THE FOLK THEOREM IN REPEATED GAMES WITH DISCOUNTING OR WITH INCOMPLETE INFORMATION[1]

By Drew Fudenberg and Eric Maskin[2]

When either there are only two players or a "full dimensionality" condition holds, any individually rational payoff vector of a one-shot game of complete information can arise in a perfect equilibrium of the infinitely-repeated game if players are sufficiently patient. In contrast to earlier work, mixed strategies are allowed in determining the individually rational payoffs (even when only realized actions are observable). Any individually rational payoffs of a one-shot game can be approximated by sequential equilibrium payoffs of a long but finite game of incomplete information, where players' payoffs are almost certainly as in the one-shot game.

1. INTRODUCTION

THAT STRATEGIC RIVALRY in a long-term relationship may differ from that of a one-shot game is by now quite a familiar idea. Repeated play allows players to respond to each other's actions, and so each player must consider the reactions of his opponents in making his decision. The fear of retaliation may thus lead to outcomes that otherwise would not occur. The most dramatic expression of this phenomenon is the celebrated "Folk Theorem" for repeated games. An outcome that Pareto dominates the minimax point is called individually rational. The Folk Theorem asserts that any individually rational outcome can arise as a Nash equilibrium in infinitely repeated games with sufficiently little discounting. As Aumann and Shapley [3] and Rubinstein [20] have shown, the same result is true when we replace the word "Nash" by "(subgame) perfect" and assume no discounting at all.

Because the Aumann–Shapley/Rubinstein result supposes literally no discounting, one may wonder whether the *exact* counterpart of the Folk Theorem holds for perfect equilibrium, i.e., whether as the discount factor tends to one, the set of perfect equilibrium outcomes converges to the individually rational set. After all, agents in most games of economic interest are not completely patient; the no discounting case is of interest as an approximation.

It turns out that this counterpart is false. There can be a discontinuity (formally, a failure of lower hemicontinuity) where the discount factor, δ, equals one, as we show in Example 3. Nonetheless the games in which discontinuities occur are quite degenerate, and, in the end, we *can* give a qualified "yes" (Theorem 2) to the question of whether the Folk Theorem holds with discounting. In particular, it always holds in two-player games (Theorem 1). This last result contrasts with the recent work of Radner–Myerson–Maskin [18] showing that, even in two-player games, the equilibrium set may not be continuous at $\delta = 1$ in

the discount factor if players' moves are not directly observable and outcomes depend stochastically on moves.

Until recently, the study of perfect equilibrium in repeated games concentrated mainly on infinite repetitions without discounting ("supergames"). One early exception was Friedman [5 and 6], who showed that any outcome that Pareto dominates a Nash equilibrium of the constituent game (the game being repeated) can be supported in a perfect equilibrium of the repeated game.[3] The repeated game strategies he specified are particularly simple: after any deviation from the actions that sustain the desired outcome, players revert to the one-shot Nash equilibrium for the remainder of the game. More recently, Abreu [1] established that a highly restricted set of strategies suffices to sustain any perfect equilibrium outcome. Specifically, whenever any player deviates from the desired equilibrium path, that player can be "punished" by players' switching to the worst possible equilibrium for the deviator regardless of the history of the game to that point.

We exploit this idea of history-free punishments, by contrast with the methods of Aumann–Shapley/Rubinstein, in the proofs of our Theorems 1 and 2, which are constructive.[4] In the proof of the two-person "discounting folk theorem" (Theorem 1), *both* players switch for a specified number of periods to strategies that minimize their opponent's maximum payoff (i.e., minimax strategies) after any deviation. Theorem 2 treats the n-person case, where "mutual minimaxing" may be impossible. In this case we impose a "full dimensionality" assumption that enables players to be rewarded for having carried out punishments. Theorem 5 also makes use of rewards to punishers to show that mixed strategies can be used as punishments even if only realized actions, and not the mixed strategies themselves, are observable. This provides a substantially stronger result, because the individually rational payoff levels are often lower with mixed strategies than with pure ones.

Although the theory of infinitely repeated games offers an explanation of cooperation in ongoing relationships, economic agents often have finite lives. If the game has a long but finite length, the set of equilibria may be much smaller than the folk theorem would suggest. The classic example here is the repeated prisoner's dilemma: with a fixed finite horizon the only equilibrium involves both players' confessing every period, in contrast with the cooperative equilibrium that is sustainable with an infinite horizon.[5] Still anecdotal and experimental evidence both suggest that cooperation is a likely outcome with a large but finite number of repetitions.

