UC Berkeley
Other Recent Work

Title
NASH and the Perfect Equilibria of Discounted Repeated Games

Permalink
https://escholarship.org/uc/item/7tr3c98t

Authors
Fudenberg, D.
Maskin, E.

Publication Date
1987-04-02

Peer reviewed

UNIVERSITY OF CALIFORNIA, BERKELEY

Department of Economics

Berkeley, California 94720

Working Paper 8736

NASH AND PERFECT EQUILIBRIA
OF DISCOUNTED REPEATED GAMES

D. Fudenberg and E. Maskin

April 2, 1987

Key words: Repeated gains, folk theorem.

Abstract

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 $\delta > \underline{\delta}$.

ACKNOWLEDGMENT
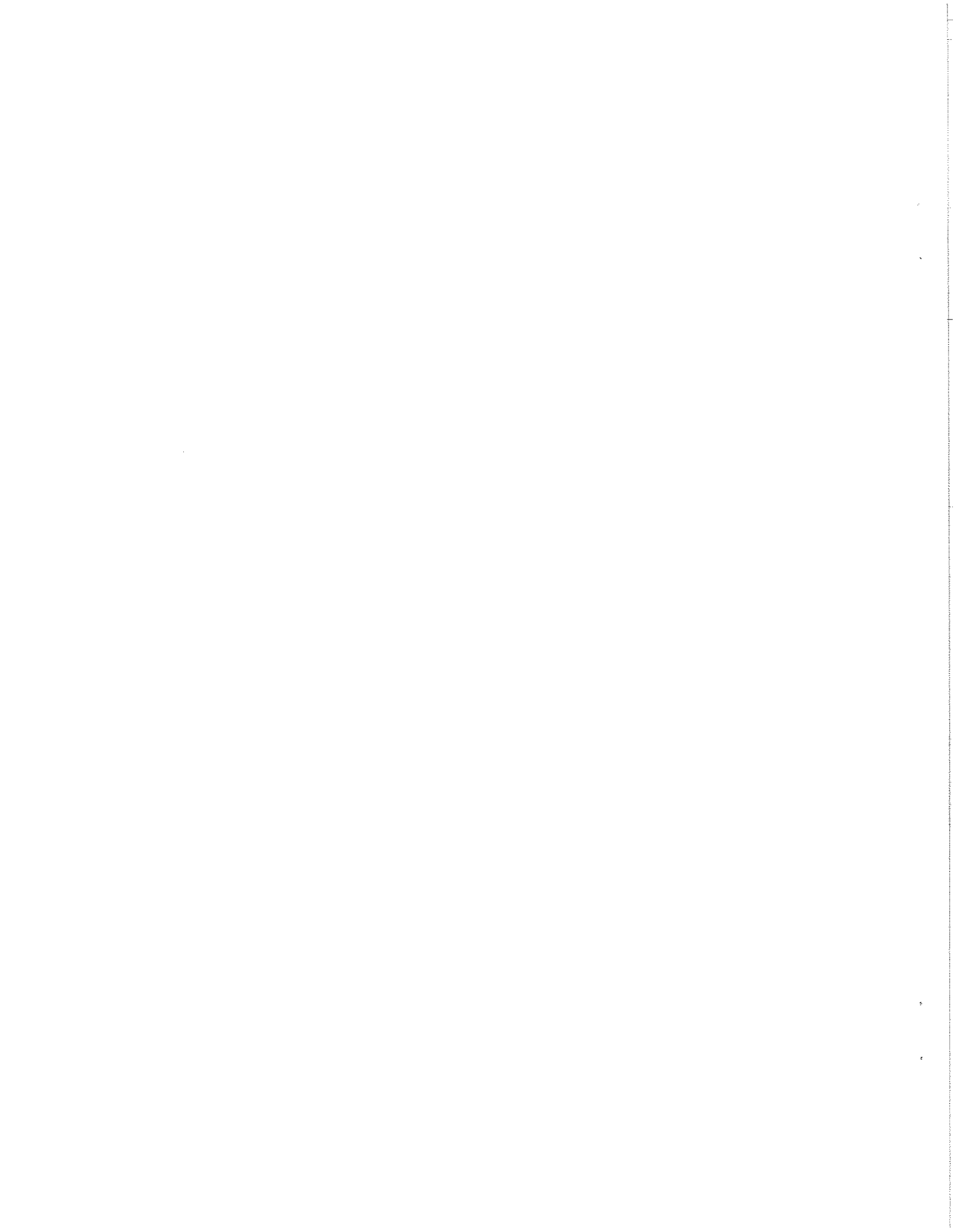
1. Introduction

The "Folk Theorem" for infinitely repeated games with
discounting asserts that any feasible, individually rational payoffs
(payoffs that Pareto dominate the minmax point) can arise as Nash
equilibria if the discount factor δ is sufficiently near one. Our
[1986] paper shows that under a "full-dimensionality" condition 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 set
of equilibrium payoffs. (Even in the limit, perfection does, of
course, rule out some Nash equilibrium strategies.)

