JOURNAL OF ECONOMIC THEORY 51, 194–206 (1990)

Nash and Perfect Equilibria of Discounted Repeated Games*

D. FUDENBERG

Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

AND

E. MASKIN

Harvard University, Cambridge, Massachusetts 02138

Received March 11, 1987; revised June 20, 1989

The "perfect Folk Theorem" for discounted repeated games establishes that the sets of Nash and subgame-perfect equilibrium payoffs are equal in the limit as the discount factor δ tends to one. We provide conditions under which the two sets coincide *before* the limit is reached. That is, we show how to compute $\underline{\delta}$ such that the Nash and perfect equilibrium payoffs of the δ-discounted game are identical for all $\delta > \underline{\delta}$. *Journal of Economic Literature* Classification Number: 022. © 1990 Academic Press, Inc.

1. INTRODUCTION

The "Folk Theorem" for infinitely repeated games with discounting asserts that any feasible, individually rational payoffs (payoffs that strictly pareto dominate the minmax point) can arise as Nash equilibria if the discount factor δ is sufficiently near one. Our paper [3] established that under a "full-dimensionality" condition (requiring the interior of the feasible set to be nonempty) the same is true for perfect equilibria, so that in the limit as δ tends to 1 the requirement of subgame perfection does not restrict the

* We thank two referees and an associate editor for helpful comments. Research support from the U.K. Social Science Research Council, National Science Foundation Grants SES-8607229 and SES-8520952, and the Sloan Foundation is gratefully acknowledged.

0022-0531/90 $3.00
Copyright © 1990 by Academic Press, Inc.
All rights of reproduction in any form reserved.

set of equilibrium payoffs. These results leave open the question of whether the restriction to perfect equilibria constrains the set of equilibrium payoffs for fixed discount factors near 1.

We show that perfectness is *not* a constraint under two simple conditions. Under these conditions, there is a discount factor $\underline{\delta}$ less than 1 such that for all larger discount factors the Nash and perfect equilibrium payoffs are identical. (An earlier version of this paper, Fudenberg and Maskin [4], provided examples to show that the perfectness constraint *does* bind when our conditions are not satisfied.)

The key to our proof is the construction of perfect equilibria, one for each player, that hold that player to exactly his reservation value. Since no strategies by the player's opponents can prevent him from obtaining his reservation value, any path of play that can arise in a Nash equilibrium can be enforced by using the reservation-value perfect equilibria to "punish" all deviations from the support of the equilibrium distribution, so the Nash and perfect equilibrium payoffs coincide. (Note that we do not assert that all Nash equilibrium *strategies* are perfect.) This observation is closely related to one of Abreu [2]. He shows that any pure-strategy perfect equilibrium path can be enforced with strategies that punish any deviation with the pure-strategy perfect equilibrium that minimizes the deviator's payoff.[1]

The equivalence between the Nash and perfect equilibrium payoffs holds even though, for any fixed $\delta < 1$, there will typically be individually rational payoffs that cannot arise in equilibrium. In other words, the payoff sets coincide before attaining their limiting values. One purpose of the paper is to clarify the connection between the Nash and perfect equilibria of repeated games and thus provide additional insight into the structure of the perfect equilibria. The paper also makes a more practical contribution. Our proof explicitly constructs the reservation–value equilibria, which are quite simple in form, and also provides an easily computed expression for the value of $\underline{\delta}$. This may make the present result easier to apply than our theorem in [3], where the discount factor required for a given payoff to arise in a perfect equilibrium depended on the particular payoff in more complicated ways.

Section 2 introduces our notation for the repeated game model. Section 3 presents the main results. Through Section 3 we make free use of the possibility of *public* randomization. That is, we suppose that there exists some random variable (possibly devised by the players themselves) whose realizations are publicly observable. Players can thus use the random variable to (perfectly) coordinate their actions. In Section 4, however, we show that our results do not require public randomization.

[1] Abreu proves that these worst pure-strategy equilibria exist in games with a continuum of actions.

2. NOTATION

We consider a finite n-player game in normal form,

$$g: A_1 \times \cdots \times A_n \to R^n,$$

where $g(a_1, ..., a_n) = (g_1(a_1, ..., a_n), ..., g(a_1, ..., a_n))$ and $g_i(a_1, ..., a_n)$ is player i's payoff from the vector of actions $(a_1, ..., a_n)$. Player i's mixed strategies, i.e., the probability distributions over A_i, are denoted Σ_i. For notational convenience, we will write "$g(\sigma)$" for the expected payoffs corresponding to the vector of mixed strategies $\sigma = (\sigma_1, ..., \sigma_n)$.

