Repeated Partnership Games with Imperfect Monitoring and No Discounting

Roy Radner

The Review of Economic Studies, Vol. 53, No. 1. (Jan., 1986), pp. 43-57.

Stable URL:
http://links.jstor.org/sici?sici=0034-6527%28198601%2953%3A1%3C43%3ARPGWIM%3E2.0.CO%3B2-4

The Review of Economic Studies is currently published by The Review of Economic Studies Ltd..

Review of Economic Studies (1986) LIII, 043-057
0034-6527/86/00030043$02.00

Repeated Partnership Games with Imperfect Monitoring and No Discounting

ROY RADNER
AT & T Bell Laboratories

In a partnership game, each player's utility depends on the other players' actions through a commonly observed consequence (e.g. output, profit, price), which is itself a function of the players' actions and an exogenous stochastic environment. If a partnership game is repeated infinitely, and each player's payoff in the infinite game (supergame) is the long-run average of his expected one-period utilities, then efficient combinations of one-period actions can be sustained as Nash equilibria of the supergame even if the players cannot observe other players' actions or information, but can only observe the resulting consequences.

1. INTRODUCTION

The background for the present study is the problem of achieving efficient decentralized decision-making in an organization. By a decentralized organization I shall mean here one with more than one decision-maker, in which different decision-makers are responsible for different decision variables and make those decisions on the basis of different information, and in which the outcome to the organization depends jointly on the several decisions and on some stochastic environmental variables. For a given system of rewards to the decision-makers, one can in a natural way model the situation as a several-person game.

If the organization is short-lived, or if the members are short-sighted, Nash equilibria of the corresponding game will typically be inefficient, in the sense that some other combination of decision rules would give each member a higher level of expected utility. This will typically be the case even if the designer of the organization has some freedom to choose the system of rewards. (See Section 8 for bibliographic notes.)

The inefficiency of Nash equilibria is not, of course, peculiar to games that are intended to model decentralized decision-making with exogenous uncertainty. This phenomenon is found in the simplest games with certainty; the Prisoners' Dilemma is perhaps the most famous example.[1] At least two approaches have been explored to explain—or prescribe—efficient behaviour in a game. One approach postulates that the players will enter into binding commitments to enforce some "cooperative solution" (Aumann, 1967). The second approach concerns the situation in which the original game is repeated a large (perhaps infinite) number of times. This theory of "repeated games" formalizes the insight that, in the context of organizational decision-making, if the organization is long-lived, and if the members can monitor the information signals and decisions of the other members *ex post*, then the members may have an opportunity to use *self-enforcing* rules of behaviour that sustain efficient decision rules. Formally, this means that there may be Nash equilibria of the repeated game, or *supergame*, that result in efficient decision-making. In particular, in the theory of supergames, the situation in which the one-period game is repeated infinitely often, *with perfect monitoring of all* *random variables and decisions after each repetition*, has been studied in some detail (see Aumann (1983)). Thus, in this situation of perfect monitoring, one might say that the theory of repeated games has explained rigorously how long-term relationships can sustain self-enforcing, efficient behaviour. An important hypothesis in this theory is that the players do not discount future utility. (This may be formalized by taking a player's supergame payoff to be the long-run average of his expected one-period utilities.) If the players do discount future utilities (but not too much), then there will typically be supergame equilibria that are more efficient than one-period equilibria, although not fully (first-best) efficient.

Unfortunately, the same circumstances that lead to the decentralization of information and decision-making usually make perfect ex post monitoring of information and decisions impractical; this consideration leads one to the study of repeated games with imperfect monitoring. As will be explained more precisely in Section 4, in a repeated game each player chooses his strategy in each repetition of the game as a function of some information about the course of the play of all of the previous repetitions. If this information includes less than the complete history of all previous observations and moves of all of the players, then I shall say that the game is one of *imperfect monitoring.*

In the present paper I shall show how, in two classes of infinitely repeated games with imperfect monitoring, one can sustain efficient behaviour as a Nash equilibrium of the repeated game, if the players do not discount future utility. Roughly speaking, after each repetition each player tests the statistical hypothesis that all of the other players have been following the prescribed efficient strategy combination. If at any time any player rejects this hypothesis, then all of the players switch to a one-period (short-run) Nash equilibrium strategy combination for some prescribed number of periods, after which all players return to the efficient strategy combination and the process of statistical testing is started anew.

In the first class of games (Sections 2–6), each player's utility in each repetition depends on the *other* players' moves through some commonly observed consequence (e.g. output, profit, or price), and this sequence of consequences that are observed in common provides the statistics on which the players base their tests. Section 3 illustrates this class of games with an example of a model of decentralized organizational decision-making. In the second class of games (Section 7), each player bases his tests only on the successive observations of his own utilities, but also has the opportunity to announce the results of his tests to the other players. In both cases, the theorems make use of a certain convexity assumption that enables one to bound the gain in expected utility that any one player can achieve from "cheating" by a linear function of the observed statistics.

The assumption that the players do not discount future utility plays an important role in the present analysis. If the players do discount future utilities, then supergame equilibria may be bounded away from full efficiency *uniformly in the players' discount rates*, provided the latter are strictly positive. An example of this phenomenon is presented in the succeeding article in this issue, and hence will not be discussed here.

The concluding section of the paper contains further comments, as well as bibliographic notes.

2. THE ONE-PERIOD GAME WITH CONSEQUENCES OBSERVED IN COMMON

Consider a game in normal form with I players ($I \geqq 2$), and for each player i let $\mathbf{D}_i$ denote i's set of strategies. In order to simplify the exposition, it will be convenient to restrict our attention, until further notice, to the case in which the players' strategies are nonrandom (pure). In Appendix 2 I shall indicate how, with a simple reinterpretation of the notation, one can extend the main result to the case of random strategies. The set of *strategy combinations* is the Cartesian product, $\mathbf{D}$, of the sets $\mathbf{D}_i$.