This paper shows that when the minmax point is in the interior
of the feasible set and a second, mild condition holds, the Nash and
perfect equilibrium payoffs coincide before the limit. That is, for
any repeated game, there is a $\underline{\delta}<1$ such that for all $\delta\varepsilon(\underline{\delta},1)$, the
Nash and perfect equilibrium payoffs of the δ-discounted game are
identical. Our proof is constructive, and gives an easily computed
expression for the value of $\underline{\delta}$. The payoff-equivalence result holds
even though for any fixed $\delta<1$ there will typically be individually
rational payoffs that cannot arise as equilibria. In other words,
the payoff sets coincide before attaining their limiting values.
Payoff-equivalence is not a consequence of the Folk Theorem; indeed,
it requires the additional conditions that we impose.

The key to our argument is the construction of "punishment
equilibria," one for each player, that hold a player to exactly his
reservation value. As in our other work on discounted repeated
games ([1986], [1986a]), and [1987]), we interweave dynamic

programming arguments and "geometrical" heuristics. The recent
papers by Abreu [1986] and Abreu-Pearce-Stacchetti [1986] use
related techniques, but are concerned with quite different issues.

Section 2 introduces our notation for the repeated game model.
Section 3 presents the main results. Section 4 provides
counterexamples to show that the additional restrictions we impose
are necessary, and that these restrictions are not so strong as to
imply that the Folk Theorem itself obtains for a fixed discount
factor less than one. Through Section 4 we make free use of the
possibility of <u>public</u> randomization. That is, we suppose that there
exists some random variable (possibly devised by the players
themselves), the realizations of which are publicly observable.
Players can thus use the random variable to (perfectly) coordinate
their actions. In Section 5, however, we show that our results do
not require public randomization.

2. <u>Notation</u>

We consider a finite n-player game in normal form

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

where $g(a_1, \ldots, a_n) = (g_1(a_1, \ldots, a_n), \ldots, g_n(a_1, \ldots, a_n))$ and
$g_i(a_1, \ldots, a_n)$ is player i's payoff from the vector of actions
$(a_1, \ldots, a_n)$. As a notational shorthand, we shall often write mixed
strategies as if they were pure. Thus, if $(p_1, \ldots, p_n)$ is a vector
of mixed strategies, and $p_i(a_i)$ is the probability that player i
uses action a_i, we shall write "$g(p_1, \ldots, p_n)$" for

$$\sum_{a_1} \ldots \sum_{a_n} p_1(a_1)\ldots p_n(a_n)g(a_1,\ldots,a_n).$$

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

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

where a(t) is the vector of actions chosen in period t. Player i's action in period t can depend on the past actions of all players and also on the result of a real-valued publicly observable random variable, distributed uniformly on [0,1].

For each player j, choose "minmax strategies" $m^j=(m_1^j,\ldots,m_n^j)$ so that

(0) $m_{-j}^j \in \arg\min_{m_{-j}} \max_{m_j} g_j(m_j,m_{-j})$,

and let

$$v_j^* = \max_{a_j} g_j(a_j,m_{-j}^j) = g_j(m^j).$$

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

We call v_j^* player j's reservation value: it is the lowest payoff to which the other players can hold j. 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 <u>Nash</u> 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 $(v_1^*, \ldots, v_n^*) = (0, \ldots, 0)$. Call $(0, \ldots, 0)$ the <u>minmax point</u>, and take $\bar{v}_i = \max\limits_a g_i(a)$. Let

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

V = Convex Hull of U,

and

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

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, individually rational payoffs.

3. <u>Nash and Perfect Equilibrium</u>

Any feasible vector of payoffs $(v_1, \ldots, v_n)$ 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 $(v_1, \ldots, v_n)$ 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, we may be
able, as in our [1986] paper, to induce players to minmax by
providing "rewards" for doing so.

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

<u>Lemma</u>: 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.

<u>Proof</u>: 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.

<u>Remark 1</u>:   When the hypotheses of the lemma are not satisfied, we would expect the sets of Nash and perfect equilibrium payoffs to differ.

<u>Remark 2</u>:   Note that the lemma does not conclude that all Nash equilibrium <u>strategies</u> are perfect.

A trivial case to which the lemma applies is 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 than to be minmaxed, i.e., a game in which $g_i(m^j) > 0$ for $i \neq j$. In this case we need not reward punishers to ensure their compliance but can simply threaten them with future punishment if they fail to punish an opponent.

<u>Theorem 1</u>:   Suppose that for all i and j, $i \neq j$, m^i_j, as defined by (0), is a pure strategy, and that $g_j(m^i) > 0$. Let $\underline{\delta}$ satisfy

$\bar{v}_i(1 - \underline{\delta}) < \min_j g_i(m^j)$ for all i.   Then for all $\delta \in (\underline{\delta}, 1)$, the sets of Nash and perfect equilibrium payoffs of the repeated game exactly coincide.

