)$. Assuming that $S = \{0,1,\ldots,N\}$, $A_i^1 = \{1,2,\ldots,\ell\}$, and $A_i^2 = \{1,2,\ldots,m\}$ we let Δ be the set of all <u>stationary policies</u> for player 1, where $\delta \in \Delta$ is a set of N+1 probability vectors

$$\begin{aligned} \vec{\delta}_0 &= (\delta_0(1),\ldots,\delta_0(\ell)) \\ &\vdots \\ \vec{\delta}_N &= (\delta_N(1),\ldots,\delta_N(\ell)), \end{aligned}$$

and $\delta_i(k)$ is the probability that player 1 will choose action k at any instant when the state is i. Similarly we let Π represent the set of all <u>stationary policies</u> for player 2, where $\pi \in \Pi$ is a set of N+1 probability vectors

$$\begin{aligned} \vec{\pi}_0 &= (\pi_0(1),\ldots,\pi_0(m)) \\ &\vdots \\ \vec{\pi}_N &= (\pi_N(1),\ldots,\pi_N(m)), \end{aligned}$$

and $\pi_i(k)$ is the probability that player 2 will choose action k whenever the state is i.

Let $\bar{V}(i,\delta,\pi)$ be the expected discounted payoff received by firm 1 during a horizon of infinite length when player 1 follows δ, player 2 follows π, and the initial state is i. Then $\underset{\sim}{\delta} \in \Delta$ and $\underset{\sim}{\pi} \in \Pi$ are <u>infinite horizon optimal</u> if

$$\bar{V}(i,\underset{\sim}{\delta},\pi) \geq \bar{V}(i,\underset{\sim}{\delta},\underset{\sim}{\pi}) \geq \bar{V}(i,\delta,\underset{\sim}{\pi}) \quad \delta \in \Delta,\ \pi \in \Pi \quad i = 0,1,\ldots,N.$$

Sobel [15] observed that an infinite horizon optimal stationary policy for each player always exists. The results of Denardo [3] again show that restricting our interest to stationary policies is without loss of generality.

For the continuous time stochastic game described above (now referred to as CG) consider a two person zero-sum stochastic game in the spirit of section 1 (labelled DG) described by the following:

State Space	S
Action Spaces	$A_i^v \quad i \in S,\ v = 1,2$
Transition Probabilities	$Q_{ij}(k_1,k_2) = \gamma_{ij} + \dfrac{a_{ij}(k_1,k_2)}{\Lambda}$ where $\gamma_{ij} = \begin{cases} 1 & i = j \\ 0 & \text{otherwise} \end{cases}$
Discount Factor	$\beta = \Lambda/(\Lambda + r)$ where $\Lambda = \max\limits_{\substack{i \in S \\ k_v \in A_i^v}} [-a_{ii}(k_1,k_2)]$
Rewards	$r_i(k_1,k_2) = \beta \bar{r}_i(k_1,k_2)$

This transformation has been used (with great success) by Lippman [10] to analyze the structure of optimal policies for several queuing control problems.

Since CG and DG have identical state and action spaces there is an obvious equivalence between the stationary policies for players in CG and DG. The chain of reasoning given on page 121 of Howard [7] shows that for $i \in S$, $\delta \in \Delta$, and $\pi \in \Pi$, $V(i,\delta,\pi) = \bar{V}(i,\delta,\pi)$. This immediately yields

Theorem 4. If $\underset{\sim}{\delta}(\underset{\sim}{\pi})$ is infinite horizon optimal for DG then $\underset{\sim}{\delta}(\underset{\sim}{\pi})$ is infinite horizon optimal for DG.

We now extend the results of sections 2 and 3 to a continuous time zero-sum stochastic game model of duopoly. We again suppose that two firms are competing for N customers. At any time both firms must choose a price with firm 1's price being a member of $\{p_1,\ldots,p_\ell\}$ and firm 2's price being a member of $\{p_1^*, \ldots, p_m^*\}$. During an instant in which firm 1 controls i customers, firm 1 charges p_r, and firm 2 charges p_s^*, the following happens.

1. Firm 1 earns a profit at the rate $i p_r \bar{D}_1(p_r) = i\bar{R}_1(p_r)$ where $\bar{D}_1(\cdot)$ measures the dependence on price of the rate at which firm 1's customers make a purchase.

2. Firm 2 earns profit at a rate $(N-i)p_s^* \bar{D}_2(p_s^*) = (N-i)\bar{R}_2(p_s^*)$, where $\bar{D}_2(\cdot)$ measures the dependence on price of the rate at which firm 2's customers make a purchase. We again assume that each firm wishes to maximize the difference between their discounted profit and their opponent's discounted profit.

3. Firm 1 gains control of an additional customer according to an exponential distribution with rate $\lambda_i(p_r, p_s^*)$ and loses a customer according to an exponential distribution with rate $\mu_i(p_r, p_s^*)$.