In the repeated version of g, each player i maximizes the normalized discounted sum π^i of his per-period payoffs, with common discount factor δ:

$$\pi^i \equiv (1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} g_i(\sigma(t)).$$

Here $\sigma(t)$ is the vector of mixed strategies chosen in period t.

Player i's strategy in period t can depend on the past *actions* of all players, that is on the sequence $\{a(\tau)\}_{\tau<t}$, but not on the past choices of randomizing probabilities $\sigma(\tau)$. Each period's play can also depend on the realization of a publicly observed random variable such as sunspots. Although we feel that allowing public randomization is as reasonable as prohibiting it, our results do not require it. Section 5 explains how, for δ close to 1, the effect of sunspots can be duplicated by deterministic cycles of play.

To give a formal definition of the strategy spaces, let $\omega(t)$ be a sequence of independent random variables with the uniform distribution on $[0, 1]$. The history at time t is $h^t = (h^{t-1}, a(t-1), \omega(t))$, and a strategy for player i is a sequence $\{s_i^t\}$, where $s_i^t: H^t \to \Sigma_i$ and H^t is the space of time t histories.

For each player j, choose minmax strategies $m^j = (m_1^j, ..., m_n^j)$, where

$$m_{-j}^j \in \arg\min_{m_{-j}} \max_{m_j} g_j(m_j, m_{-j}) \quad \text{and} \quad g_j(m^j) = \max_{a_j} g_j(a_j, m_{-j}^j). \tag{0}$$

Let

$$v_j^* = g_j(m^j).$$

(Here "m_{-j}" is a mixed strategy selection for players other than j, and $g_j(a_j, m_{-j}^j) = g_j(m_1^j, ..., m_{j-1}^j, a_j, m_{j+1}^j, ..., m_n^j)$.)

We call v_j^* player j's reservation value. Since one feasible strategy for player j is to play in each period a static best response to that period's play of his opponents, player j's average payoff must be at least v_j^* in any equilibrium of g, whether or not g is repeated. Note that any *Nash* equilibrium path of the repeated game can be enforced by the threat that any deviation by j will be punished by the other players' minmaxing j (i.e., playing m^j_{-j}) for the remainder of the game.

Henceforth we shall normalize the payoffs of the game g so that $(v_1^*, ..., v_n^*) = (0, ..., 0)$. Call $(0, ..., 0)$ the *minmax point*, and take $\bar{v}_i = \max_a g_i(a)$. Let

$$U = \{(v_1, ..., v_n) \mid \text{there exists } (a_1, ..., a_n) \in A_1 \times \cdots \times A_n \text{ with } g(a_1, ..., a_n) = (v_1, ..., v_n)\},$$

$$V = \text{convex hull of } U,$$

and

$$V^* = \{(v_1, ..., v_n) \in V \mid v_i > 0 \text{ for all } i\}.$$

The set V consists of feasible payoffs, and V^* consists of feasible payoffs that strictly Pareto dominate the minmax point. That is, V^* is the set of feasible, strictly individually rational payoffs.

3. Nash and Perfect Equilibrium

Any feasible vector of payoffs $(v_1, ..., v_n)$ that gives each player i at least $(1-\delta)\bar{v}_i$ is attainable in a Nash equilibrium, since Nash strategies can specify that any deviator from the actions sustaining $(v_1, ..., v_n)$ will be minmaxed forever. In a subgame-perfect equilibrium, however, the punishments must themselves be consistent with equilibrium play, so that the punishers must be given an incentive to carry out the prescribed punishments. One way to try to arrange this is to specify that players who fail to minmax an opponent will be minmaxed in turn. However, such strategies may fail to be perfect, because minmaxing an opponent may be more costly than being minmaxed oneself. Still, even in this case, one may be able, as in Abreu [1] and our paper [3], to induce players to minmax by providing "rewards" for doing so.

In fact, the present paper demonstrates that under certain conditions these rewards can de devised in such a way that the punished player is held to *exactly* her minmax level. When this is possible, the sets of Nash and perfect equilibrium payoffs coincide, as the following lemma asserts.