<u>Proof</u>:   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 $\bar{v}_i(1-\delta)$,[1] which is less than $g_j(m^i)$ by construction.

So the hypotheses of the lemma are satisfied.

<div align="center">Q.E.D.</div>

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.[2]

Remark 2: The proof of Theorem 1 actually shows that all feasible

payoff vectors that give each player at least $\min_{i \neq j} g_j(m_j)$ 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

----

1. Recall that we are expressing players' payoffs in the
repeated game as average payoffs.
2. 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.

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 Section 4. We must strengthen it to require that the minmax point $(0,\ldots,0)$ is itself in the interior of V. Moreover, we need 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 a $\underline{\delta} < 1$ such that for all $\delta \in (\underline{\delta}, 1)$, the sets of Nash and perfect equilibrium average payoffs of the repeated game exactly coincide.

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

Remark: The full dimensionality condition and hypothesis (ii) of the theorem are 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.

Proof of Theorem 2: For convenience, we will give the proof for the case in which mixed strategies are observable. We explain below how the methods of our [1986] paper can be used to extend this construction to the case in which only the realized outcomes are observable.

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 $y_j^i = -g_j(\hat{a}_i, m_{-i}^i)$. 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 $v^i = (v_1^i, \ldots, v_n^i)$, which are to be determined. If player j cheats in a reward state, switch to punishment state j.

Choose $v^i \gg 0$ in V^* so for $i \neq j$, $v_j^i - v_i^i y_j^i > 0$ (this is possible because $0 \in \text{int } V$). Now set $p_i(\delta) = (1-\delta)x_i/\delta v_i^i$, and choose

$\tilde{\delta} > x_i/(v_i^i + x_i)$, so that for $\delta > \tilde{\delta}$, $p_i(\delta) < 1$. This choice of $p_i(\delta)$ sets player i's payoff starting in state i equal to zero if she plays as specified. Player j's payoff starting in state i if he does not deviate, which we denote w_j^i, solves the functional equation

(1) $\quad w_j^i = (1-\delta)(-y_j^i) + \delta p_i(\delta) v_j^i + \delta(1-p_i(\delta))w_j^i,$

so that

(2) $\quad w_j^i = (v_j^i - y_j^i v_i^i)/(v_i^i + x_i).$

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

We claim that for all $\delta \varepsilon(\underline{\delta}, 1)$, the specified strateiges are a perfect equilibrium. First consider punishment state i. In this state, player i receives average 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 an average payoff of zero from tomorrow on. Thus player i cannot gain by a single deviation, and the usual dynamic programming argument shows that she cannot gain by multiple deviations. Player j's average payoff in state i is w_j^i(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 average payoff is zero, so player j cannot profit from deviating in state i. Finally, in reward state n+i, each player k obtains an

average payoff v_k^i exceeding $\bar{v}_k(1-\delta)$, and so the threat of switching to punishment state k prevents deviations. The theorem now follows from our lemma.

<div align="center">Q.E.D.</div>

The proof of Theorem 2 involves punishing player j if, in state i, he fails to use his minmax strategy m_j^i. Of course, m_j^i may be a mixed strategy, and so implicitly we have assumed that such strategies are observable (i.e., that player j's private randomizing device is observable). Following our [1986] article, however, we can readily modify the argument so that this asumption is not needed. Specifically, suppose that we make player j's reward in state n+i dependent on the actions he took in state i. Thus, if $\hat{a}_j$ and $\hat{\hat{a}}_j$ are both in the support of m_j^i and player j obtains a higher one-shot payoff from the former than the latter when the other players use m_{-j}^i, we can reward him more (once state n+i is reached) for using $\hat{\hat{a}}_j$ in such a way that he is exactly indifferent between the two actions (the possibility of doing this without affecting other players' rewards depends on full dimensionality). Similarly, by a judicious choice of rewards, we can make player j indifferent among all actions in the support of m_j^i, so that he is willing to play this mixed strategy after all (for further details, see Section 6 of the [1986] paper).

4. Counterexamples

We turn next to a series of examples designed to explore the roles of the hypotheses of Theorems 1 and 2. Three of these examples demonstrate that none of the hypotheses can be dropped without rendering the theorems false.

Example 1 shows that Theorem 1 need not hold when the minmax strategies are mixed. Example 2 shows 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. Examples 3 and 4 show that both hypotheses (i) and (ii) of Theorem 2 are necessary.

Several of the examples make use of the following facts: (a) for any fixed δ, the sets of Nash and perfect equilibrium payoffs are closed; and (b) if a player's equilibrium payoff is exactly zero, he must be minmaxed in the first period (otherwise he could obtain a positive payoff in the first period, and, because he cannot subsequently be held below zero, he would thus have a positive payoff overall.)

