MASSACHUSETTS INSTITUTE of TECHNOLOGY
LIBRARIES

Digitized by the Internet Archive
in 2011 with funding from
Boston Library Consortium Member Libraries

http://www.archive.org/details/ondispensability00fude

HB31
.M415
no.467

DEWEY
M.I.T. LIBRARIES

working paper
department
of economics

ON THE DISPENSABILITY OF PUBLIC RANDOMIZATION
IN DISCOUNTED REPEATED GAMES

Drew Fudenberg
Eric Maskin

Number 467 August 1987

massachusetts
institute of
technology
50 memorial drive
cambridge, mass. 02139

ON THE DISPENSABILITY OF PUBLIC RANDOMIZATION IN DISCOUNTED REPEATED GAMES

Drew Fudenberg
Eric Maskin

Number 467 August 1987

Dewey

MASS. INST. TECH.
MAY 23 1988
LIBRARIES

On the Dispensability of Public Randomization in Discounted Repeated Games

by

Drew Fudenberg*

and

Eric Maskin**

August 1987

* Massachusetts Institute of Technology

** Harvard University

We would like to thank David Levine for encouraging us to study this topic. Research support from National Science Foundation Grants SES-8607229 and SES-8520952 is gratefully acknowledged.

2009829

1. Introduction

The Folk Theorem for repeated games asserts that any feasible, individually rational payoffs for a one-shot game can arise as Nash equilibrium average payoffs when the game is infinitely repeated. In our [1986] paper, which extends this result to subgame perfect equilibrium and discounting, we assumed that the players can condition their play on the realization of a publicly observed random variable. We asserted, however, that abandoning the assumption would lead to only a slight weakening of the results; viz., any feasible, individually rational payoffs can be approximated by a perfect equilibrium where there is sufficiently little discounting. This note shows that, in fact, our extension of the Folk Theorem holds in a strong sense even without public randomization: all feasible individually rational payoffs can be exactly attained in equilibrium.

Although this stronger result is of some interest by itself, its true significance appears in connection with mixed strategies. Early analyses of repeated games with little or no discounting (Aumann-Shapley [1976], Friedman [1971] and Rubinstein [1979]) restricted players to pure strategies, or equivalently, assumed that a player's choice of a *mixed* strategy in any period is observable by his fellow players. The assumption of pure strategies is restrictive because typically the range of individually rational payoffs is greater when players are allowed to use mixed strategies to punish their opponents. The alternative hypothesis--that a player's randomizations are *ex post* observable--is, likewise, strong.

Section 6 of our [1986] paper showed how to extend the Folk Theorem to allow for mixed strategies when only a player's realized actions, and not his choices of randomizing probabilities, are observable. The key was the observation that a player can be induced to use a mixed strategy to minimax an opponent by making her continuation payoff depend on her current action in a way that renders her exactly indifferent among the various choices in the mixed strategy's support.

Our argument relied on public randomization to ensure that any individually rational continuation payoffs can be exactly attained. If, without public randomization, the continuation payoffs could merely be approximated, a minimaxing player might not be exactly indifferent over the support of his mixed strategy, and our construction would fail. Thus, if we obtain only an approximate version of the Folk Theorem without public randomization, our construction cannot accommodate unobservable mixed strategies.

Attaining payoffs exactly is also essential for the argument in our [1987a] paper, which provided sufficient conditions for the sets of Nash and perfect equilibrium payoffs to coincide for discount factors less than one. Although the body of that paper assumed the possibility of public randomization, our results here imply that this assumption, as in the Folk Theorem paper, is unnecessary.

2. The Model

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

$g: A \to R^n$,

where $A = A_1 \times \ldots \times A_n$ and A_i is player i's action space. Let Σ_i be the set of player i's mixed strategies, i.e., the probability distributions over A_i, and set $\Sigma=\Sigma_1\times\ldots\times\Sigma_n$. To simplify notation, we will write $g_i(\sigma)$ for player i's payoff given the mixed strategy vector $\sigma\in\Sigma$.

In repeated versions of g, each player's probability mixture over actions at time t can depend on the actions chosen at all previous times. More formally, let $h(t) \in A^{t-1} \equiv H(t)$ be the realized actions from time zero through time t-1. Player i's strategy is a sequence of maps (one for each period) from H(t) to Σ_i. Note that, at any time t, player i's strategy does not depend on the past randomizing probabilities of his opponents, but only on their realized actions.

In the infinitely repeated game G_δ, each player i's payoff is the average discounted sum π_i of his per-period payoffs, with common discount factor δ:

$$\pi_i \equiv (1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} g_i(\sigma(t)),$$

where $\sigma(t)$ is the probability distribution of actions chosen in period t.