The utility (payoff) to each player will depend on the strategy combination and on a random variable, X, the state of the environment. I shall assume a special structure for this dependence, namely, one in which each player's utility depends on the other players' strategies through some common "consequence", say C, which is the same for all players. To be precise, if the strategy combination is $D=(D_1, \dots, D_I)$ in $\mathbf{D}$, and the state of the environment is X, then the consequence is

$$C = G(D, X),$$

and the utility to player i is

$$\begin{aligned} U_i &= V_i(C, D_i, X) \\ &= V_i(G[D, X], D_i, X). \end{aligned}$$

The functions G and V_i are real-valued, and are assumed to have suitable measurability properties so that, for each D, C and U_i are random variables.[2]

Assume further that the functions G and V_i are bounded. For every strategy combination D, player i's expected utility is

$$\bar{V}_i(D) = EV_i(G[D, X], D_i, X).$$

Recall that a strategy combination D^* is a *Nash equilibrium* (or equilibrium, for short) if no individual player i can increase his own expected utility by unilaterally changing his own strategy. A strategy combination D is *Pareto-superior* to a strategy combination D' if D yields every player at least as much expected utility as D' does, and yields some player strictly more. A strategy combination is *Pareto-optimal* if there is no strategy combination that is Pareto-superior to it.

In what follows, we shall be concerned with a situation in which an equilibrium strategy combination D^* is not Pareto-optimal, and one wishes to find a means of "enforcing" the use of a strategy combination that is Pareto-superior to it. This will be done by repeating the game, and allowing each player to condition his choice of strategy in any one period on some information about the previous history of play.

3. AN EXAMPLE: PROFIT-SHARING AMONG DECENTRALIZED DECISION-MAKERS

As an example of the one-period game of Section 2, consider an enterprise with I decision-makers or members. Member i observes a random variable Y_i and then chooses an action A_i from some set. The resulting profit to the firm is

$$C = g(A, Z), \tag{3.1}$$

where $A=(A_1, \dots, A_I)$ and Z is a random variable. Suppose that member i receives a share r_iC of the profit, where $r_1, \dots, r_I$ are positive numbers whose sum is 1, and that the resulting utility to member i is

$$P_i(r_iC) - Q_i(A_i), \tag{3.2}$$

where the first term represents his utility of his share of the profit, and the second term represents his "disutility of effort".

Suppose that member i uses a decision rule D_i to determine his action ("effort") as a (measurable) function of his information signal Y_i; thus

$$A_i = D_i(Y_i). \tag{3.3}$$

Let $X = (Y_1, \ldots, Y_I, Z)$, and $D = (D_1, \ldots, D_I)$; then in the notation of Section 2,

$$C = G(D, X) = g[D_1(Y_1), \ldots, D_I(Y_I), Z], \tag{3.4}$$

$$U_i = V_i(C, D_i, X) = P_i(r_iC) - Q_i[D_i(Y_i)], \tag{3.5}$$

which is a summary of (3.1)-(3.3). The strategy of member (player) i is the decision rule, D_i, and the common consequence is the enterprise profit, C.

To illustrate the source of the typical inefficiency of an equilibrium, I consider a special case. Suppose that there are no information signals, so that each player's strategy is the same as his action. Suppose further that, for every realized value of the environment, the profit is an increasing function of the players' efforts. Each player will then have an incentive to be a "free rider" on the other players' efforts. Formally, assume that (1) for every X, $G(A, X)$ is concave, differentiable, and strictly increasing in A, (2) for every i, P_i and Q_i are differentiable and strictly increasing, (3) P_i is concave and Q_i is convex. Let A^* denote an equilibrium combination of actions, and $C^* = G(A^*, X)$. Suppose that the problem is sufficiently regular so that the equilibrium satisfies the first-order conditions: for each i,

$$Er_iP_i'(r_iC^*)G_i'(A^*, X) - Q_i'(A_i^*) = 0, \tag{3.6}$$

where G_i' denotes the partial derivative of G with respect to A_i. For $j \neq i$, the partial derivative of EU_i with respect to A_j is

$$Er_iP_i'(r_iC^*)G_j'(A^*, X),$$

which is strictly positive. Hence, if each player j increases his effort by a "small" amount dA_j, then the change in player i's expected utility will be

$$[Er_iP_i'(r_iC^*)G_i'(A^*, X) - Q_i'(A_i^*)]dA_i + \textstyle\sum_{j \neq i} Er_iP_i'(r_iC^*)G_j'(A^*, X)dA_j. \tag{3.7}$$

The first term of (3.7) is zero, by the first-order condition (3.6), and the second term of (3.7) is strictly positive. Hence a small simultaneous increase in effort by all players will make all players better off. In economic jargon, the second term of (3.7) represents the "positive externalities" generated by the players' efforts, which are ignored by each player in the (noncooperative) equilibrium.

4. THE REPEATED GAME (SUPERGAME)

Suppose that the one-period game described in Section 2 is repeated infinitely often. The resulting game, which I shall now describe in detail, is called the *supergame*. In each period $t = 1, 2, \ldots$, *ad infinitium*, each player i chooses a one-period strategy, D_{it}, the state of the environment is X_t, the resulting consequence is C_t, and the one-period utility to player i is U_{it}.

In the course of the play of the one-period game in period t, player i will have observed some information signal Y_{it}, a random variable. Assume that at the end of period t, player i knows

$$H_{it} = (Y_{i1}, \ldots, Y_{it}, D_{i1}, \ldots, D_{it}, C_1, \ldots, C_t, U_{i1}, \ldots, U_{it}). \tag{4.1}$$

I shall call H_{it} *player i*'s information history at t. In the supergame, each player is allowed to condition his choice of one-period strategy in period t on his information history at $t-1$. Formally, a *supergame strategy* for player i is a sequence of functions, $s_i = (s_{i1}, s_{i2}, \ldots)$ such that, for t, s_{it} is a (measurable) function from the set of his information histories $H_{i,t-1}$ at $t-l$ to the set D_i of his one-period strategies. (Note that, in the language of the example in Section 3, player i's *action* in period t can depend on both his information history $H_{i,t-1}$ at $t-1$ and his information signal Y_{it} in period t.)