Example 1

Consider the game in Table 1

13

Player 2

	L	R
U	1,−1	−1,2
D	−1,−1	1,0

Player 1

Table 1

Note that player 1 is minmaxed if and only if player 2 mixes with equal probabilities between L and R. When minmaxed, player 1's payoff is zero regardless of how he plays, but if he plays U with more than probability one-half, player 2's payoff is positive. Moreover, Player 1's payoff is positive when he minmaxes 2 (by playing D) and 2 responds optimally (by playing R). Thus, the hypotheses of Theorem 1 are satisfied except for the assumption that 2's minmax strategy be pure.

To show that the conclusion of Theorem 1 does not hold, we first establish two preliminary claims.

Claim 1: In any Nash equilibrium, player one's payoff must be positive.

Proof: Suppose that there is a Nash equilibrium where player 1's payoff is zero. Then 1 must be minmaxed in the first period, which, in this game, means that his first-period payoff must be zero. Thus, his payoff from period two onwards must be zero, implying that he must be minmaxed in the second period, and, from induction, in every subsequent period as well. This requires player 2 to play L

with probability $\frac{1}{2}$ in every period. But L is dominated by R, and so such a strategy cannot be optimal for 2.

<div align="right">Q.E.D.</div>

Now let $v^n(\delta)$ be the smallest payoff for player 1 in a Nash equilibrium (of the δ-discounted repeated game) where player 2's payoff is zero, and let $v^p(\delta)$ be the lowest payoff for player 1 in a <u>perfect</u> equilibrium where player 2's payoff is zero.

<u>Claim 2</u>: There exists a $\underline{\delta}<1$ such that $v^n(\delta)\le1-\delta$ for all $\delta\varepsilon(\underline{\delta},1)$.

<u>Proof</u>: We will construct a Nash equilibrium that attains the desired bound. The strategies have three phases along the equilibrium path. In the initial phase, players play (D,L). At the end of each period, play switches to phase 2 with probability p_1; play remains in the initial phase with complementary probability $1-p_1$. In phase 2, players play (U,R), and there is a probability p_2 of switching to phase 3, in which players play (D,R) forever. If any deviations occur, players minmax each other forever.

Let $\hat{v}_1$ be player one's (average) continuation payoff at the beginning of phase 2 in this equilibrium. If we take

$$p_1 = \frac{(1-\delta)(2-\delta)}{\delta(\hat{v}_1-(1-\delta))} \qquad \hat{v}_1 = \frac{\delta p_2-(1-\delta)}{1-\delta(1-p_2)}$$

and

$$P_2 = \frac{(1-\delta)(6-3\delta)}{\delta^2},$$

it can be verified that player one's overall payoff is $1-\delta$, and player two's is 0. The above probabilities lie between 0 and 1 for all δ greater than some $\underline{\delta} < 1$.

<div align="right">Q.E.D.</div>

Now we can show that $v^n(\delta) < v^p(\delta)$ for all $\delta > \underline{\delta}$. Suppose to the contrary that $v^n(\delta) = v^p(\delta) = v^*$ for some $\delta > \underline{\delta}$. Fix a perfect equilibrium s^* that yields payoffs $(v^*, 0)$. We will show how to construct a Nash equilibrium with payoffs $(\hat{v}, 0)$, $\hat{v} < v^*$.

In the perfect equilibrium s^*, player one must play D with probability one in the first period (or else player two could obtain a positive payoff). Let q^* be the probability that player two plays L in the first period, let x_1 be player one's expected payoff from period two on if there are no deviations from equilibrium, and let w_1 be his expected payoff from period 2 on if he deviates in the first period. For player one to prefer D to U in the first period, we must have

(3) $v^* = (1-\delta)(1-2q^*) + \delta x_1 \geq (1-\delta)(2q^*-1) + \delta w_1.$

In a perfect equilibrium, w_1 must be player one's payoff in some Nash equilibrium, and thus, by Claim 1, w_1 must be strictly positive. If $q^*=1$, the right side of inequality (3) exceeds $1-\delta$, a contradiction of Claim 2. If $q^*=0$, then x_1 would have to be zero (otherwise, v^* would exceed $1-\delta$), which is a contradiction of Claim

1. Therefore $0 < q^* < 1$.

Now consider the strategy pair $\hat{s}$ that differs from s^* in the following two ways. First, if player one deviates, he is minmaxed forever, and, second, the probability that player two plays L in the first period is $\hat{q}$, $q^* < \hat{q} < 1$. Since player two randomizes in the first period under s^*, he is indifferent between his choices, and changing his randomizing probabilities will not change his payoff. Thus, player two's payoff under $\hat{s}$ is zero, and he will not choose to deviate.