Recently Kreps–Wilson [14], Milgrom–Roberts [17], and Kreps–Milgrom–Roberts–Wilson [13] have proposed a reason why a finite number of repetitions might allow cooperation. Their explanation supposes that players are uncertain about the payoffs or possible actions of their opponents. Such "incomplete

---

[3] Actually Friedman was concerned explicitly only with Nash equilibria of the repeated game. The strategies that he proposed, however, constitute perfect equilibria (See Theorem C of Section 2).

[4] Lockwood [16] characterizes the (smaller set of) equilibrium payoffs that are possible when one restricts attention to punishments of the Aumann–Shapley/Rubinstein variety.

[5] See, however, Benoit and Krishna [4] and Friedman [7] who show that when a game with multiple equilibria is repeated even only finitely many times, "Folk-Theorem-like" results may emerge.

information" in the prisoner's dilemma precludes applying the backwards-induction argument that establishes that the players must confess each period. Players can credibly threaten to take suboptimal actions if there is some (small) probability that the action is indeed optimal, because they have an interest in maintaining their reputation for possible "irrationality."

The examples of reputation games analyzed to date exhibit the apparent advantage, compared with infinite-horizon models, of having substantially smaller sets of equilibria. However, the equilibrium set depends on the precise form of irrationality specified. Our "incomplete information" Folk Theorem shows that by varying the kind of irrationality specified, but still keeping the probability of irrationality arbitrarily small, one can trace out the entire set of infinite-horizon equilibria. Thus, in a formal sense, the two approaches, infinite and finite horizon, yield the same results. However, those who are willing to choose among different forms of irrationality may still find the incomplete information approach useful. One may argue for or against certain equilibria on the basis of the type of irrationality needed to support them.

We provide two different theorems for repeated games of incomplete information. Our first result (Theorem 3) parallels Friedman's work on repeated games with discounting: after a deviation the "crazy" player switches to a Nash-equilibrium strategy of the constituent game. This simple form of irrationality suffices to support any outcome that Pareto-dominates a (one-shot) Nash equilibrium. Our second, and main, result (Theorem 4) uses a more complex form of irrationality. However, the basic approach is the same as in our Folk Theorem with discounting: after a deviation each player switches to his minimax strategy for a specified number of periods.

It is not surprising that similar kinds of arguments should apply to both infinite horizon games with discounting and finite horizon games. Each type of game entails the difficulty, not present in infinite horizon games without discounting, that deviators from the equilibrium path cannot be "punished" arbitrarily severely. This limitation is a problem because of the requirement of perfection. Deviators must be punished, but it must also be in the interest of the punishers to punish. That is, they must themselves be threatened with punishment if they fail to punish a deviator. Such considerations give rise to an infinite sequence of potential punishments that, at each level, enforce the punishments of the previous level. Depending on how these punishments are arranged, they may have to become increasingly severe the farther out in the sequence they lie. This creates no problem in supergames but may be impossible for the two types of games that we consider. It seems natural, therefore, to study these two types together.

Section 2 presents the classical Folk Theorem and the Aumann–Shapley/Rubinstein and Friedman variants. Section 3 discusses continuity of the equilibrium correspondence as a function of the discount factor and develops Folk Theorems for infinitely repeated games with discounting. Section 4 provides a simple proof that any payoffs that Pareto dominate a (one-shot) Nash equilibrium can be sustained in an equilibrium of a finitely repeated game with incomplete information. This result is the analog of the Friedman [5] result. Section 5 uses a more complex approach to prove a Folk Theorem for these finitely repeated games.

Sections 2-5 follow previous work on repeated games in assuming that, if mixed strategies are used as punishments, they are themselves observable. In Section 6 we drop this assumption but show that our results continue to hold under the more natural hypothesis that players can observe only each other's past *actions*.

2. THE CLASSICAL FOLK THEOREM

Consider a finite n-person game in normal form

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

For now, we shall not distinguish between pure and mixed strategies, and so we might as well suppose that the A_i's consist of mixed strategies. Thus, we are assuming either that a player can observe the others' past mixed strategies, as in previous work on repeated games, or restricting players to pure strategies. Mixed strategies can be made observable if the outcomes of players' randomizing devices are jointly observable *ex-post*. (More importantly, we show in Section 6 that the assumption is not necessary.) Moreover, for convenience, we assume that the players can make their actions contingent on the outcome of a *public* randomizing device. That is, they can play correlated strategies.[6] Even if a correlated strategy over vectors of actions cannot literally be adopted, it can still be approximated if the action vectors are played successively over time and the frequency of any given vector corresponds to its probability in the correlated strategy. To see how to modify the statements of the theorems if correlated strategies cannot be used, see the Remark following Theorem A.