Let each Y_{it} be some function of X_t, the same function for all t (but not for all i), and assume that the successive X_t's are independent and identically distributed.[3]

Given the *supergame strategy* combination $s = (s_1, \ldots, s_I)$, the resulting sequence (U_{it}) of one-period utilities for each player i is determined recursively in the obvious way. These utilities are, of course, random variables. Assume that player i's criterion for the supergame is

$$W_i(s) = \liminf_{T\to\infty} E(1/T) \textstyle\sum_{t=1}^{T} U_{it}, \tag{4.2}$$

where E denotes the mathematical expectation operator. This is one way of formalizing the idea that the players do not discount future utilities. One can now define Nash equilibrium, Pareto-superiority, and Pareto-optimality for the supergame in the obvious way, just as in the one-period game.

5. OPTIMAL EQUILIBRIA OF THE SUPERGAME

Let D^* be a Nash equilibrium strategy combination in the one-period game, yielding one-period expected utilities (u_i^*). It is easy to check that the supergame strategy combination s^*, in which each player i uses the one-period strategy D_i^* in each period t, whatever be his information history at $t-1$, is a Nash equilibrium of the supergame, and yields the long-run average expected utilities (u_i^*).

We shall be interested in constructing Nash equilibria of the supergame that are Pareto-superior to D^*, or even Pareto-optimal. We shall do this by constructing supergame strategy combinations in which for "most of the time" the players use a Pareto-optimal one-period strategy combination, say $\hat{D}$, that is Pareto superior to D^*. Suppose that $\hat{D}$ yields expected one-period utilities $(\hat{u}_i)$, and the corresponding expected one-period consequence $\hat{c}$. Since $\hat{D}$ is Pareto-superior to D^*,

$$\hat{u}_i \geqq u_i^*, \qquad i = 1, \ldots, I. \tag{5.1}$$

To motivate the next assumption, consider the special case in Section 3. Starting from a Pareto-optimal one-period strategy combination, any one member i could increase his own expected utility only by decreasing someone else's, and this would imply that he could increase his own expected utility only by decreasing the *expected* total profit. Furthermore, this relationship between the change in expected profit and the change in i's expect utility would be concave, so that the "frontier" could be supported by a positive linear function. This motivates the following definition of Property PLS ("positive linear support"):

Property PLS. Given the one-period strategy combination $\hat{D}$, for every player i there is a strictly positive number K_i such that: for any (one-period) strategy D_i for i, if i uses D_i and every member j other than i uses $\hat{D}_j$, let c be the corresponding expected consequence and let u_i be i's one-period expected utility; then

$$(u_i - \hat{u}_i) + K_i(c - \hat{c}) \leqq 0. \tag{5.2}$$

Theorem 5.1. *If Property PLS is satisfied, and if one-period utilities and consequences are uniformly bounded, then there exists a Nash equilibrium of the supergame that yields each member i the long-run-average expected utility $\hat{u}_i$, i.e.*

$$\lim_{T\to\infty} E\frac{1}{T}\sum_{t=1}^{T} U_{it} = \hat{u}_i. \tag{5.3}$$

Proof. The required equilibrium of the supergame will be constructed with "relenting" strategies, which were informally described in Section 1. Let $\hat{C}_t$ be the realized consequence in period t if the one-period strategy combination $\hat{D}$ is used in that period, and let

$$\bar{C}_t = \frac{1}{t}(\hat{C}_1 + \cdots + \hat{C}_t) \tag{5.4}$$

be the corresponding average consequence for the first t periods. The consequences $\hat{C}_t$ are independent and identically distributed, with expected value $\hat{c}$, and are bounded. By the strong law of large numbers, the sequence $(\bar{C}_t)$ converges to $\hat{c}$, almost surely. Therefore (see the Appendix), there is a sequence (b_t^0) with the following properties:

$$b_t^0 \text{ converges to } 0; \tag{5.5}$$

$$tb_t^0 \text{ is nondecreasing in } t; \tag{5.6}$$

$$|\bar{C}_t - \hat{c}| \text{ is eventually not greater than } b_t, \text{ almost surely;} \tag{5.7}$$

furthermore, for any $d > 0$, there is a $k > 1$ such that

$$\text{Prob}\{|\bar{C}_t - \hat{c}| \leqq kb_t^0, \text{ for all } t\} \geqq 1 - d. \tag{5.8}$$

For example, by the law of the iterated logarithm, one may take (for t sufficiently large)

$$b_t^0 = k_0\left(\frac{\ln\ln t}{t}\right)^{1/2}, \tag{5.9}$$

$$k^0 > (2 \text{ Var } C_t)^{1/2}.$$

Pick any number d strictly between 0 and 1, let k correspond to d as in (5.8), and define the sequence $(b(t))$; by

$$b(t) = kb_t^0, \qquad t = 1, 2, \ldots \tag{5.10}$$

I shall now define the "cooperative" and "noncooperative" phases, recursively. Let (ε_m) be any strictly positive sequence converging to zero. For any given joint supergame strategy, let C_t be the realized consequence in period t, and define the random times N_m and $(N_m + N'_m)$, with the latter possibly infinite, as follows.

$$N_0 = N'_0 = N''_0 = 0. \tag{5.11}$$

For $m \geqq 0$, if $N_0, \ldots, N_m$ are finite, let

$$N'_{m+1} = \inf\{n : \textstyle\sum_{t=N_m+1}^{N_m+n} (C_t - \hat{c}) \leqq -nb(n) \cdot n \geqq 1\}. \tag{5.12}$$

If $N'_{m+1} = \infty$, let $N = N_m$ and $M = m$. If $N'_{m+1} < \infty$, let N''_{m+1} be the smallest integer such that, for every i,

$$N_m(u_i^* + \varepsilon_m) + N'_{m+1}\hat{u}_i + K_i[N'_{m+1}b(N'_{m+1}) + B] + N'_{m+1}u_i^*$$
$$\leqq (N_m + N'_{m+1} + N''_{m+1})(u_i^* + \varepsilon_{m+1}), \tag{5.13}$$

where B is a uniform bound on $|C_t|$, and K_i is as in Property PLS; let

$$N_{m+1} = N_m + N'_{m+1} + N''_{m+1}. \tag{5.14}$$

Inequality (5.13) may be rewritten as

$$N''_{m+1}\varepsilon_{m+1} \geqq N'_{m+1}(\hat{u}_i - u_i^* - \varepsilon_{m+1}) + N_m(\varepsilon_m - \varepsilon_{m+1}) + K_i[N'_{m+1}b(N'_{m+1}) + B]. \tag{5.15}$$

If, for every m, $N'_m < \infty$, then let $M = \infty$.