LEMMA 1. *For discount factor δ, suppose that, for each player i, there is a perfect equilibrium of the discounted repeated game in which player i's payoff is exactly zero. Then the sets of Nash and perfect equilibrium payoffs (for δ) coincide.*

Proof. Fix a Nash equilibrium s^*, and construct a new strategy $\hat{s}$ that agrees with s^* along the equilibrium path, but specifies that, if player i is the first to deviate from s^*, play switches to the perfect equilibrium that holds player i's payoff to zero. (If several players deviate simultaneously, the deviations are ignored.) Since zero is the worst punishment that player i could have faced in s^*, he will not choose to deviate from the new strategy $\hat{s}$. By construction, $\hat{s}$ is a perfect equilibrium with the same payoffs as s^*. Q.E.D.

Remark. Note that the lemma does not conclude that all Nash equilibrium *strategies* are perfect.

A trivial application of Lemma 1 is to a game, like the Prisoners' Dilemma, in which there is a one-shot equilibrium that gives all players their minmax values. An only slightly more complex case is a game where each player prefers to minmax rather than to be minmaxed, i.e., a game in which $g_i(m^j) > 0$ for all $i \neq j$. In such a game we need not reward punishers to ensure their compliance but can simply threaten them with future punishment if they fail to punish an opponent.

THEOREM 1. *Suppose that for all i and j, $i \neq j$, m_j^i, as defined by* (0), *is a pure strategy, and that $g_j(m^i) > 0$. Let $\underline{\delta}$ satisfy $\bar{v}_j(1-\underline{\delta}) < \min_{i \neq j} g_j(m^i)$ for all j. Then for all $\delta \in (\underline{\delta}, 1)$, the sets of Nash and perfect equilibrium payoffs of the repeated game exactly coincide.*

Proof. For each player i, define the ith "punishment equilibrium" as follows. Players play according to m^i until some player $j \neq i$ deviates. If this occurs, they switch to the punishment equilibrium for j. Player i has no incentive to deviate from the ith punishment equilibrium because in every period he is playing his one-shot best response. Player $j \neq i$ may have a short-run gain to deviating, but doing so results in his being punished, so that the maximum payoff to deviation is $\bar{v}_j(1-\delta)$,[2] which is less than $\min_{i \neq j} g_j(m^i)$ by assumption. So the hypotheses of Lemma 1 are satisfied. Q.E.D.

Remark 1. If the minmax strategies are mixed instead of pure, the construction above is inadequate because player j may not be indifferent among all actions in the support of m_j^i. Example 1 of Section 4 shows that in this case Theorem 1 need not hold.[3]

[2] Recall that we are expressing players' payoffs in the repeated game as *discounted average* payoffs, and not as present values.

[3] However, in two-player games we can sharpen Theorem 1 by replacing its hypotheses with the condition that for all i and j, $i \neq j$, and all a_j in the support of m_j^i, $g_j(a_j, m_{-j}^i)$ is positive. Note that this condition reduces to that of Theorem 1 if all the m_j^i are pure strategies.

Remark 2. The proof of Theorem 1 actually shows that all feasible payoff vectors that give each player j at least $\min_{i \neq j} g_j(m^i_j)$ can be attained in equilibrium if δ exceeds the $\underline{\delta}$ defined in the proof.

Although the hypotheses of Theorem 1 are not pathological (i.e., they are satisfied by an open set of payoffs in nontrivial normal forms), they do not apply to many games of interest. We now look for conditions that apply even when minmaxing an opponent gives a player less than her reservation utility.

In this case, to induce a player to punish an opponent we must give him a "reward" afterwards, as we explained earlier. To construct equilibria of this sort, it must be possible to reward one player without also rewarding the player he punishes. This requirement leads to the "full-dimensionality" requirement we introduced in our earlier paper: the dimensionality of V should equal the number of players. However, full dimensionality is not sufficient for the stronger results of this paper, as we show in [4]. We must strengthen it to require that the minmax point $(0, ..., 0)$ is itself in the interior of V. Moreover, we need to assume that each player i has an action $\hat{a}_i$ such that $g_i(\hat{a}_i, m^i_{-i}) < 0$, so that when minmaxed, a player has an action for which he gets a strictly negative payoff. (From our normalization, his maximum payoff when minmaxed is zero.)

