PUBLIC RANDOMIZATION IN DISCOUNTED REPEATED GAMES

Drew Fudenberg*
and
Eric Maskin**

*University of California, Berkeley

**Harvard University

Discussion Paper No. 1315

May, 1987

Harvard Institute for Economic Research
Harvard University
Cambridge, Massachusetts

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.
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 [1986b] 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.

Although this stronger result is an improvement over previous work, on its own it is of no great moment. The result assumes greater importance when we simultaneously allow players to use mixed strategies. Previous analyses of repeated games with little or no discounting (Aumann-Shapley [1976], Friedman [1971] and Rubinstein [1979]) had restricted players to pure strategies, or equivalently, had 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 [1986b] 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 if her continuation payoff depends on her current action in a way that renders her exactly indifferent between 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 the continuation payoffs could only be approximated without public randomization, the minimaxing player need not be exactly indifferent, and our construction would fail. Thus, if we obtain only an approximate version of the Folk Theorem without public randomization, we cannot accommodate unobservable mixed strategies.

Attaining payoffs exactly is also essential for the argument in our [1986a] paper, which provides sufficient conditions for the sets of Nash and perfect equilibrium payoffs to coincide for a range of discount factors less than one. Although the body of that paper assumes 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:

\[ g: A \rightarrow \mathbb{R}^n, \]

where \( A = A_1 \times \ldots \times A_n \) and \( A_i \) is player \( i \)'s action space. Let \( \Sigma_i \) 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 the 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 \( h(t) \in A^{t-1} = H(t) \) be the realized actions from time zero through time \( t-1 \). Player \( i \)'s strategy \( s_i \) is a sequence of maps \( s_i(t) \) from \( H(t) \) to \( \Sigma_i \). Note that player \( i \)'s strategy does not depend on the past randomizing probabilities of his opponents, but only on their realized actions.

In the repeated game \( G_\delta \), each player \( i \)'s payoff is the average discounted sum \( n_1 \) of his per-period payoffs, with common discount factor \( \delta \):

\[
\lim_{n \to \infty} (1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} g_i(\sigma(t)),
\]

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

For each player \( j \), choose "minimax strategies" \( m^j = (m^j_1, \ldots, m^j_n) \) so that
\[ m_j \leftarrow \arg \min_{\sigma_j} \max_{\sigma_{-j}} g_j(\sigma_j, \sigma_{-j}), \]

and let

\[ \nu_j^* = \max_{a_j} g_j(a_j, m_{-j}^j) = g_j(m_j^j). \]

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

We call \( \nu_j^* \) player j's reservation value. Player j's average
payoff must be at least \( \nu_j^* \) in any equilibrium of \( g \), whether or not \( g \)
is repeated.

Henceforth we shall normalize the payoffs of the game \( g \) so that
\( (\nu_1^*, \ldots, \nu_n^*) = (0, \ldots, 0) \). Call \( (0, \ldots, 0) \) the minimax point. Take

\[ \bar{\nu}_i = \max_{a_i} g_i(a). \]

Take

\[ U = \{(\nu_1, \ldots, \nu_n) \mid \text{there exists } \sigma \in \Sigma \text{ with } g(\sigma) = (\nu_1, \ldots, \nu_n)\}, \]

\[ V = \text{Convex Hull of } U, \]

and

\[ V^* = \{(\nu_1, \ldots, \nu_n) \in V \mid \nu_i > 0 \text{ for all } i\}. \]

3. **The Folk Theorem without Public Randomization**

Our [1986b] paper showed that if public randomization is
allowed, for any payoff vector \( v \in V^* \), there exists a discount factor
\( \delta < 1 \), such that for all \( \delta \in (\delta, 1) \), there is a perfect equilibrium of \( G_\delta \)
with payoffs $v$. We now demonstrate that public randomization is inessential for this result.  

In Lemma 1, we first show that, for $\delta$ 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 average discounted payoff vector. This is not sufficient to establish the Folk Theorem, because the sequence $\{a(t)\}$ might have the property that, for some period $\tau$, the continuation payoffs beginning at $\tau$ 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.