For $1 \leqq m \leqq M$ and m finite define:

$$\begin{aligned} \mathbf{N}'_m &= \{t\colon N_{m-1} < t \leqq N_{m-1} + N'_m\}, \\ \mathbf{N}''_m &= \{t\colon N_{m-1} + N'_m < t \leqq N_m\}, \\ \mathbf{N}_m &= \mathbf{N}'_m \cup \mathbf{N}''_m. \end{aligned} \tag{5.16}$$

For $M < \infty$, define

$$\mathbf{N}'_{m+1} = \{T\colon t > N_m\}. \tag{5.17}$$

The interval $\mathbf{N}'_m$ is the m-th "cooperative phase," and for $m \leqq M$ and finite $\mathbf{N}''_m$ is the m-th "noncooperative phase" and $\mathbf{N}_m$ is the mth *complete epoch* M is the number of completed epochs. Let b denote the sequence $(b(t))$, and define the strategy $\sigma_j(b)$ by:

$$j \text{ uses } \begin{cases} \hat{D}_j \text{ in the cooperative phases,} \\ D_j^* \text{ in the noncooperative phases.} \end{cases} \tag{5.18}$$

I shall show that $(\sigma_j(b))$ is the required equilibrium of the supergame.

If each member j uses $\sigma_j(b)$, then $C_t = \hat{C}_t$ in every cooperative phase. By property (5.8), $M < \infty$ almost surely; hence, for each i, by the strong law of large numbers,

$$\begin{aligned} \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} U_{it} &= \lim_{T\to\infty} \frac{1}{T}\left[\sum_{t=1}^{N} U_{it} + \sum_{t=N+1}^{T} \hat{U}_{it}\right]. \\ &= \hat{u}_i, \quad \text{a.s.} \end{aligned}$$

The successive averages are uniformly bounded; hence, by the Lebesgue Dominated Convergence theorem,

$$\lim_{T\to\infty} E\frac{1}{T}\sum_{t=1}^{T} U_{it} = \hat{u}_i,$$

so that the strategy $(\sigma_j(b))$ yields the required utility.

It remains to show that $(\sigma_j(b))$ is an equilibrium of the supergame. Suppose that every member j other than i uses $\sigma_j(b)$, and i uses some strategy $\bar{\sigma}_i$. Fix T (for the moment) and let $M(T)$ be the number of epochs completed by T, $N(T)$ be the end of the last epoch completed by T, $\mathbf{N}'(T)$ be all the "cooperative periods" from 1 to T, and $\mathbf{N}''(T)$ be all the "noncooperative periods" from 1 to T. To be precise,

$$\begin{aligned} M(T) &= \max\{m\colon N_m \leqq T\}, \\ N(T) &= N_{M(T)}, \\ \mathbf{N}'(T) &= [\textstyle\bigcup_{m \leqq M(T)} \mathbf{N}'_m] \cup \{t\colon N(T) < t \leqq T\}, \\ \mathbf{N}''(T) &= \textstyle\bigcup_{m \leqq M(T)} \mathbf{N}''_m. \end{aligned}$$

Let H_t denote the partial history $(X_1, \ldots, X_t)$. By Property PLS,

$$E(U_{it}|H_{t-1}) \leqq \hat{u}_i + K_i[\hat{c} - E(C_t|H_{t-1})], \quad \text{for } t \text{ in } \mathbf{N}'(T),$$

and since D^* is a short-term Nash equilibrium,

$$E(U_{it}|H_{t-1}) \leqq u_i^*, \quad \text{for } t \text{ in } N''(T).$$

Therefore, by the optional sampling theorem for supermartingales (see, e.g. Chung (1974, Exercise 16, p. 330), and the detailed argument in the Appendix),

$$E \sum_{t=1}^{T} U_{it} \leqq E \sum_{t\in N'(T)} [\hat{u}_i + K_i(\hat{c} - C_t)] + E \sum_{t \in N''(T)} u_i^*. \tag{5.19}$$

By the stopping rule for ending the cooperative phases,

$$\sum_{t\in N'(T)} (C_t - \hat{c}) \geqq -\sum_{m \leqq M(T)} [N'_m b(N'_m) + B] - [T - N(T)]b[T - N(T)] - B. \tag{5.20}$$

Combining (5.13), (5.19), and (5.20) one gets

$$E \sum_{t=1}^{T} U_{it} \leqq EN(T)(u_i^* + \varepsilon_{M(T)}) + E[T - N(T)][\hat{u}_i + K_i b(T - N(T))] + K_i B,$$

$$\leqq T\hat{u}_i + EN(T)\varepsilon_{M(T)} + K_i E[T - N(T)]b[T - N(T)] + K_i B. \tag{5.21}$$

Note that to get the last inequality one uses the hypothesis that $u_i^* \leqq \hat{u}_i$.