THEOREM 2. *Assume that* (i) *the minmax point is in the interior of* V, *and that* (ii) *for each player i there exists* $\hat{a}_i$ *such that* $g_i(\hat{a}_i, m^i_{-i}) < 0$. *Then there exists* $\underline{\delta} < 1$ *such that for all* $\delta \in (\underline{\delta}, 1)$, *the sets of Nash and perfect equilibrium payoffs of the repeated game exactly coincide.*

COROLLARY. *Under the conditions of Theorem* 2, *for* $\delta > \underline{\delta}$, *any feasible payoff vector* v *with* $v_i \geqslant \bar{v}_i(1-\delta)$ *for all* i *can be attained by a perfect equilibrium.*

Remark. Hypothesis (ii) of the theorem is satisfied by generic payoffs in normal forms with three or more actions per player. The interiority condition (i) is, however, not generic: the minmax point can be outside V for an open set of payoffs. The condition is important because it ensures that in the construction of equilibria to reward a punisher, the ratio of his payoff to that of the deviator can be made as large as we like.

Proof of Theorem 2. We begin in part (A) with the case in which all the minmax strategies m^i_j are pure, or, equivalently, each player's choice of a mixed strategy is observable. Part (B) explains how to use the methods of our paper [3] to extend the construction to the case where the minmax strategies are mixed.

(A) Assume that each m^i_j is a pure strategy. For each player i, choose an action $\hat{a}_i$ such that $g_i(\hat{a}_i, m^i_{-i}) \equiv -x_i < 0$. For $j \neq i$, let $y^i_j \equiv -g_j(\hat{a}_i, m^i_{-i})$. The equilibrium strategies will have $2n$ "states," where n is the number of players. States 1 through n are the "punishment states," one for each player; states $n+1$ to $2n$ are "reward states." In the punishment state i, the strategies are: Play $(\hat{a}_i, m^i_{-i})$ today. If there are no deviations, switch to state $n+i$ tomorrow with probability $p_i(\delta)$ (to be determined), and remain in state i with complementary probability $1-p_i(\delta)$. If player j deviates, then switch to state j tomorrow. In reward state $n+i$, players play actions to yield payoffs $v^i = (v^i_1, ..., v^i_n)$, which are to be determined. If player j deviates in a reward state, switch to punishment state j.

Choose v^i in V^* such that, for $i \neq j$, $x_i v^i_j - v^i_i y^i_j > 0$ (this is possible because $0 \in \operatorname{int} V$). Now set $p_i(\delta) = (1-\delta)x_i/\delta v^i_i$ and choose $\bar{\delta} > x_i/(v^i_i + x_i)$, such that for $\delta > \bar{\delta}$, $p_i(\delta) < 1$. This choice of $p_i(\delta)$ sets player i's payoff starting in state i equal to zero if she plays as specified. If player j does not deviate, his payoff starting in state i, which we denote w^i_j, solves the functional equation

$$w^i_j = (1-\delta)(-y^i_j) + \delta p_i(\delta)\, v^i_j + \delta(1-p_i(\delta))\, w^i_j, \tag{1}$$

so that

$$w^i_j = (x_i v^i_j - y^i_j v^i_i)/(v^i_i + x_i). \tag{2}$$

By construction, the numerator in (2) is positive. The interiority condition has allowed us to choose the payoffs in the reward states so as to compensate the punishing player j without raising i's own payoff above zero. Choose $\hat{\delta} < 1$ large enough so that, for all i and j, $v^i_j > \bar{v}_j(1-\hat{\delta})$, and so that for $i \neq j$, $w^i_j > \bar{v}_j(1-\hat{\delta})$. Set $\underline{\delta} = \max(\hat{\delta}, \bar{\delta})$.

We claim that for all $\delta \in (\underline{\delta}, 1)$, the specified strategies are a perfect equilibrium. First consider punishment state i. In this state, player i receives payoff zero by not deviating. If player i deviates once and then conforms, she receives at most zero today (since she is being minmaxed) and she has a normalized payoff of zero from tomorrow on. Thus player i cannot gain by a single deviation, and the usual dynamic programming argument implies that she cannot gain by multiple deviations. Player j's payoff in state i is w^i_j (which exceeds $\bar{v}_j(1-\delta)$). A deviation could yield as much as $\bar{v}_j$ today, but will shift the state to state j, where j's payoff is zero, so player j cannot profit from deviating in state i. Finally, in reward state $n+i$, each player k obtains payoff v^i_k exceeding $\bar{v}_k(1-\delta)$, and so the threat of switching to punishment state k prevents deviations. The theorem now follows from Lemma 1.