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

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

and

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

(Here "m_{-j}^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. Clearly, player j's average payoff must be at least v_j^* in any equilibrium of g, whether or not g is repeated.

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>minimax point</u>. Take

$\bar{v}_i = \max_a g_i(a)$. Moreover, let

$$U = \{(v_1,\ldots,v_n) \mid \text{there exists } \sigma\in\Sigma \text{ with } g(\sigma)=(v_1,\ldots,v_n)\},$$

$$V = \text{Convex Hull of } U,$$

and

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

3. The Folk Theorem without Public Randomization

Our [1986] paper showed that if public randomization is allowed and either n=2 or the dimension of V^* equals n, then for any payoff vector $v\in V^*$, there exists a discount factor $\underline{\delta}<1$ such that, for all $\delta\in(\underline{\delta},1)$, there is a perfect equilibrium of G_δ with payoffs v. We now demonstrate that public randomization is inessential for this result.[1]

Lemma 1 establishes that, for δ sufficiently large, all points in V are feasible and can be obtained without using mixed strategies. That is, for any $v\in V$ there is a deterministic sequence of actions $\{a(t)\}_{t=1}^{\infty}$ for which v is the payoff vector. This is not sufficient to establish the Folk Theorem, however, because, even if

----

1. Of course, for low discount factors, public randomization does make a difference. If δ is near zero, the payoff vector for the sequence $\{\sigma(t)\}$ is approximately $g(\sigma(1))$, and so, quite apart from equilibrium considerations, many payoffs in V^* are not feasible.

$v \in V^*$, the sequence $\{a(t)\}$ might have the property that, for some period τ, the continuation payoffs beginning at τ do not belong to V^*. In that case, some player would prefer to deviate from the sequence, even if so doing caused his opponents to minimax him thereafter.

Building on Lemma 1, Lemma 2 shows that payoffs in V^* can be generated by a deterministic sequence in such a way that the continuation payoffs always lie in V^*. Following Lemma 2, we explain how our results allow us to do without public randomization in the proof of the Folk Theorem.

Write $A=\{a^1,\ldots,a^m\}$ and, for each j, let $w^j = g(a^j)$. Thus, $\{w^1,\ldots,w^m\}$ is the set of payoff vectors corresponding to <u>pure</u> strategies.

<u>Lemma 1</u>: If $\delta > 1-\frac{1}{m}$, then for any $v \in V$ there is a sequence $\{a(t)\}$ of pure strategies whose average payoff is v.

<u>Proof</u>: Let $v=\Sigma\lambda^j w^j$, where $0 \le \lambda^j \le 1$, and $\sum_{j=1}^{m} \lambda^j=1$. We construct $\{a(t)\}$ as follows. Let $I^j(t)$ be an index variable, which is 1 if $a(t)=a^j$ and 0 otherwise. Set $N^j(1)=0$ for all j, and let $N^j(t) = \sum_{\tau=1}^{t-1} (1-\delta)\delta^{\tau-1} I^j(\tau)$ for $t>1$. $N^j(t)$ is the "average discounted weight" given to strategy vector a^j before time t. Let $C(t) = \{j \mid \lambda^j - N^j(t) > \delta^{t-1}(1-\delta)\}$. Now define

$$j^*(t) = \arg\max_{j \in C(t)} \{\lambda^j - N^j(t)\},^2$$

and set $a(t)=a^{j^*(t)}$. This defines an algorithm for computing a(t).

----

2. If there is a tie, make a deterministic selection.

Claim 1: The algorithm is well-defined, i.e., the set C(t) is never empty.

To prove the claim, assume to the contrary that at some time(s) t, C(t) is empty, and let s be the first such time. Then $\delta^{s-1}(1-\delta)\geq\lambda^j-N^j(s)$ for all j. Summing over j, we have

$$(1)\quad m(1-\delta)\delta^{s-1}\geq 1-\sum_{j=1}^{m}N^j(s) = 1-\sum_{j=1}^{m}\sum_{\tau=1}^{s-1}(1-\delta)\delta^{\tau-1}I^j(\tau)$$

$$= 1-\sum_{\tau=1}^{s-1}(1-\delta)\delta^{\tau-1} = \delta^{s-1}.$$

But (1) contradicts our assumption that $m(1-\delta)<1$, establishing the claim.

Let $N^j(\infty) \equiv \lim_{t\to\infty} N^j(t)$. (Because $N^j(t)$ is increasing and bounded, this limit exists.)

Claim 2: For all j, $N^j(\infty)=\lambda^j$.