We now examine the behaviour of the right side of (5.21), as T increases without limit. If $M = \infty$, then $M(T) \to \infty$, and

$$\frac{N(T)\varepsilon_{M(T)}}{T} \leqq \varepsilon_{M(T)} \to 0.$$

If $M < \infty$, then

$$\frac{N(T)}{T} \to 0 \quad \text{and} \quad \frac{N(T)\varepsilon_{M(T)}}{T} \to 0.$$

Therefore, by the Lebesgue Dominated Convergence Theorem,

$$\lim_{T\to\infty} E\frac{N(T)\varepsilon_{M(T)}}{T} = 0. \tag{5.2v}$$

By the definition of the sequence $b(t)$, $tb(t)$ is increasing in t; hence

$$[T - N(T)]b[T - N(T)] \leqq Tb(T),$$

so that, since $b(T) \to 0$,

$$\lim_{T\to\infty} E(1/T)[T - N(T)]b[T - N(T)] = 0 \tag{5.23}$$

(use Lebesgue again). From (5.21)-(5.23) one immediately gets

$$\lim\sup_{T\to\infty} E(1/T) \sum_{t=1}^{T} U_{it} \leqq \hat{u}_i, \tag{5.24}$$

which is actually somewhat stronger than is needed. ||

6. OTHER SUPERGAME EQUILIBRIA

The equilibria described in Theorem 5.1 are not the only supergame equilibria. In particular, for any one-period Nash equilibrium strategy combination D^*, the supergame strategy combination in which each player i plays D_i^* in every period is a supergame equilibrium.

Furthermore, the set of supergame equilibria has a certain convexity property, which I shall now explain. For any supergame strategy combination s, let $w(s)$ denote the vector with coordinates $w_i(s)$, as in (4.2), and let W^* denote the set of vectors $w(s)$ such that s is a supergame equilibrium. I shall call a vector in W^* a *supergame equilibrium utility vector.* A supergame strategy combination s is called *summable* if for every player i the lim inf in (5.2) is actually an ordinary limit; let W^{**} denote the set of supergame equilibrium utility vectors corresponding to summable strategies, or supergame *summable equilibrium* utility vectors. Note that the strategies described in Theorem 5.1, as well as the strategy in which D^* is played in every period, are summable.

For two vectors $w=(w_i)$ and $w'=(w_i')$, I shall use the notation $w \geqq w'$ to mean that $w_i \geqq w_i'$ for each coordinate i.

Theorem 6.1. *If w and w' are in W^*, and k is a number between* 0 *and* 1, *then there is a w'' in W^* such that*

$$w'' \geqq kw + (1-k)w'; \tag{6.1}$$

if, furthermore, w and w' are in W^{**}, then the inequality in (6.1) *can be replaced by an equality.*

Corollary. *The set W^{**} of supergame summable equilibrium utility vectors is convex.*

(The proof of Theorem 6.1 uses well-known methods, and is omitted.)

Let W^0 denote the set of vectors $w(s)$ such that s is a supergame strategy in which a single equilibrium one-period strategy is played in every period. As was noted above, W^0 is a subset of W^{**}. According to Theorem 5.1, if u is the vector of expected utilities corresponding to a one-period strategy that satisfies Property PLS, and $u \geqq w$ for some w in W^0, then u is in W^{**}. In fact, one can show that the last statement is true if w is any convex combination of vectors in W^0. Thus let W' denote the set of all expected utility vectors u that correspond to a one-period strategy that satisfies Property PLS, and such that $u \geqq w$ for some w that is a convex combination of vectors in W^0; then W' is contained in W^{**}. If one combines this last statement with Theorem 6.1 one gets:

Theorem 6.2. *W^{**} contains all convex combinations of vectors in the union of W^0 and W'.*

It is not known to me whether there are vectors in W^{**} other than those described in Theorem 6.2.

7. GAMES WITHOUT CONSEQUENCES OBSERVED IN COMMON

I now turn to the case in which the players cannot rely on the observation of some commonly observed consequences to determine, before each period, whether cooperative behaviour should continue. I shall show how the players need use only the observations of their own respective utilities on which to base their rules for stopping cooperation, *provided that they have the means to signal to the other players their intent to do so.*

Suppose that the formulation in Section 2 is replaced by the more general formulation.

$$U_{it} = V_i(D_t, X_t), \tag{7.1}$$

where $D_t = (D_{1t}, \dots, D_{It})$ is the joint one-period strategy used in period t. Again, assume that the random variables X_t are independent and identically distributed.

Let $\bar{U}$ be the set of all vectors of one-period expected utilities that correspond to some one-period joint strategy. If $\bar{U}$ were convex, and if $\hat{u}$ were Pareto-optimal in $\bar{U}$, then there would be a nonnegative, nonzero vector L that supports $\bar{U}$ at $\hat{u}$, i.e. such that

$$L(u-\hat{u}) \leqq 0 \tag{7.2}$$

for all u in U. Call $\hat{u}$ *regular* if there is a corresponding supporting vector L that is *strongly* positive (that is, all of whose coordinates are strictly positive). It is known that the set of regular Pareto-optimal vectors in $\bar{U}$ is dense in the set of all Pareto-optimal vectors (see Arrow, Barankin, and Blackwell, 1953). This motivates the following definition, of Property PLSU ("positive linear support of utilities").

Property PLSU. A vector $\hat{u}$ in $\bar{U}$ has Property PLSU if there is a strongly positive vector L such that (7.2) holds for all u in $\bar{U}$.

Suppose now that at the end of each one-period game each player i announces a signal, say M_{it}, which can be 0 or 1. The interpretation of this signal will be that, if the players are in a cooperative phase at t, then $M_{it}=0$ signals that player i intends to begin a *non*cooperative phase in period $t+1$, whereas $M_{it}=1$ signals that player i is willing to continue the cooperative phase into period $t+1$ (provided that all the others are also willing).

Again, during period t player i may have observed some information signal Y_{it}. In this section, I shall assume that at the end of period t player i's information history is

$$H_{it}=(Y_{i1},\ldots,Y_{it}, D_{i1},\ldots,D_{it}, U_{i1},\ldots,U_{it}, M_1,\ldots,M_t), \tag{7.3}$$

where M_t is the vector with coordinates M_{it}.