If we assume that $p\bar{D}_1(p)$ and $p^*\bar{D}_2(p^*)$ are differentiable concave functions attaining their maximum at p_ℓ and p_m^*, respectively and define

$$\Omega = \{(i,p,p^*) : i = 0,1,\ldots,N,\ p \in [p_1, p_\ell],\ p^* \in [p_1^*, p_m^*],$$

$$\frac{\partial \lambda_i(p,p^*)}{\partial p} \leq 0 \quad \text{and} \quad \inf_{\Omega} \frac{\partial \lambda_i(p,p^*)}{\partial p} = c_1 < 0,$$

$$\frac{\partial \mu_i(p,p^*)}{\partial p} \geq 0 \quad \text{and} \quad \sup_{\Omega} \frac{\partial \mu_i(p,p^*)}{\partial p} = c_2 > 0,$$

$$\frac{\partial \lambda_i(p,p^*)}{\partial p^*} \geq 0 \quad \text{and} \quad \sup_{\Omega} \frac{\partial \lambda_i(p,p^*)}{\partial p^*} = c_1^* > 0,$$

$$\frac{\partial \mu_i(p,p^*)}{\partial p^*} \leq 0 \quad \text{and} \quad \inf_{\Omega} \frac{\partial \mu_i(p,p^*)}{\partial p^*} = c_2^* < 0,$$

$$C = \left(\frac{\Lambda + r}{r}\right)\left(p_\ell\, \bar{D}_1(p\ ) + p_m^*\, \bar{D}_1(p_m^*)\right),$$

$$\bar{i}_1(k_0) = \frac{C(c_2 - c_1)}{\bar{R}_1'(p_{k_0})}$$

and

$$\bar{i}_2(k_0) = \frac{C(c_1^* - c_2^*)}{\bar{R}_2'(p_{k_0})},$$

then the following result follows immediately from Theorems 2 and 4

Theorem 5. Given any k_0, $i > \bar{i}_1(k_0)$ $(N-i > \bar{i}_2(k_0))$ and $k < k_0$ imply $\underset{\sim}{\delta}_i(k) = 0$ $(\underset{\sim}{\pi}_i(k) = 0)$.

Finally we observe that if

(14) $\bar{r}_i(k_1,k_2) = \bar{r}_i(k_1) + \bar{r}_i(k_2)$,

and

(15) $a_{ij}(k_1,k_2) = a_{ij}(k_1) + a_{ij}(k_2)$

then Theorems 3 and 4 immediately imply

Theorem 6. Any continuous time zero-sum stochastic game satisfying (14) and (15) has pure infinite horizon optimal strategies.

Hopefully the techniques discussed in this section will be an aid in the determination of the structure of optimal policies for other continuous time stochastic games.

References

1. T. Bewley and E. Kohlberg, "Theory of Stochastic Games with Zero-Stop Probabilities, Discussion Paper Number 456, Harvard Institute of Economic Research, January 1976.

2. S. Clemhous, G. Leitman, and H.Y. Wan, Jr., "A Differential Game Model of Duopoly," Econometrica, 39 (1971), 911-938.

3. E. Denardo, "Contraction Mappings in the Theory Underlying Dynamic Programming," SIAM Review, 9 (1967), 165-177.

4. C. Derman, "On Optimal Replacement Rules When Changes of State are Markovian," published in Mathematical Optimization Techniques, R. Rellman (ed.), 1963.

5. S. Deshmukh and W. Winston, "A Controlled Birth and Death Model of Optimal Product Pricing Under Stochastically Changing Demand," submitted to J. Appl. Prob., June 1976.

6. J.W. Friedman, "A Non-Cooperative View of Oligopoly," International Economic Review, 12 (1971), 106-122.

7. R. Howard, "Dynamic Programming and Markov Processes," M.I.T. Press, Cambridge, Mass., 1960.

8. A. Kirman and M. Sobel, "Dynamic Oligopoly with Inventories," Econometrica, 42 (1974), 279-284.

9. R. Levitan and M. Shubik, "Price Variation Duopoly with Differentiated Products and Random Demand," Journal of Economic Theory, 3 (1971), 23-29.

10. S. Lippman, "Applying a New Technique in the Optimization of Exponential Queuing Systems," Operations Research, 23 (1975), 687-711.

11. T. Rasor and Yu-Chi-Ho, "Informational Properties of the Nash Solutions of Two Stochastic Non-Zero-Sum Games," Journal of Economic Theory, 7 (1974), 370-387.

12. J. Roemer, "A Cournot Duopoly Problem," International Economic Review, 11 (1970), 548-552.

13. P. D. Rogers, "Nonzero-Sum Stochastic Games," Report ORC 69-8, Operations Research Center, University of California, Berkeley, 1969.

14. L. Shapley, "Stochastic Games," Proceedings of the National Academy of Sciences, USA, 39 (1953), 1095-1100.

15. M. Sobel, "Noncooperative Stochastic Games,"Ann. Math. Statist., 42 (1971), 1930-1935.

16. M. Sobel, "Notes on Stochastic Games," 1976 (unpublished).