(B) The strategies in part (A) punish player j if, in state i, he fails to use his minmax strategy m^i_j. The players can detect all deviations from m^i_j

only if it is a pure strategy, or if the players' choices of mixed strategies are observable. (Otherwise, player j would not be detected if he deviated to a different mixed strategy with the same support). However, following our [3] paper, we can readily modify the construction of (A) to allow for mixed minmax strategies. The idea is to specify that player j's payoff in reward state $n+i$ depend on the last action he took in punishment state n in such a way that player i is exactly indifferent among all the actions in the support of m_j^i.

To begin, let $\{a_j^i(k)\}$ be the pure strategies in the support of m_j^i, where the indexation is chosen so that

$$y_j^i(k) \equiv -g_j(\hat{a}_i, a_j(k), m^i_{-j-i}) \leqslant -g_j(\hat{a}_i, a_j^i(k+1), m^i_{-j-i}) \equiv y_j^i(k+1),$$

where m^i_{-j-i} is the vector of minmax strategies against i by players other than j. Thus $-y_j^i(k)$ is player j's expected payoff in punishment state i if she plays her kth-best strategy in the support of m_j^i. Next define

$$c = \max_{i, a_i, a_i^0, \sigma_{-i}} |g_i(a_i, \sigma_{-i}) - g_i(a_i^0, \sigma_{-i})|.$$

This is the maximum variation a player's choice of action can induce in his own payoff. Also, let $\varepsilon > 0$ be such that all payoff vectors v with $0 < v_i < 3\varepsilon$ for all i are in the interior of V. (This is possible because $0 \in \text{int } V$.)

As in part (A), our strategies will have n punishment states and n reward states, with a probability $p_i(\delta)$ of switching from state i to state $n+i$ if player i played $\hat{a}_i$ and each $j \neq i$ played an action in the support of m_j^i. However, when play switches to state $n+i$, player j's payoff depends on the action that she took in the preceding period. Denote these payoffs by $v_j^i(k_j)$, where k_j is the index (as defined in the preceding paragraph) of the action last played by player j. Thus the vector of payoffs in state $n+i$ is

$$\begin{aligned} &v^i(k_1, ..., k_{i-1}, k_{i+1}, ..., k_n) \\ &\quad = (v_1^i(k_1), ..., v_{i-1}^i(k_{i-1}), v_i^i, v_{i+1}^i(k_{i+1}), ..., v_n^i(k_n)). \end{aligned}$$

This v^i is defined as follows: First choose v_i^i and $v_j^i(1)$ for each j to satisfy

$$x_i v_j^i(1) - v_i^i y_j^i(1) > 0, \tag{3}$$

$$v_i^i < \varepsilon x_i / c, \tag{4}$$

and

$$v_j^i(1) < \varepsilon. \tag{5}$$

These conditions can be satisfied because $0 \in \text{int } V$. As in part (A), let $p_i(\delta) = (1-\delta)\, x_i/\delta v_i^i$. Now for each j and k_j, set

$$v_j^i(k_j) = v_j^i(1) + (1-\delta)[y_j^i(k_j) - y_j^i(1)]/\delta p_i(\delta). \tag{6}$$

With this specification of the reward payoffs, player j's payoff in state i is the same for each strategy in the support of m_j^i, and equals $[x_i v_j^i(1) - y_j^i(1)\, v_i^i]/(v_i^i + x_i)$, which is positive from inequality (3).

Now we must check that the specified payoffs for state $n+i$ are all feasible, and that, for δ near one, no player wishes to deviate. Substituting the definition of $p(\delta)$ into (6), we have

$$v_j^i(k_j) = v_j^i(1) + [y_j^i(k_j) - y_j^i(1)](v_i^i/x_i). \tag{7}$$

Referring to the bounds (4) and (5), we see that $v_j^i(k_j) < 3\varepsilon$ for all j and k_j, and thus the vector v^i is feasible for all values of the k_j's. Finally, since player i's payoff in state i is zero, and his payoff in state $n+j$ is bounded away from zero for all j, no player will wish to deviate in the reward states $n+j$. In state i, player $j \neq i$ obtains the same positive payoff from any strategy in the support of m_j^i, and she will be punished with zero for deviating from the support. Thus, for δ close to 1, player j will be willing to play m_j^i. The argument that player i will not deviate in state i is exactly as in Case A. Q.E.D.