To establish Claim 2, note first that, by construction, $\sum_{j=1}^{m}N^j(\infty)=1$. Moreover, $N^j(\infty)$ cannot exceed λ^j, because N^j increases (by $\delta^{t-1}(1-\delta)$) only when $N^j(t)<\lambda^j-\delta^{t-1}(1-\delta)$. Thus $N^j(\infty)\leq\lambda^j$ for each j, and, since $\sum_{j=1}^{m}\lambda^j=1$, $N^j(\infty)=\lambda^j$, proving the claim.

Now, by construction, the payoffs corresponding to {a(t)} are

$$(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}g(a(t)) =$$

$$(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}[\sum_{j=1}^{m}I^j(t)w^j] = \sum_{j=1}^{m}w^j(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}I^j(t)=\Sigma w^jN^j(\infty)=\Sigma\lambda^jw^j=v.$$

Q.E.D.

Roughly speaking, the algorithm of Lemma 1 works as follows. By definition, v is a convex combination $\Sigma\lambda^jw^j$ of the pure strategy payoff vectors $w^1,\ldots,w^m$. To generate v as a discounted average payoff over time, choose that pure strategy vector a^j at time t for which the difference between λ^j and the fraction of times a^j has been used up until t (suitably weighted for discounting) is largest.

The <u>continuation payoffs</u> at time s associated with the sequence $\{a(t)\}$ are simply $\sum_{t=s}^{\infty}\delta^{t-s}g(a(t))$, i.e. the discounted sum of per-period payoffs starting at time τ and discounted to time s.

<u>Lemma 2</u>: For every $\varepsilon>0$ and closed set $\hat{V}\subset V^*$ with $\min_{i,\hat{v}\in\hat{V}}\hat{v}_i>5\varepsilon$, there exists $\underline{\delta}<1$ such that, for all $\delta\in(\underline{\delta},1)$ and every $v\in\hat{V}$, there is a sequence $\{a(t)\}$ of pure strategies whose discounted average payoffs are v, and whose continuation payoffs at any time t are at least ϵ for each player.

<u>Remark</u>: To prove that public randomization is inessential for the Folk Theorem with observable mixed strategies, it would suffice to show that, for any individually rational payoff v, there exists $\underline{\delta}$ such that, if δ exceeds $\underline{\delta}$, v can be generated by a deterministic

sequence whose continuation payoffs are individually rational. Lemma 2 establishes a stronger property; it asserts that a fixed $\underline{\delta}$ works uniformly for the entire set $\hat{V}$. We provide the stronger result because it is needed to show that public randomization is inessential with unobservable mixed strategies and in our [1987] paper.

Proof: Let $\tilde{Z}$ be the polygon corresponding to the intersection of the set V with the inequality constraints $v_i \geq 3\varepsilon$, and let $\{\tilde{z}^j\}$ be the J vertices of $\tilde{Z}$. Clearly, $\hat{V} \subset \tilde{Z}$. Let Z be a polygon with vertices $\{z^j\}$ such that (i) each z^j is within ε of $\tilde{z}^j$; (ii) $z^j = \tilde{z}^j$ if $\tilde{z}^j \in \{w^1, \ldots, w^m\}$; and (iii) z^j can be expressed as a weighted average $\sum_k \lambda^k(j) w^k$, where each weight $\lambda^k(j)$ is a rational number. Observe that $\hat{V} \subset Z$.

Because the $\lambda^k(j)$'s are rational, we can find integers $\{r^k(j)\}_{k=1}^m$ and d such that for all j and k, $\lambda^k(j) = r^k(j)/d$. Let "cycle j" be the d-period sequence of pure strategies in which a^1 is played for the first $r^1(j)$ periods; a^2 is played for the next $r^2(j)$ periods; and so on for all k between 3 and m. (Recall that $w^k = g(a^k)$). Let $z^j(\delta)$ be the discounted average payoffs corresponding to cycle j. Note that if z^j is on the boundary of V, then $z^j(\delta)$ will be on the boundary as well. If we set

$$R^k(j) = \sum_{s=1}^{k} r^s(j), \text{ with } R^0(j) = 0,$$

then

$$z^j(\delta) = \sum_{k=1}^{m} \sum_{s=R^{k-1}(j)}^{R^k(j)-1} (1-\delta)\delta^s w^k/(1-\delta^d).$$

Choose $\underline{\delta}$ so that, for all δ greater than $\underline{\delta}$ and all j, $z^j(\delta)$ is within ε of z^j. By construction, for all $\delta > \underline{\delta}$, $\hat{V}$ is contained in the polygon $Z(\delta)$ whose vertices are the $z^j(\delta)$'s.