Player one's payoff under $\hat{s}$ is $\hat{v} = (1-\delta)(1-2\hat{q}) + \delta x_1$, which is less than v^* (hence, $\hat{s}$ is not a _perfect_ equilibrium). For $\hat{q}$ near q^*, $\hat{v}$ will exceed $(1-\delta)(2\hat{q}-1)$, and so $\hat{s}$ is a _Nash_ equilibrium in which player two's payoff is zero, and player one's is less than v^*. We conclude that for all $\delta > \underline{\delta}$, the Nash and perfect equilibrium payoff sets fail to coincide.

Example 2

Consider the game in Table 2. It satisfies the assumptions of both Theorems 1 and 2, so that for sufficiently large discount factors the Nash and perfect equilibrium payoffs coincide.

	L	M	R
U	1,0	2,-1	-2,-1
D	3,0	1,2	0,1

Table 2

Now, the feasible point (3,0) is contained in the limit of the Nash

equilibrium payoff sets as δ tends to 1. However, for any fixed discount factor $\delta < 1$, the Nash equilibrium payoffs are bounded away from (3,0). To see this observe that if, for $\delta < 1$, there existed a Nash equilibrium with average payoffs (3,0), players would necessarily play (D,L) every period. But in any given period, player two could deviate to M and thereby obtain a positive payoff, a contradiction.

We next show that the interiority assumption of Theorem 2 is necessary.

Example 3

Consider the two-player game in Table 3, in which (0,0) lies on the boundary of V instead of the interior.

	L	R
U	0,0	-1,-4
D	2,1	0,-5

Table 3

Observe first that player one's payoff cannot be exactly zero in a Nash equilibrium. The only way that player one's payoff could be zero along a path in which player 2's payoff is nonnegative would be for players to choose (U,L) with probability one in every period. But then player one could deviate to D in the first period and guarantee himself $2(1-\delta)$.

We will show that for $\delta > 9/14$, there is a Nash equilibrium in

which player one's payoff is $2(1-\delta)$, but that the lowest player one's payoff can be in a <u>perfect</u> equilibrium is $7(1-\delta)/3$. To show the latter we first prove that $14(1-\delta)/9$ is a lower bound for player one's Nash equilibrium payoffs.

Fixing a Nash equilibrium, let p be the probability that player one plays U in the first period, let q be the probability that player two plays L in the first period, and, for $i=1,2$, let v_i be player i's equilibrium payoff, w_i his expected first-period payoff, and x_i his expected payoff from period 2 on.

Computation yields

$$(4) \quad \begin{cases} w_2 = 5q-pq-4-(1-q)(1-p), \\[2mm] w_1 = w_2+1-qp+2(1-q)(1-p)+2(1-q). \end{cases}$$

The individual-rationality requirement $v_2 \geq 0$ implies that $x_2 \geq -(1-\delta)w_2/\delta$, and since $v_1 \geq 2v_2$ for all $v \varepsilon V^*$, we have $x_1 \geq -2(1-\delta)w_2/\delta$. This implies

$(5) \quad v_1 \geq (1-\delta)(w_1-2w_2),$

and, from formula (4),

$(6) \quad v_1 \geq (1-\delta)[7(1-q)+3(1-p)(1-q)] \geq 7(1-\delta)(1-q).$

In words, if q is small, player two has a large negative payoff in period 1 and so must be rewarded in the future to ensure that he recieves his reservation value. Given the geometry of the feasible set, rewarding player two requires rewarding player one more, so that if player one's payoff is to be small, q must be large. However, as we will see, a large q conflicts with player one's

incentive constraint.

If player one plays D in the first period, he obtains 2q immediately, and at least zero in future periods, so that

(7) $v_1 \geq 2(1-\delta)q.$

Combining (6) and (7), we have $v_1 \geq 14(1-\delta)/9$, as claimed.

In a perfect equilibrium, player one's worst possible continuation payoff is thus no less than $14(1-\delta)/9$, so that by playing D in the first period player one can guarantee at least $2q(1-\delta)+\delta 14(1-\delta)/9$.   Combining this bound with (6), we have

(8) $v_1 \geq (1-\delta) \min_{q} \max \{7(1-q), 2q+\delta 14/9\}.$

For $\delta>9/14$, this expression is minimized by $q \leq 2/3$, so that $v_1 \geq 7(1-\delta)/3$.

Finally, we must show that player one's payoff can in fact be less than $7(1-\delta)/3$ in a Nash equilibrium.   Consider the following strategies: Play begins in Phase (A), in which players play (U,L). If there have been no deviations in Phase (A), players jointly switch to (D,L) forever with probability $p=(1-\delta)/\delta$, remaining in Phase (A) with the complementary probability.   If there are any deviations, the deviator is minmaxed forever.   These strategies form a Nash equilibrium in which player one's payoff is $2(1-\delta)$.

Our final example establishes that the conclusion of Theorem 2 need not hold if some player cannot obtain less than his reservation value when being minmaxed by the others.

<u>Example 4</u>

Consider the game in Table 4.