Lemma 2 builds on Lemma 1 to show that any payoff vector in $V^*$ can be generated by a deterministic sequence $\{a(t)\}$ in such a way that the continuation payoffs beginning in any period $\tau$ on are always in $V^*$. We then explain how Lemma 2 allows us to do without public randomization in our proof of the Folk Theorem.

Let $\{w^1, \ldots w^m\}$ be the possible payoffs corresponding to vectors of pure strategies.

**Lemma 1**: 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 $v$.

---

1. Of course, for low discount factors, public randomization does make a difference. If $\delta$ is near zero, the average payoffs for the sequence $\{\sigma(t)\}$ are approximately $g(\sigma(1))$, and so, equilibrium considerations aside, many payoffs in $V^*$ are not feasible.
Proof: Let \( v = \sum_{j=1}^{m} \lambda^j w^j \), where \( 0 \leq \lambda^j \leq 1 \), and \( \sum_{j=1}^{m} \lambda^j = 1 \). We construct \( \{a(t)\} \) as follows. Let \( a^j \) be a strategy vector such that \( g(a^j) = w^j \), and 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 \( N^j(t) = \sum_{\tau=1}^{t-1} (1 - \delta)^{\tau-1} I^j(\tau) \). \( N^j(t) \) is the "average discounted weight" given to strategy \( a^j \) before time \( t \). Let \( C(t) = \{ j | \lambda^j N^j(t) > (1 - \delta) \} \). 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 \( a(t) \).

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 \( \tau \) be the first such time. Then \( (1 - \delta)^{\tau-1} \lambda^j N^j \) for all \( j \). Summing over \( j \), we have
\[
\sum_{j=1}^{m} N^j(\tau) = 1 - \sum_{t=1}^{\tau-1} (1 - \delta)^{t-1} I^j(t)
\]
But (1) contradicts our assumption that \( m(1 - \delta) < 1 \), establishing the claim.

2. If there is a tie, make a deterministic selection.
Let $N^j(\infty) = \lim_{t \to \infty} N^j(t)$.

(Because $N^j$ 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 $\lambda^j$, because $N^j$ increases (by $\delta^{t-1}(1 - \delta)$) only when $N^j < \lambda^j - \delta^{t-1}(1 - \delta)$. 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

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

$$(1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} \left[ \sum_{j=1}^{m} I^j(t) w^j \right] = \sum_{j=1}^{m} w^j (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} I^j(t) = \sum_{j=1}^{m} w^j \sum_{t=1}^{\infty} \delta^{t-1} I^j(t) = \sum_{j=1}^{m} w^j N^j(\infty) = \sum_{j=1}^{m} w^j \lambda^j = v$$

Q.E.D.

Roughly speaking, the algorithm of Lemma 1 works as follows. By definition $v$ is a convex combination $\sum \lambda^j w^j$ 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 $\lambda^j$ and the fraction of times $a^j$ has been used up until $t$ (suitably weighted for discounting) is largest.
The continuation payoffs at time $\tau$ associated with the sequence 
\[ \{a(t)\} \text{ are simply } \sum_{t=\tau}^{\infty} \delta^{t-\tau} g(a(t)), \text{ i.e. the discounted sum of} \]
per-period payoffs starting at time $\tau$ and discounted to time $\tau$.

**Lemma 2**: For every $v \in V^*$ and every $\epsilon > 0$, there exists $\delta \in (0, 1)$ such that for all $\delta \in (\delta, 1)$ there is a sequence $\{a(t)\}$ of pure strategies whose discounted average payoffs are $v$, and all of whose continuation payoffs within $\epsilon$ of $v$.

**Proof**: Choose $n+1$ points $z(1), \ldots, z(n+1)$ in $V^*$ such that
(a) the point $v$ is in the interior of the polygon $P$ whose vertices are the $z(j)$'s. (If $v$ lies on the boundary of $P$, then $v$ should be in the relative interior of $P$);
(b) each $z(j)$ can be expressed as a convex combination of the pure strategy payoffs $w^k$, where $w^k$ has rational weight $x^k(j)$; and
(c) each $z(j)$ is at a distance less than $\epsilon$ from $v$.