For each j, choose $M^j = (M^j_1, \ldots, M^j_n)$ so that

$$(M^j_1, \ldots, M^j_{j-1}, M^j_{j+1}, \ldots, M^j_n) \in \arg \min_{a_{-j}} \max_{a_j} g_j(a_j, a_{-j}),$$

and

$$v^*_j \equiv \max_{a_j} g_j(a_j, M^j_{-j}) = g_j(M^j).^7$$

The strategies $(M^j_1, \ldots, M^j_{j-1}, M^j_{j+1}, \ldots, M^j_n)$ are minimax strategies (which may not be unique) against player j, and v^*_j is the smallest payoff that the other players can keep player j below.[8] We will call v^*_j player j's reservation value and refer to $(v^*_1, \ldots, v^*_n)$ as the minimax point. Clearly, in any equilibrium of g— whether or not g is repeated—player j's expected average payoff must be at least v^*_j.

---

[6] See Aumann [2]. More generally, a correlated strategy might entail having each player make his action contingent on a (private) signal correlated with some randomizing device. We shall, however, ignore this possibility.

[7] The notation "a_{-j}" denotes "$(a_1, \ldots, a_{j-1}, a_{j+1}, \ldots, a_n)$", and "$g_j(a_j, M^j_{-j})$" denotes "$g_j(M^j_1, \ldots, M^j_{j-1}, a_j, M^j_{j+1}, \ldots, M^j_n)$".

[8] Actually, if $n \geq 3$, the other players may be able to keep player j's payoff even lower by using a correlated strategy against j, where the outcome of the correlating device is not observed by j (another way of putting this is to observe that, for $n \geq 3$, the inequality $\max_{a_j} \min_{a_{-j}} g_j(a_j, a_{-j}) \leq \min_{a_{-j}} \max_{a_j} g_j(a_j, a_{-j})$ can hold strictly). In keeping with the rest of the literature on repeated games, however, we shall rule out such correlated strategies.

Henceforth we shall normalize the payoffs of the game g so that $(v_1^*, \ldots, v_n^*) = (0, \ldots, 0)$. Let

$$U = \{(v_1, \ldots, v_n)|\exists (a_1, \ldots, a_n) \in A_1 \times \cdots \times A_n$$

$$\text{with } g(a_1, \ldots, a_n) = (v_1, \ldots, v_n)\},$$

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

and

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

The set V consists of feasible payoffs, and V^* consists of feasible payoffs that Pareto dominate the minimax point. That is, V^* is the set of individually rational payoffs. In a repeated version of g, we suppose that players maximize the discounted sum of single period payoffs. That is, if $(a_1(t), \ldots, a_n(t))$ is the vector of actions played in period t and δ is player i's discount factor, then his payoff is $\sum_{t=1}^{\infty} \delta^{t-1} g_i(a_1(t), \ldots, a_n(t))$ and his average payoff is $(1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} g_i(a_1(t), \ldots, a_n(t))$. We can now state a version of the Folk Theorem (see Hart [10] for more details).

THEOREM A (The Folk Theorem): *For any* $(v_1, \ldots, v_n) \in V^*$, *if players discount the future sufficiently little, there exists a Nash equilibrium of the infinitely repeated game where, for all i, player i's average payoff is* v_i.[9]

PROOF: Let $(s_1, \ldots, s_n) \in A_1 \times \cdots \times A_n$ be a vector of strategies[10] such that $g(s_1, \ldots, s_n) = (v_1, \ldots, v_n)$. Suppose that in the repeated game each player i plays s_i until some player j deviates from s_j (if more than one player deviates simultaneously, we can suppose that the deviations are ignored). Thereafter, assume that he plays M_i^j. These strategies form a Nash equilibrium of the repeated game if there is not too much discounting; any momentary gain that may accrue to player j if he deviates from s_j is swamped by the prospect of being minimaxed forever after.   Q.E.D.