Consider now the following situation. Let D^* be a one-period Nash equilibrium joint strategy yielding a vector u^* of one-period expected utilities, let $\hat{D}$ be another one-period joint strategy, yielding a vector $\hat{u}$ of one-period expected utilities, and suppose that $\hat{u}$ has property PLSU and $\hat{u} \geqq u^*$. I shall indicate how to construct a Nash equilibrium of the supergame that yields the long-run average expected utility vector $\hat{u}$, by using a method similar to that of Theorem 5.1.

Let $(b(t))$; be a sequence as in (5.5)-(5.10). The first cooperative phase will end after the first period n such that, for some player j,

$$U_{j1}+\cdots+U_{jn} \leqq n\hat{u}_j - nb(n); \tag{7.4}$$

call this period N'; if no such finite period exists, let $N'=\infty$. Let $M_{jN'}=0$, and let $M_{it}=1$ for all players i in all periods $t<N'$, and for all players i other than j in period n', where j is any player for which (7.4) holds.

Consider a particular player, say i, and suppose that all players j other than i use $\hat{D}_j$ during periods 1 through N'. By property PLSU, for any strategy of player i, and any period $t \leqq N'$,

$$\textstyle\sum_j L_j E(U_{jt}-\hat{u}_j \mid H_{t-1}) \leqq 0, \tag{7.5}$$

where $H_t=(X_1,\ldots,X_t)$. Hence, by (7.5) and the optional sampling theorem for supermartingales,

$$\begin{aligned} E\textstyle\sum_{t=1}^{N'}(U_{it}-\hat{u}_i) &\leqq -\textstyle\sum_{j\neq i}\left(\dfrac{L_j}{L_i}\right)E\textstyle\sum_{t=1}^{N'}(U_{jt}-\hat{u}_j) \\ &\leqq \textstyle\sum_{j\neq i}\left(\dfrac{L_j}{L_i}\right)E(N'b(N')+B) \\ &= K_i E(N'b(N')+B), \end{aligned} \tag{7.6}$$

where

$$K_i = \sum_{j \neq i} \left(\frac{L_j}{L_i} \right), \tag{7.7}$$

and B is a uniform bound on $|U_{jt}|$ (all j and t). One can now define N'', the length of the first noncooperative phase, as in (5.13). Proceeding in this way, as in the proof of Theorem 5.1, one can define a (possibly finite) sequence of alternating cooperative and noncooperative phases. Thus one can prove:

Theorem 7.1. *If u^* is the vector of one-period expected utilities corresponding to a one-period Nash equilibrium, $\hat{u}$ is a vector of one-period expected utilities corresponding to some one-period strategy, $\hat{u}$ satisfies property PLSU and $u \geqq u^*$, and players' one-period utilities are uniformly bounded, then there is a supergame equilibrium that yields each player i the long-run-average expected utility $\hat{u}_i$.*

One can show that the analogues of Theorems 6.1 and 6.2 also hold for the model of this section.

8. FURTHER COMMENTS AND BIBLIOGRAPHIC NOTES

The particular problem studied here was inspired in part by the important paper, "Incentives in Teams", by T. Groves (1973). In that paper Groves showed how, in a model like that of Section 3, efficient decentralized decision-making could be sustained in the one-period game by a suitably chosen reward function if (1) the decision-makers were neutral towards risk, and (2) there was an additional player, the "organizer", who chose the reward functions, and who received the difference between the realized enterprise profit and the total rewards to the decision-makers. An organization with property (2) was called a *foundation* by Marschak (1954). (For other aspects of the design of incentives in organzations see (Hurwicz, 1972, 1979).)

A special case of the foundation is the "principal-agent" model, in which there is one decision-maker (the "agent") and an organizer (the "principal"). For repeated principal-agent games, results analogous to those obtained here have been derived by Radner (1981, 1985), Rubinstein (1979*b*), and Rubinstein and Yaari (1983).

The model of the present paper is not a foundation, however, in that there is no organizer other than the group of decision-makers themselves; the decision-makers are the only players in the game, and are the only ones whose interests are considered.

One-period partnership games have been studied by Holmstrom (1982). Green (1980, Section 5), Porter (1983), and Green and Porter (1984) study a particular class of repeated partnership games, with discounting, that correspond to a dynamic version of the Cournot oligopoly model. In these games, the "partners" are the firms in an oligopolistic market. In each period, each firm chooses a quantity of output; the resulting market price is a function of the total industry output and some exogenous stochastic variable (a "disturbance" in the inverse demand function). Each firm can observe the market price, but not the individual outputs of the other firms. In this context, "efficiency" corresponds to an effective cartel arrangement in which total industry expected profit is maximized, whereas a competitive outcome is "inefficient" from the point of view of the firms. Attention is focused on "trigger strategies." Let q^* be a one-period equilibrium output combination, and let q be another output combination that gives each firm a higher expected profit than q^* does. Let B be some price that is lower than the price that would be expected if the output combination q were used. In a trigger-strategy combination, the firms produce the output combination q until the first time that the market price is lower than B; thereafter they produce the output combination q^*. Green (1980) gives conditions under which, as the number of firms increases without limit, the market outcomes of supergame trigger-strategy equilibria approach competitive outcomes. Porter (1983) and Green and Porter (1984) study similar supergames with a fixed number of firms that use relenting trigger strategies, in which each noncooperative phase lasts only a finite number of periods, with a subsequent return to a cooperative phase. These trigger-strategy equilibria cannot be fully efficient (i.e., cannot achieve the maximum expected profit for a cartel), but do give each firm a higher expected profit than it would get in a noncooperative one-period equilibrium.

With respect to efficiency, repeated partnership games with discounting appear to be markedly different from repeated principal-agent games, and from repeated games with perfect monitoring. Taken together with the results of the present paper, the succeeding article in this issue shows that the correspondence that maps the players' discount rate into the corresponding set of supergame equilibria can be discontinuous at the point at which the discount rate is zero. References to the relevant literature are given in that article.

APPENDIX 1

The first lemma justifies (4.5)-(4.8) in the proof of Theorem (5.1).