Because the $x^k(j)$'s are rational, we can find integers
\[ \{r^k(j)\}_{k=1}^{m} \text{ and } d \text{ such that, for all } j \text{ and } k, \frac{x^k(j)}{r^k(j)} = \frac{r^k(j)}{d}. \]
Let "cycle $j" be the d-period sequence of play consisting of choosing the pure strategy vector (i) $a^1$ for the first $r^1(j)$ periods (recall that $a^i$ is the vector leading to payoffs $w^i$); (ii) $a^2$ for the next $r^2(j)$ periods; and continuing similarly for $i=3 \ldots m$. Let $z^*(j)$ be
\[ R^k(j) = \sum_{s=1}^{k} r^s(j), \text{ where } R^0(j) = 0. \]

Then

\[ \hat{z}_\delta(j) = \frac{m}{\sum_{k=1}^{k} \sum_{s=R^k(j)/1}^{(1-\delta)\delta^t w^k/(1-\delta^d)}.} \]

For \( \delta \) near 1, \( z_\delta(j) \) is near \( z(j) \). Conditions (a) and (c) imply, therefore, that, for \( \delta \) greater than some \( \delta \), \( v \) lies in the interior of the polygon formed by the \( z_\delta(j) \)'s and is at a distance less than from each \( z_\delta(j) \).

To obtain the conclusion of Lemma 2, we apply the algorithm of Lemma 1. However, where before we used the \( w^j \)'s, we now use the \( z_\delta(j) \)'s to generate \( v \). Previously, when the payoffs \( w^j \) were called for in a given period \( t \), we set \( a(t) = a^j \). Now, if the payoffs \( z_\delta(j) \) are called for by the algorithm, we assign the cycle \( j \) as actions (note that because the cycle \( j \) is of length \( d \), we are treating \( d \) periods as a single period for the purposes of the construction).

The Lemma 1 algorithm guarantees that we can generate the average payoffs \( v \) by moving among the different cycles \( j \) in a deterministic way when \( \delta \geq \delta \). Because, moreover, \( d \) is finite and each \( z_\delta(j) \) is within \( \epsilon \) of \( v \), the continuation payoffs starting at any time \( \tau \) are also within \( \epsilon \) of \( v \) for \( \delta \) near enough 1.

Q.E.D.
We noted above that the algorithm of Lemma 1 suffers the shortcoming that, although it generates the overall average payoff \( v \), the continuation payoffs need not all be positive. The construction of Lemma 2 overcomes this deficiency by building up \( v \) out of per-period payoffs \( z_{\delta}(j) \) that are very close to \( v \): if in all future periods players are getting something close to \( v \), all continuation payoffs must be near \( v \).

Let us now explain how Lemma 2 allows us to do without public randomization in the proof of the Folk Theorem. To do so, we first remind the reader of the form of the strategies constructed in our [1986b] paper. In the case where the individual players' mixed strategies were observable, we used public randomization to obtain payoffs \( v \in V^* \) along the equilibrium path using the stationary strategies "in each period play actions that yield \( v \) in expectation, as long as there have been no deviations." To show that this path can arise in a perfect equilibrium we then specified \( n \) "punishment equilibria," one for each player \( i \), with the \( i \)th punishment equilibrium to be followed if player \( i \) deviates. Our proof showed that for discount factors near 1, (1) the "punishment equilibria" are indeed equilibria, and (2) each player \( i \) strictly prefers \( v^i \) to deviating. From Lemma 2, we can replace the stationary strategies yielding \( v \) with a deterministic sequence whose continuation payoffs are close to \( v \). Because, in the original equilibrium, each player strictly preferred \( v \) to deviating, no player will wish to deviate from the deterministic sequence. Thus public randomization is inessential when mixed strategies are observable.
The case where only the player's realized actions (and not the randomizations themselves) are observable presents an additional complication. If a player's punishment strategy is mixed, the player must be indifferent among the various actions over which he randomizes. Our [1986b] proof ensured this indifference by making the player’s continuation payoff after the punishment phase contingent on his actions during the phase. It is important here that precise 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. Thus public randomization is inessential in this case too.3

3. We should point out that our [1986b] 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


Fudenberg, D. and E. Maskin [1986a], "Nash and Perfect Equilibria of Discounted Repeated Games," mimeo, Harvard University.