REMARK: If we disallowed correlated strategies, the same proof would establish that any positive vector in U could be enforced as an equilibrium. For other points $(v_1, \ldots, v_n)$ in V^* the statement of the theorem must be modified to read: for all $\varepsilon > 0$ there exists $\underline{\delta} < 1$ such that, for all $\delta > \underline{\delta}$, there exists a subgame perfect equilibrium of the infinitely repeated game in which each player i's average payoff is within ε of v_i, when players have discount factor δ. The ε qualification is needed because discounting and the requirement that each vector of actions be played an integral number of times limit the accuracy of approximating a correlated strategy by switching among action vectors over time.

---

[9] The hypothesis that the v_i's are positive is important, as a recent example by Forges, Mertens, and Neyman demonstrates.

[10] Or, if necessary, correlated strategies.

Of course, the strategies of Theorem A do not, in general, form a (subgame) perfect equilibrium (such an equilibrium is a configuration of strategies that form a Nash equilibrium in all subgames), because, if a player deviates, it may not be in others' interest to go through with the punishment of minimaxing him forever. However, Aumann and Shapley [3] and Rubinstein [20] showed that, when there is no discounting, the counterpart of Theorem A holds for perfect equilibrium.

THEOREM B (Aumann–Shapley/Rubinstein): *For any* $(v_1, \ldots, v_n) \in V^*$ *there exists a perfect equilibrium in the infinitely repeated game with no discounting, where, for all i, player i's expected payoff each period is* v_i.[11]

REMARK: The Aumann–Shapley and Rubinstein arguments assume that past mixed strategies are observable (or, alternatively, that only pure strategies are ever played, which, in general, implies a smaller equilibrium set). However, the methods of Section 6 can be used to establish Theorem B in the case where only past actions are observable.

The idea of the proof is simple to express. Once again, as long as everyone has previously conformed, players continue to play their s_i's, leading to payoff v_i. If some player j deviates, he is, as before, minimaxed but, rather than forever, only long enough to wipe out any possible gain that he obtained from this deviation. After this punishment, the players go back to their s_i's. To induce the punishers to go through with their minimaxing, they are threatened with the prospect that, if any one of them deviates from his punishment strategy, he in turn will be minimaxed by the others long enough to make such a deviation not worthwhile. Moreover, *his* punishers will be punished if any one of *them* deviates, etc. Thus, there is a potential sequence of successively higher order punishments, where the punishment at each level is carried out for fear the punishment at the next level will be invoked.

Theorem B is not an exact counterpart of Theorem A because it allows no discounting at all (we investigate in Section 3 when an exact counterpart holds). Moreover, the strategies of the proof are a good deal more complex than those of Theorem A. One well-known case that admits both discounting and simple strategies is where the point to be sustained Pareto dominates the payoffs of a Nash equilibrium of the constituent game g.

THEOREM C (Friedman [5] and [6]): *If* $(v_1, \ldots, v_n) \in V^*$ *Pareto dominates the payoffs* $(y_1, \ldots, y_n)$ *of a* (*one-shot*) *Nash equilibrium* $(e_1, \ldots, e_n)$ *of g, then, if players discount the future sufficiently little, there exists a perfect equilibrium of the infinitely repeated game where, for all i, player i's average payoff is* v_i.

---

[11] If there is no discounting, the sum of single-period payoffs cannot serve as a player's repeated game payoff since the sum may not be defined. Aumann and Shapley use (the lim infinum of) the average payoff; Rubinstein considers both this and the overtaking criterion, and the sketch of the proof we offer corresponds to this latter rule. (See also Hart [10]). The average payoff criterion allows more outcomes to be supported as equilibria than the overtaking criterion because for a player to strictly prefer to deviate he must gain in infinitely many periods. Indeed, for the former criterion, Theorem B holds for the closure of V^*.

PROOF: Suppose that players play actions that sustain $(v_1, \ldots, v_n)$ until someone deviates, after which they play $(e_1, \ldots, e_n)$ forever. With sufficiently little discounting, this behavior constitutes a perfect equilibrium.   *Q.E.D.*

Because the punishments used in Theorem C are less severe than those in Theorems A and B, its conclusion is correspondingly weaker. For example, Theorem C does not allow us to conclude that a Stackelberg outcome can be supported as an equilibrium in an infinitely-repeated quantity-setting duopoly.