Lemma A.1. *Let (Ω, F, μ) be a probability space, and let (f_t) be a uniformly bounded sequence of measurable functions converging to 0 a.e. on Ω; then there is a sequence (b_t^0) of numbers such that*

$$(b_t^0) \text{ converges to } 0; \tag{A.1}$$

$$(tb_t^0) \text{ is nondecreasing}; \tag{A.2}$$

$$\text{for a.e. } \omega,\ |f_t(\omega)| \text{ is eventually not greater than } b_t^0. \tag{A.3}$$

Furthermore, for any $d > 0$ there is a $k > 1$ such that

$$\text{Prob}\,\{|f_t(\omega)| \leqq kb_t^0 \text{ for all } t\} \geqq 1-d. \tag{A.4}$$

Proof. Let $\varepsilon > 0$ be fixed. By Egorov's theorem (see, e.g., Halmos (1950, p. 88)) for each positive integer m there is a set R_m such that $\mu(R_m) \geqq 1-\varepsilon^m$, and (f_t) converges uniformly on R_m. Let

$$S_m = \bigcup\nolimits_{k \leqq m} R_k;$$

then $\mu(S_m) \geqq 1-\varepsilon^m$, the sequence (S_m) is monotone nondecreasing, and (f_t) converges uniformly on S_m, and hence, for every positive integer n that there is a t_{mn} such that, for all ω in S_m and all $t \geqq t_{mn}$,

$$|f_t(\omega)| \leqq \frac{1}{n}.$$

Let

$$u_n = \max\{t_{mn}: m \leqq n\}.$$

For ω in S_m, $n \geqq m$, and $t \geqq u_n$, one has $t \geqq t_{mn}$ and hence

$$|f_t(\omega)| \leqq \frac{1}{n}.$$

Let B be a uniform bound on (f_t), and define

$$b_t = \begin{cases} B, & \text{for } 1 \leqq t \leqq u_1, \\ \dfrac{1}{n}, & \text{for } u_n \leqq t < u_{n+1} \end{cases}$$

(Note that (b_t) converges to 0.) For ω in S_m and $t \geqq u_m$,

$$|f_t(\omega)| \leqq b_t.$$

But $S = \bigcup_m S_m$ has μ-measure one, so that (A.3) is satisfied for a.e. ω.

To get (A.2), one can modify the sequence (b_t) by letting $b_1^0 = b_1$ and

$$b_t^0 = \max\left\{b_t, \left(\frac{t-1}{t}\right)b'_{t-1}\right\}, \qquad t > 1;$$

then the sequence (b_k^0) satisfies (A.1)-(A.3).

To prove (A.4), let

$$g(\omega) = \sup_t \frac{|f_t(\omega)|}{b_t^0};$$

by (A.3), $g(\omega)$ is finite almost surely. Hence for every $\varepsilon > 0$ there is a $k > 0$ such that

$$\text{Prob}\,\{g(\omega) \leqq k\} \geqq 1 - \varepsilon.$$

Condition (A.4) now follows immediately, which completes the proof of Lemma A.1.

The next lemma establishes (5.19) in the proof of Theorem 5.1.

Lemma A.2. *Under the hypotheses of Theorem* 5.1, *condition* (5.19) *is satisfied.*

Proof. Let $\tilde{A}_{it}$ denote the action taken by i in period t if he uses strategy $\tilde{\sigma}_i$ and every $j \neq i$ uses $\sigma_j(b)$. Each $\tilde{A}_{it}$ is, of course, a random variable, measurable with respect to H_t. Now let $\tilde{C}_t$ and $\tilde{U}_{it}$ be the consequence and i's utility in period t, respectively, if i uses the action $\tilde{A}_{it}$ and every $j \neq t$ uses $\hat{D}_j$. By Property PLS,

$$E(\tilde{U}_{it}|H_{t-1}) \leqq \hat{u}_i - K_i[E(\tilde{C}_t|H_{t-1}) - \hat{c}],$$

or

$$E(Z_t|H_{t-1}) \geqq 0,$$

where

$$Z_t \equiv \hat{u}_i - K_i(\tilde{C}_t - \hat{c}) - \tilde{U}_{it}, \quad t \geqq 1$$

If we let $Z_0 \equiv 0$, then

$$Z_0 + \cdots + Z_t$$

is a submartingale.

Let $v_0 = 0$, and $v_t = 1$ *or* 0 according as t is or is not in $N'(T)$. Then

$$v_0 Z_0 + \cdots + v_t Z_t$$

is also a submartingale (Chung (1974, p. 330, Exercise 16)). Hence

$$E \sum_{t \in N'(T)} Z_t = E(v_0 Z_0 + \cdots + v_T Z_T) \leqq 0.$$

But for $t \in N'(T)$,

$$Z_t = \hat{u}_i - K_i(C_t - \hat{c}) - U_{it},$$

so that

$$E \sum_{t \in N'(T)} U_{it} \leqq E \sum_{t \in N'(T)} [\hat{u}_i - K_i(C_t - \hat{c})]. \tag{A.5}$$

By a similar argument

$$E \sum_{t \in N''(T)} U_{it} \leqq E \sum_{t \in N''(T)} u_i^*. \tag{A.6}$$

Adding (A.5) and (A.6) yields (5.19). ||

APPENDIX 2. MIXED STRATEGIES

Many games have no Nash equilibrium in pure strategies, but do have equilibria in mixed strategies. (For example, a game in which each player has only a finite number of pure strategies always has a Nash equilibrium in mixed strategies.) In order for Theorem 5.1 to be applicable to such (one-period) games in a nonvacuous way, one wants to be assured that the formulation of the model in Sections 2 and 4 can accommodate mixed strategies in the one-period game. This can be done if the spaces of pure strategies are sufficiently regular. (In this section, I shall be referring only to strategies in the one-period game.)