Remark. Our working paper [4] provides a series of examples to explore the roles of the hypotheses of Theorems 1 and 2. One example shows that Theorem 1 need not hold when the minmax strategies are mixed. A second example establishes that the hypotheses of Theorems 1 and 2 do not imply that all individually rational payoffs can be attained for some δ strictly less than one. We also provide counterexamples to Theorem 2 when either hypothesis (i) or (ii) is dropped. These counterexamples, which are two-player games with two actions per player, suggest that it may be difficult to obtain Theorem 2 under weaker conditions.

The interiority hypothesis of Theorem 2 implies that the set V has full dimension, i.e., that $\dim V = n$. Let us briefly consider the connection between Nash and perfect equilibrium when V has lower dimension. When the number of players n exceeds two, our article [3] shows by example that the Nash and perfect equilibrium payoff sets need not coincide even in the limit as δ tends to 1. Thus, for such examples, these sets do not coincide for $\delta < 1$. However, we obtain a result much like that of Theorem 2 for two player games when $\dim V = 1$. In this case, the V is simply a line segment. It if has negative slope, the game is zero-sum. If it has positive slope, the payoffs can be normalized so that the two players' payoffs are always equal, i.e., the game is one of pure coordination.

THEOREM 3. *In a two-player game where* dim $V=1$, *there exists* $\underline{\delta}<1$ *such that, for all* $\delta \in (\underline{\delta}, 1)$, *the sets of Nash and perfect equilibrium payoffs coincide.*

Proof. Let (m_1^2, m_2^1) be a pair of minmax strategies. (If there are multiple such pairs, choose one that maximizes player 1's payoff.) If $g(m_1^2, m_2^1)=(0,0)$, then (m_1^2, m_2^1) forms a Nash equilibrium of the one-shot game. In this case, as infinite repetition of a one-shot Nash equilibrium constitutes a perfect equilibrium of the repeated game, application of Lemma 1 establishes the theorem. Suppose, therefore, that $g_i(m_1^2, m_2^1)<0$ for some i. But then g cannot be a constant sum game and so, since dim $V<2$, we can normalize the players' payoffs so that, for all (a_1, a_2), $g_1(a_1, a_2)=g_2(a_1, a_2)$. Take $\underline{v}=g_1(m_1^2, m_2^1)$ $(=g_2(m_1^2, m_2^1))$. Because $\underline{v}<0$, there must exist (a_1^*, a_2^*) such that $g(a_1^*, a_2^*)=(v^*, v^*)$, where $v^*>0$ (otherwise (m_1^1, m_2^1), where m_1^1 is a best response to m_2^1, is a minmax pair for which $g(m_1^1, m_2^1)=(0,0)$, contradicting the choice of (m_1^2, m_2^1)). We will show that, for δ near 1, there exists a perfect equilibrium of the repeated game in which both players' payoffs are zero. When m_1^2 and m_2^1 are pure strategies this is easily done by choosing $p(\delta)$ such that

$$(1-\delta)\,\underline{v}+\delta p(\delta)\,v^*=0. \tag{8}$$

The equilibrium strategies consist of playing (m_1^2, m_1^1) in the first period and then either switching (permanently) to (a_1^*, a_2^*) with probability $p(\delta)$ or else starting again with probability $1-p(\delta)$. Deviations from this path are punished by restarting the equilibrium. Suppose that the support of m_1^2 is $\{a_1(1), ..., a(R)\}$ and that of m_2^1 is $\{a_2(1), ..., a_2(S)\}$. Suppose that the probability of $a_h(k)$ is $q_h(k)$. By definition,

$$\sum_j \sum_i q_1(i)\, q_2(j)\, g_1(a_1(i), a_2(j))=\underline{v}, \tag{9}$$

and since the m_j^i are minmax strategies,

$$\sum_i q_1(i)\, g_2(a_1(i), a_2(j))\leqslant 0 \qquad \text{for all } j \tag{10}$$

and

$$\sum_j q_2(j)\, g_1(a_1(i), a_2(j))\leqslant 0 \qquad \text{for all } i. \tag{11}$$