We now apply the algorithm of Lemma 1 to generate each $v \in \hat{V}$ by a deterministic sequence of the $z^j(\delta)$'s for $\delta > \hat{\underline{\delta}}$, where $\hat{\underline{\delta}} = \max(\underline{\delta}, 1-1/J)$. Earlier, when the payoffs w^j were called for in a given period t, we set $a(t) = a^j$. In our current application, we replace the w^j's with the $z^j(\delta)$'s. Moreover, when the algorithm calls for payoffs $z^j(\delta)$, we assign <u>cycle j</u> as the actions for the next d periods. The Lemma 1 algorithm so modified guarantees that we can generate each of the payoffs in $\hat{V}$ by a deterministic sequence of these cycles. Because each cycle is of length d and each $z^j(\delta)$ gives each player a payoff of at least 2ε, the continuation payoffs starting at any time τ give each player at least ε if $\underline{\delta}$ is taken large enough to satisfy $(1-\delta^d)\underline{g} + \delta^d(2\varepsilon) > \varepsilon$, where $\underline{g} = \min_{i,a} g_i(a)$ is the lowest possible value of any player's payoff.

Q.E.D.

To summarize, the algorithm of Lemma 1 shows how to attain any $v \in V$ by a deterministic sequence of w^j's. Lemma 2 replaces each w^j that is not individually rational with a payoff vector $z^j(\delta)$ that itself can be attained through a finite cycle of w^j's. Hence, to obtain $v \in V^*$ through a deterministic sequence, (i) apply the Lemma 1 algorithm using the $z^j(\delta)$'s instead of the w^j's; and (ii) whenever the algorithm calls for $z^j(\delta)$, replace it with the corresponding d-period cycle

To see how Lemma 2 enables us to do without public randomization in the proof of the Folk Theorem, we first recall the form of the strategies in our [1986] paper. To obtain the point $v \in V^*$, we had players use publicly correlated action (generating v in each period) as long as no player deviates. If player i deviates, we provided a "punishment equilibrium" in which, (a) for a certain number of periods, the player's opponents minimax him and he responds optimally and then (b) the players revert to a more "cooperative" mode in which their payoffs are $\hat{v}^i$, where this vector is chosen so as to induce i's punishers to go through with their minimaxing and so player i's overall payoff is less than ϵ. Like v, it is generated by publicly correlated actions. From Lemma 2, we can replace the publicly correlated actions yielding v and $\hat{v}^i$ with deterministic sequences whose continuation payoffs are greater than ϵ. Because deviation leads to payoffs less than ϵ, no player will wish to deviate from such a sequence. In the case where mixed strategies are observable, this is the only change to the proof required to eliminate public randomization.

The case where only players' realized actions (and not the randomizations themselves) are observable presents an additional complication. If, to minimax player i, player j uses a mixed strategy, he must must be indifferent among the various actions over which he randomizes. Our [1986] proof ensured this indifference by making j's continuation payoff after the punishment phase contingent on his actions during the phase. It is important here that precisely specified values for the continuation payoffs be attainable; it would not suffice merely to approximate them. Lemma 2 shows, however, these exact values can, in fact, be attained.[3] Thus public randomization is inessential in this case too.[4]

----

3. The strong version of Lemma 2 -- guaranteeing a uniform choice of $\underline{\delta}$ -- is needed here because the particular values of the continuation payoffs will depend on the discount factor. Without uniformity we would run into the difficulty that if δ were chosen large enough to invoke Lemma 2 for particular continuation payoffs, the continuation payoffs required to ensure indifference would themselves change, necessitating a new δ, etc.

4. We should point out that our [1986] paper was misleading in its assertion that the Folk Theorem holds approximately when public randomization is not feasible. Without the possibility of attaining certain continuation payoffs exactly, it is not clear that even an approximate version of the Folk Theorem holds for the case of unobservable mixed strategies.

References

Aumann, R., and L. Shapley [1976], "Long Term Competition: A Game Theoretic Analysis," mimeo, Hebrew University.

Friedman, J. [1971], "A Noncooperative Cooperative Equilibrium For Supergames," Review of Economic Studies, 38, 1-12.

Fudenberg, D. and E. Maskin [1986], "The Folk Theorem in Repeated Games With Discounting or With Incomplete Information," Econometrica, Vol. 54, No. 3, May 1986, pp. 533-554.

______________ and _________ [1987], "Nash and Perfect Equilibria of Discounted Repeated Games," mimeo, Harvard University.

Rubinstein, A. [1979], "Equilibrium in Supergames with the Overtaking Criterion," Journal of Economic Theory, No. 21, 1-9.

3611 096

Date Due

Lib-26-67

MIT LIBRARIES
3 9080 005 130 312