To be specific, assume that for each player i the set D_i of pure strategies is a separable metric space whose measurable sets are the Borel sets. A *mixed strategy* for player i is a probability measure on the Borel sets of D_i. Adjoin to the environment variable X a random vector, say Z, that is uniformly distributed on the unit cube, $[0, 1]^I$ in I-dimensional space and independent of X, and adjoin to player i's information signal Y_i the i'th coordinate of Z. Thus let the new environment be $X' = (X, Z)$, and i's new information signal be $Y'_i = (Y_i, Z_i)$. Note that $X, Z_1, \ldots, Z_I$ are independent, and that each Z_i is uniformly distributed on the unit interval. With this augmented information signal, a pure strategy for i is a measurable function from the unit interval to D_i. It can be shown[4] that for any probability measure m_i on the Borel sets of D_i there is a measurable function, say f_i, from the unit interval to D_i such that if Z_i is uniformly distributed on the unit interval then $f_i(Z_i)$ will be distributed in D_i according to the measure m_i. Hence a mixed strategy for i in the original formulation of the game can be represented as a pure strategy with the augmented information.

First version received August 1981; *final version accepted July* 1985 (*Eds.*)

I am grateful to R. J. Aumann, S. Hart, and R. W. Rosenthal for helpful discussions on the subject of this research.

NOTES

1. On the Prisoners' Dilemma, see Luce and Raiffa (1957).
2. It is to be understood in the models described in this paper that there is an underlying probability space, that all random variables under consideration are measurable functions from that probability space to some measurable space (e.g. the real line with its Borel sets), and that all functions considered are suitably measurable.
3. This implies that information signals from different periods are statistically independent. However, it does not exclude the possibility that, in any one period, the information signals of different players may be statistically dependent.

4. For example, use Halmos (1950, p. 173, Theorem C), together with the fact that a probability space can have at most a countable number of atoms.

REFERENCES

ARROW, K. J., BARANKIN, E. W. and BLACKWELL, D. (1953), "Admissible Points of Convex Sets", in Kuhn H and Tucker, A. W. (eds.) *Contributions to the Theory of Games*, Vol. II (Princeton, N.J.: Princeton Univ. Press) 87-92.

AUMANN, R. J. (1967), "A Survey of Cooperative Games without Side Payments", in Shubik M. (ed.) *Essays in Mathematical Economics in Honor of Oscar Morgenstern* (Princeton, N.J.: Princeton Univ. Press) 3-27.

AUMANN, R. J. (1983), "Survey of Repeated Games", in Aumann, R. *et al.*, (eds.) *Essays in Game Theory*, (Mannheim: Bibliographisches Institut).

CHUNG, K. L. (1974). *A Course in Probability Theory*, second edition (New York: Academic Press).

GREEN, E. J. (1980), "Noncooperative Price-Taking in Large Dynamic Markets", *Journal of Economic Theory*, **22**, 37-64.

GREEN, E. J. and PORTER, R. H. (1984), "Noncooperative Collusion under Imperfect Price Information", *Econometrica*, **52**, 87-100.

GROVES, T. (1973), "Incentives in Teams", *Econometrica*, **41**, 617-631.

HALMOS, P. (1950) *Measure Theory* (New York: Van Nostrand).

HOLMSTROM, B. (1982), "Moral Hazard in Teams", *Bell Journal of Economics*, **13**, 324-340.

HURWICZ, L. (1972), "On Informationally Decentralized Systems", in McGuire C. B. and Radner, R (eds.) *Decision and Organization* (Amsterdam: North-Holland).

HURWICZ, L. (1979), "On the Interaction Between Information and Incentives in Organizations", in Krittendorff, K. (ed.) *Communication and Control in Society* (New York: Gordon and Breach) 123-147.

LUCE, D. and RAIFFA, H. (1957) *Games and Decisions* (New York: Wiley).

MARSCHAK, J. (1954), "Towards an Economic Theory of Organization and Information", in Thrall, R. M. Coombs, C. H. and Davis, R. L. (eds.) *Decision Processes* (New York: Wiley) 187-220.

PORTER, R. H. (1983), "Optimal Cartel Trigger-Price Strategies", *Journal of Economic Theory*, **29**, 313-338.

RADNER, R. (1981), "Monitoring Cooperative Agreements in a Repeated Principal-Agent Relationship", *Econometrica*, **49**, 1127-1148.

RADNER, R. (1985), "Repeated Principal-Agent Games with Discounting", *Econometrica*, **53**, 1173-1198.

RADNER, R., MYERSON, R. and MASKIN, E. (1986), "An Example of a Repeated Partnership Game with Discounting and with Uniformly Inefficient Equilibria", *Review of Economic Studies*, this issue.

RUBINSTEIN, A. (1979), "Equilibrium in Supergames with the Overtaking Criterion", *Journal of Economic Theory*, **21**, 1-9.

RUBINSTEIN, A. (1979), "Offenses That May Have Been Committed by Accident—An Optimal Policy of Retribution," Brams S. J., Schotter, A. and Schrödiauer, G. (eds.) *Applied Game Theory*, (Wurzburg: Physica-Verlag).

RUBINSTEIN, A. and YAARI, M. E. (1983), "Repeated Insurance Contracts and Moral Hazzard," *Journal of Economic Theory*, **30**, 74-97.

References

Noncooperative Collusion under Imperfect Price Information
Edward J. Green; Robert H. Porter
Econometrica, Vol. 52, No. 1. (Jan., 1984), pp. 87-100.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198401%2952%3A1%3C87%3ANCUIPI%3E2.0.CO%3B2-6

Incentives in Teams
Theodore Groves
Econometrica, Vol. 41, No. 4. (Jul., 1973), pp. 617-631.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28197307%2941%3A4%3C617%3AIIT%3E2.0.CO%3B2-E

Monitoring Cooperative Agreements in a Repeated Principal-Agent Relationship
Roy Radner
Econometrica, Vol. 49, No. 5. (Sep., 1981), pp. 1127-1148.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198109%2949%3A5%3C1127%3AMCAIAR%3E2.0.CO%3B2-T

Repeated Principal-Agent Games with Discounting
Roy Radner
Econometrica, Vol. 53, No. 5. (Sep., 1985), pp. 1173-1198.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198509%2953%3A5%3C1173%3ARPGWD%3E2.0.CO%3B2-M