	L	R
U	0,-1	-1,0
D	0,-1	1,1

Table 4

When player two plays L, player one obtains zero regardless of what he does. Note that the origin is in the interior of V. As in Example 3, player one's payoff cannot be exactly zero in a Nash equilibrium: If player two plays R with positive probability in any period, player one can obtain a strictly positive payoff by playing D in that period, and it is not individually rational for player two to always play L. However, for $\delta > 1/2$ player two's payoff can be zero, even in a <u>perfect</u> equilibrium, as shown by the following strategies: Play begins in Phase A, where players choose (U,L). If neither player deviates, play switches to Phase B with probability $(1-\delta)/\delta$ and remains in Phase A with the complementary probability $(2\delta-1)/\delta$. In Phase B, players take (D,R) forever. If there are any deviations, they start over. It is easy to see that player one cannot gain by deviating. The switching probability has been chosen to set player two's equilibrium payoff to zero, so that in Phase A player two is indifferent between playing L or R.

For each $\delta > 1/2$, let $v^n(\delta)$ be player one's lowest payoff among Nash equilibria of the δ-discounted game where player two's payoff

is zero. That is, $v^n(\delta) = \min_{v_1} \{v_1 \mid (v_1, 0) \in N(\delta)\}$, where $N(\delta)$ is the set of Nash equilibrium payoffs (because $N(\delta)$ is closed, the minimum is attained). Notice that if player one plays D with positive probability in the first period, player two can obtain a strictly positive payoff by playing R forever. Thus in any Nash equilibrium that attains payoffs $(v^n(\delta), 0)$, player one must play U with probability one in the first period.

We can show that $v^n(\delta) < 2(1-\delta)/3$ with a slight modification of the strategies we introduced to show that player two's payoff could be zero in a perfect equilibrium. As before, play begins in Phase A. In this phase, player one plays U, and player two plays L with probability $q = (2-\delta)/2$ (in the earlier equilibrium, q was equal to one). If the Phase A outcome is (U,L), then as before play switches to Phase B with probability $(1-\delta)/\delta$, and play remains in Phase A with the complementary probability $(2\delta-1)/\delta$. In Phase B, players choose (D,R) forever. If the Phase A outcome is (U,R), play remains in Phase A. If player one deviates by playing D instead of U, then player two plays L for the rest of the game. Deviations from (D,R) are ignored.

Player two is willing to randomize in Phase A because his expected total payoff from either pure action is zero. Phase B consists of one-shot Nash equilibrium play, which is self-enforcing. Let us check that player one will not deviate in Phase A, and that his payoff is less than $2(1-\delta)/3$. If player one plays as specified, his payoff v_1 satisfies the functional equation

(9) $v_1 = q[(1-\delta)\cdot 0+(1-\delta)+(2\delta-1)v_1] + (1-q)[-(1-\delta)+\delta v_1]$,

so that

(10) $v_1 = (2q-1)/(1+q) = 2(1-\delta)/(4-\delta) < 2(1-\delta)/3$.

If player one plays D in the first period, he obtains $(1-\delta)(1-q)=$ $(1-\delta)\delta/2$, which is less than v_1. So the prescribed play forms an equilibrium, and thus $v^n(\delta)<2(1-\delta)/3$.

Let $v^P(\delta)$ be player one's lowest payoff in a perfect equilibrium that holds player two's payoff to zero, and let $\underline{v}(\delta)$ be player one's lowest payoff in any Nash equilibrium, including those where player two's payoff is positive (recall that $\underline{v}(\delta)>0$ for all δ). We will show that, for all $\delta>1/2$, $v^P(\delta)>v^n(\delta)$. To do so, suppose to the contrary that, for some δ, $v^P(\delta)=v^n(\delta)=v^*$, and fix a perfect equilibrium s^* that yields $(v^*,0)$. Contrary to our supposition, we will construct a Nash equilibrium with payoffs $(\hat{v},0)$, where $\hat{v}<v^*$.

In the perfect equilibrium s^*, let q^* be the probability that player two plays L in the first period (recall that player one must play U with probability one) and let x_i^L and x_i^R be the expected payoffs for player i from period two on if L or R occurs in the first period. Player one can obtain $1-q^*$ in the first period by playing D, and cannot receive less than $\underline{v}(\delta)$ in the future, because perfection requires that future play be Nash. This gives us the "perfect equilibrium incentive constraint"

(11) $v^* = -(1-\delta)(1-q^*)+\delta q^* x_1^L+\delta(1-q^*)x_1^R \geq (1-\delta)(1-q^*)+\delta\underline{v}(\delta)$.

If player two plays R in the first period, he receives zero today, so, for two's equilibrium payoff to be zero, the continuation

payoff x_2^R must equal zero as well. Since $v^*=v^P(\delta)$ is player one's lowest perfect equilibrium payoff consistent with two's payoff being zero, we have $x_1^R \geq v^*$.