Now, in Lemma 2 below, we will show that, for all i and j, there exists $c_{ij}\geqslant 0$ such that

$$\sum_i q_1(i)[g_2(a_1(i), a_2(j))+c_{ij}]=0 \tag{12}$$

and

$$\sum_j q_2(j)[g_1(a_1(i), a_2(j)) + c_{ij}] = 0. \tag{13}$$

Take

$$p_{ij}(\delta) = \frac{c_{ij}(1-\delta)}{\delta v^*}.$$

Then, if players play (m_1^2, m_2^1) in the first period and switch (forever) to (a_1^*, a_2^*) with probability $p_{ij}(\delta)$ if the outcome is $(a_1(i), a_2(j))$, their expected payoffs are $(0, 0)$ (from (12) and (13)). Furthermore, (12) and (13) imply that the players are indifferent among actions in the supports of their minmax strategies. Q.E.D.

Remark. An immediate corollary of Theorem 3 is that the Folk Theorem holds for two player games of less than full dimension even when mixed strategies are not observable. (This case was not treated in our paper [3].)

LEMMA 2. *Suppose that $B = (b_{ij})$ is an $R \times S$ matrix and that $p = (p(1), ..., p(R))$ and $q = (q(1), ..., q(S))$ are probability vectors such that*

$$p \circ B \leqslant 0 \qquad \text{and} \qquad B \circ q \leqslant 0. \tag{14}$$

Then there exists an $R \times S$ matrix $C = (c_{ij})$ such that $p \circ (B + C) = (B + C) \circ q = 0$ and $c_{ij} \geqslant 0$ for all i, j.

Proof. Consider a row $b_{i.}$ such that $p \circ b_{i.} < 0$. Now if, for all j, $b_{.j} \circ q = 0$, then (14) implies that $p \circ B = 0$, a contradiction of the choice of $b_{i.}$. Hence there exists j such that we can increase b_{ij} while (14) continues to hold. Indeed, we can increase b_{ij} until either $p \circ b_{i.} = 0$ or $b_{.j} \circ q = 0$. Continuing by increasing other elements of B in the same way, we eventually obtain $p \circ B = 0$ and $B \circ q = 0$. Let C be the matrix of the increases that we make to B. Q.E.D.

4. NO PUBLIC RANDOMIZATION

In the proof of Theorem 2, we contructed strategies in which play switches probabilistically from a "punishment" phase to a "reward" phase, with the switching probability chosen to make the punished player's payoff equal to zero. This switch is coordinated by a public randomizing device. The reward phase also relies on public randomization when the vector $v^i = (v_1^i, ..., v_n^i)$ lies outside the set U of payoffs attainable with pure strategies. Although public randomizing devices help to simplify our arguments by convexifying the set of payoffs, they are not essential to Theorem 2. Convexification can also be achieved with deterministic cycles over pure-strategy payoffs when the discount factor is close to one.

Our paper [5] showed that public randomization is not needed for the proof of the perfect equilibrium Folk Theorem, even if players' mixed strategies are not observable. Lemma 2 of that paper established that, for all $\varepsilon > 0$, there exists $\underline{\delta}$ such that, for all vectors $v \in V^*$ with $v_i > \varepsilon$ for i, and for all $\delta > \underline{\delta}$, there is a sequence $\{a(t)\}_{t=1}^{\infty}$, where $a(t)$ is a vector of actions in period t, whose corresponding normalized payoffs are v (i.e., $(1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} g(a(t)) = v$) and for which the continuation payoffs at any time τ are within ε of v (i.e., for all τ, $\|(1-\delta)\sum_{t=\tau}^{\infty} \delta^{t-\tau} g(a(t)) - v\| < \varepsilon$.) This result implies that, for δ large enough so that $\bar{v}_i(1-\delta) < \varepsilon$, we can sustain the vector v^i as the payoffs of a perfect equilibrium, where the equilibrium path is a deterministic sequence of action vectors whose continuation payoffs are always within ε of v^i, and where deviators are punished by assigning them a subsequent payoff of zero. Hence, attaining v^i, even when it does not belong to U, does not require public randomization.

Nor do we need public randomization to devise punishments whose payoffs are exactly zero. We can replace the punishment phase of Theorem 2 with one of deterministic length.