We will first argue that we need only consider the case $x_1^R=v^*$. If $x_1^R>v^*$ and $q^*=1$, then we construct new strategies $\hat{s}$ which agree with s^* except that, if player two plays R in the first period, the continuation payoffs are $(\hat{x}_1^R, 0)$, where $\hat{x}_1^R=v^*$. The only effect of this substitution is to lower player one's payoff if player <u>two</u> deviates, so $\hat{s}$ is a perfect equilibrium. We can thus consider $\hat{s}$ rather than s^*. If $x_1^R>v^*$ and $q^*<1$, we construct new strategies $\hat{s}$ that differ from s^* in two ways. First, the continuation payoffs after two plays R in the first period are $(x_1^R/(1+\theta), 0)$, for some small positive θ. These payoffs can be attained by a public randomization between $(x_1^R, 0)$ and $(v_1^*, 0)$. Second, if player one plays D in the first period, he is minmaxed for the rest of the game. Because player two does not gain by deviating from s^*, he will not wish to deviate from $\hat{s}$ along the equilibrium path. Player one's payoff, $\hat{v}$, from $\hat{s}$ is less than v_1^*, and so may violate the perfect equilibrium incentive constraint $\hat{v} \geq (1-\delta)(1-q^*)+\delta\underline{v}(\delta)$. However, for small θ, $\hat{v}$ is near v^*, and so, since $\underline{v}(\delta)>0$, $\hat{v}$ exceeds $(1-\delta)(1-\hat{q})$. Thus $\hat{s}$ is a <u>Nash</u> equilibrium that gives player two a zero payoff and gives player one less than v^*, a contradiction. If $q^*<1$, therefore, we must have $x_1^R=v^*$.

Consider the case $x_1^R=v^*$. Player one's overall payoff if player two plays R in the first period is $v_1^R \equiv -(1-\delta)+\delta v^* < 0$, since, by

assumption $v^* = v_1^n(\delta) = x_1^R$, and $v_1^n(\delta) \leq 2(1-\delta)/3$. Because $v_1^R < 0$, $q^* > 0$; otherwise, player one's equilibrium payoff would be negative. The nonnegativity constraint also implies that player one's overall payoff if player two plays L in the first period must be positive. Now, let us construct new strategies $\hat{s}$ which differ from s^* only in that $\hat{q} < q^*$. This change lowers player one's equilibrium payoff $\hat{v}$, and so, as before, the perfect equilibrium incentive constraint need not hold. However, for $\hat{q}$ near q^*, $\hat{v}$ exceeds $(1-\delta)(1-\hat{q})$, so $\hat{s}$ is a Nash equilibrium with payoff $\hat{v} < v^*$, a contradiction.

We conclude that $v^n(\delta) < v^P(\delta)$ after all. Thus for all $\delta > 1/2$, the Nash and perfect equilibrium payoff sets differ.

5. No Public Randomization

In the proof of Theorem 2, we constructed strategies in which play switches probabilistically from a "punishment" phase to a "reward" phase, with the switching probability chosen so as to make the punished player's average payoff exactly equal to zero. This switch was coordinated by the players' observation of a public randomizing device. Also, our strategies specified that play in the ith reward phase yield a specific vector $v^i = (v_1^i, \ldots, v_n^i)$ of positive utilities. This aspect of the strategies relied on public randomizations to overcome the difficulty that v^i need not lie in U, the set of payoffs attainable with pure strategies. Public randomizing devices are sometimes available, and are very helpful in simplifying the type of dynamic programming-plus-geometry arguments of the sort we have developed. However, we will show in this

section that public randomization is not essential to the proof of Theorem 2.

Our [1987] paper showed that public randomizations are not needed for the proof of the perfect Folk Theorem, even if each player's privately mixed strategies are not observable. Lemma 2 of that paper established that for any vector $v \in V^*$ and every $\alpha > 0$, there exists, for δ close enough to 1, a sequence $\{a(t)\}_{t=1}^{\infty}$ of vectors of actions (where $a(t)$ is the vector of actions in period t) whose corresponding discounted average payoffs are v (i.e.,

$(1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} g(a(t)) = v$) and such that, the continuation payoffs from any time τ on are within α of v (i.e., for all τ,

$|(1-\delta) \sum_{t=\tau}^{\infty} \delta^{t-\tau} g(a(t)) - v| < \alpha$). This result implies that, for δ near 1, we can sustain the vector v^i as the average payoffs of a perfect equilibrium where the equilibrium path is a deterministic sequence of action vectors and where deviators are punished by assigning them a subsequent average payoff of zero. Hence, the problem that v^i may not belong to U does not require public randomization as a cure.