Proof of Theorem 2 *without Public Randomization.* (A) As in the earlier proof, we begin with the case of pure minmax strategies. Let y_j^i, x_i, and $\hat{a}_i$ be as before. Because $0 \in \operatorname{int} V$ we can choose $\varepsilon > 0$ and, for each i, a vector v^i with $v_j^i > 2\varepsilon$ and $v_j^i x_i > y_j^i v_i^i + 2\varepsilon$ for all $j \neq i$. For δ near enough 1, we can choose function $v_i^i(\delta)$ such that

$$\|v_i^i(\delta) - v_i^i\| < \varepsilon, \tag{15}$$

$$v_j^i x_i > y_j^i v_i^i(\delta) + \varepsilon \qquad \text{for all} \quad j \neq i, \tag{16}$$

and such that there exists an integer $t(\delta)$ such that

$$(1 - \delta^{t(\delta)})(-x_i) + \delta^{t(\delta)} v_i^i(\delta) = 0. \tag{17}$$

Let

$$w_j^i = (1 - \delta^{t(\delta)})(-y_j^i) + \delta^{t(\delta)} v_j^i. \tag{18}$$

Substituting using (16) and (17), we have

$$k w_j^i > \varepsilon / x_i, \tag{19}$$

where $k > 0$ is the limit of $\delta^{t(\delta)}$ as $\delta \to 1$.[4] Take δ close enough to one so that for all i and j, $\bar{v}_j(1-\delta) < \min\{\varepsilon, k\varepsilon/x_i\}$.

[4] See [5] for the proof that this limit is strictly positive.

Now consider the following strategies: In state i, play $(\hat{a}_i, m^i_{-i})$ for $t(\delta)$ periods, and then switch to state $n+i$, where play follows a deterministic sequence whose payoffs are $v^i(\delta) = (v^i_i(\delta), v^i_{-i})$. For the play in state $n+i$ we appeal to Lemma 2 of our paper [5], which guarantees that the continuation payoffs in state $n+i$ are at least ε. If player i ever deviates from his strategy, switch to the beginning of state i. By construction, player i's payoff is exactly zero at the beginning of state i, and increases thereafter, and since i is minmaxed in state i, he cannot gain by deviating. If player $j \neq i$ deviates in state i, he can obtain at most $\bar{v}_j(1-\delta)$, and from not deviating obtains at least w^i_j, which is larger. Finally, since payoffs at every date of each reward state are bounded below by ε, and deviations result in a future payoff of zero, no player will wish to deviate in the reward states.

(B) To deal with mixed minmax strategies, we must make player j's payoff in $n+i$ depend on how he played in state i, as in the earlier proof of Theorem 2. We will be a bit sketchier here than before because the argument is essentially the same.

As before, let $-y^i_j(k)$ be player j's expected payoff from his kth best action in the support of m^i_j, $j \neq i$. Let

$$R^i_j = -(1-\delta) \sum_{\tau=0}^{\tau=t(\delta)-1} \delta^\tau [y^i_j(1) - y^i_j(k(\tau))];$$

this is the amount that player j sacrifices in state i relative to always playing $a^i_j(1)$. Now take

$$v^i_j(\delta) \equiv v^i_j + R^i_j(1-\delta)/\delta^{t(\delta)}. \tag{20}$$

With these payoffs in state $n+i$, each player j is indifferent among all the actions in the support of m^i_j. If v^i_i and ε are taken small enough, $t(\delta)$ (as defined by (17)) will also be small, and so the right hand side of (20) will be feasible. Thus, the reward payoffs can be generated by deterministic sequences. Q.E.D.

References

1. D. Abreu, Extremal equilibria of oligopolistic supergames, *J. Econ. Theory* **45** (1986), 191–225.
2. D. Abreu, On the theory of infinitely repeated games with discounting, *Econometrica* **56** (1988), 383–396.
3. D. Fudenberg and E. Maskin, The folk theorem in repeated games with discounting or with incomplete information, *Econometrica* **54** (1986), 533–554.
4. D. Fudenberg and E. Maskin, "Nash and Perfect Equilibria of Discounted Repeated Games," Harvard University Working Paper #1301, 1987.
5. D. Fudenberg and E. Maskin, On the dispensability of public randomization in discounted repeated games, mimeo, 1988.