Nor do we need public randomization to ensure that player i can be punished in such a way that his average payoff is zero. To see this, choose $\hat{a}_i$, as above, so that $g_i(\hat{a}_i, m_{-i}^i) = -x_i < 0$ (we are assuming the hypotheses of Theorem 2). For $j \neq i$, let $y_j^i = -g_j(\hat{a}_i, m_{-j}^i)$. Because $0 \in \text{int } V$, we can choose the vector $v^i \in V^*$ such that

(12) $v_j^i > \dfrac{y_j^i}{x_i} v_i^i$ for all $j \neq i$.

Now, choose δ close enough to 1 and an integer $t(\delta)$ so that

(13) $v_j^i > \bar{v}_i (1-\delta)$

and

(14) $(1-\delta^t)(-x_i) + \delta^{t(\delta)} v_j^i = 0$.

Since (14) implies that $\delta^{t(\delta)}$ cannot tend to 0 as δ tends to 1,

formulas (13) and (14) together imply that

$(1-\delta^t)(-y_j^i) + \delta^t v_j^i > 0$, $j \neq i$.

Indeed, for δ near enough 1, we have

(15) $(1-\delta^t)(-y_j^i) + \delta^t v_j^i > (1-\delta)\bar{v}_j$, $j \neq i$.

But (13)-(15) imply that player i can be punished by players

choosing $(\hat{a}_i, m_{-i}^i)$ for t periods followed by their switching to v^i.

As in the proof of Theorem 2, each player is deterred from deviating

from this path by the threat of a similar punishment. Notice, from

(14), that player i's average payoff is 0, and so we have eliminated

the need for punishments of random length.

References

Abreu, D. [1986], "Extremal Equilibria of Oligopolistic Supergames," Journal of Economic Theory, 39, 191-225.

_____, D. Pearce, and E. Stachetti [1986], "Optimal Cartel Equilibria with Imperfect Monitoring," Journal of Economic Theory, 39, 251-269.

Fudenberg, D. and E. Maskin [1986], "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," Econometrica, 54, 533-554.

_____ and _____ [1986a], "Discounted Repeated Games with Unobservable Actions, I: One-Sided Moral Hazard," mimeo.

_____ and _____ [1987], "Public Randomization in Discounted Repeated Games," mimeo.

RECENT ISSUES OF THE WORKING PAPER SERIES
OF THE DEPARTMENT OF ECONOMICS
UNIVERSITY OF CALIFORNIA, BERKELEY

Copies may be obtained from the Institute of Business and Economic
Research. See the inside cover for further details.

8730   Richard J. Gilbert and Marvin Lieberman
        INVESTMENT AND COORDINATION IN OLIGOPOLISTIC INDUSTRIES

        Feb-87.

8731   Paul Milgrom and John Roberts
        BARGAINING AND INFLUENCE COSTS AND THE ORGANIZATION OF ECONOMIC ACTIVITY

        Feb-87.

8732   Paul Milgrom
        AN ESSAY ON PRICE DISCRIMINATION

        Feb-87.

8733   Lloyd Ulman
        INDUSTRIAL RELATIONS

        Feb-87.

8734   Dale J. Poirier and Paul A. Ruud
        PROBIT WITH DEPENDENT OBSERVATIONS

        Mar-87.

8735   Pranab Bardhan
        ALTERNATIVE APPROACHES TO DEVELOPMENT ECONOMICS: AN EVALUATION

        Mar-87.

8736   D. Fudenberg and E. Maskin
        NASH AND PERFECT EQUILIBRIA OF DISCOUNTED REPEATED GAMES

        Apr-87.

RECENT ISSUES OF THE WORKING PAPER SERIES
OF THE DEPARTMENT OF ECONOMICS
UNIVERSITY OF CALIFORNIA, BERKELEY

Copies may be obtained from the Institute of Business and Economic
Research. See the inside cover for further details.

8723   Pranab Bardhan and Nirvikar Singh
       MULTINATIONAL RIVALRY AND NATIONAL ADVANTAGE: SOME THEORETICAL CONSIDERATIONS

       Jan-87.

8724   Albert Fishlow
       LESSONS OF THE 1890S FOR THE 1980S

       Jan-87.

8725   Jeffrey A. Frankel
       THE SOURCES OF DISAGREEMENT AMONG INTERNATIONAL MACRO MODELS,
       AND IMPLICATIONS FOR POLICY COORDINATION
       Jan-87.

8726   Jeffrey A. Frankel and James H. Stock
       REGRESSION VS. VOLATILITY TESTS OF THE EFFICIENCY
       OF FOREIGN EXCHANGE MARKETS
       Jan-87.

8727   Joseph Farrell and Carl Shapiro
       DYNAMIC COMPETITION WITH LOCK-IN

       Jan-87.

8728   James A. Mirrlees
       ECONOMIC POLICY AND NONRATIONAL BEHAVIOUR

       Feb-87.

8729   Jeffrey A. Frankel
       THE IMPACT OF U.S. ECONOMIC POLICIES ON A COMMODITY-EXPORTING DEBTOR:
       THE CASE OF THAILAND
       Feb-87.