3. THE FOLK THEOREM IN INFINITELY REPEATED GAMES WITH DISCOUNTING

We now turn to the question of whether Theorem A holds for perfect rather than Nash equilibrium. Technically speaking, we are investigating the *lower hemicontinuity*[12] of the perfect equilibrium average payoff correspondence (where the independent variable is the discount factor, δ) at $\delta = 1$. We first remind the reader that this correspondence is *upper* hemicontinuous.[13]

THEOREM D: *Let* $V(\delta) = \{(v_1, \ldots, v_n) \in V^* | (v_1, \ldots, v_n) \text{ are the average payoffs of a perfect equilibrium of the infinitely repeated game where players have discount factor } \delta\}$. *The correspondence* $V(\cdot)$ *is upper hemicontinuous at any* $\delta < 1$.

It is easy to give examples where $V(\cdot)$ fails to be lower hemicontinuous at $\delta < 1$.

EXAMPLE 1: Consider the following version of the Prisoner's Dilemma:

	C	D
C	1, 1	−1, 2
D	2, −1	0, 0

For $\delta < 1/2$ there are no equilibria of the repeated game other than players' choosing D every period. However at $\delta = 1/2$ many additional equilibria appear, including playing C each period until someone deviates and thereafter playing D. Thus $V(\cdot)$ is not lower hemicontinuous at $\delta = 1/2$.

3A. *Two-Player Games*

Our particular concern, however, is the issue of lower hemicontinuity at $\delta = 1$, and we begin with two-player games. It turns out that, in this case, the exact analog of Theorem A holds for perfect equilibrium. We should point out, however,

---

[12] A correspondence $f: X \to Y$ is lower hemicontinuous at $x = \bar{x}$ if for any $\bar{y} \in f(\bar{x})$ and any sequence $x^m \to \bar{x}$ there exists a sequence $y^m \to \bar{y}$ such that $y^m \in f(x^m)$ for all m.

[13] If Y is compact the correspondence $f: X \to Y$ is upper hemicontinuous at $\bar{x}$ if for any sequence $x^m \to \bar{x}$ and any sequence $y^m \to \bar{y}$ such that $y^m \in f(x^m)$ for all m, we have $\bar{y} \in f(\bar{x})$.

that to establish this analog we *cannot* use the Aumann-Shapley/Rubinstein (AS/R) strategies of Theorem B once there is discounting. To see that, if there is discounting, such strategies may not be able to sustain all individually rational points, consider the following example.

EXAMPLE 2:

	C	D
C	1, 1	0, −2
D	−2, 0	−1, −1

For this game the minimax point is $(0, 0)$, and so a "folk theorem" would require that we be able to sustain, in particular, strategies that choose (C, C) a fraction $(\varepsilon + 1)/2$ of the time and (D, D) the remainder of the time, with $0 < \varepsilon < 1$ (note that for δ near 1, these strategies yield average payoffs of approximately $(\varepsilon, \varepsilon)$, which are individually rational). However such behavior cannot be part of an AS/R type of equilibrium.[14] Suppose, for example, that one of the players (say, player I) played C in a period where he was supposed to play D. In an AS/R equilibrium, player II would "punish" I by playing D sufficiently long to make I's deviation unprofitable. I's immediate gain from deviation is 1, and I's best response to D is C, resulting in a payoff 0. Therefore if the punishment lasts for t_1 periods, t_1 must satisfy

$$\frac{\delta\varepsilon(1-\delta^{t_1})}{1-\delta} > 1 + t_1 \cdot 0 = 1.$$

That is,

$$(1) \qquad t_1 > \frac{\log\left(\dfrac{\varepsilon\delta - 1 + \delta}{\delta\varepsilon}\right)}{\log \delta}.$$

Condition (1) can be satisfied as long as

$$(2) \qquad \delta > \frac{1}{1+\varepsilon}.$$

---

[14] Although Rubinstein's [19] theorem applies to stationary strategies without public correlation (and so does not directly imply that $(\varepsilon, \varepsilon)$ can be enforced without discounting) it does permit a continuum of strategies. Thus we can think of an "enlarged" version of this game, with a continuum of strategies indexed by ε, $0 < \varepsilon < 1$, as well as the original strategies C and D. Suppose that if both players play ε, they each receive ε. If one plays ε and the other plays ε', they each get 1/2. If player 2 plays ε and 1 plays C, 1 receives $1 + \varepsilon$ and 2 receives 0. If 2 plays ε and 1 plays D, 1 gets $-1 + \varepsilon$ and 2 gets ε. The payoffs are permuted if the roles are reversed. In this enlarged game, D is still the minimax strategy and $(0, 0)$ is the minimax point. Hence the pure strategy pair $(\varepsilon, \varepsilon)$ is sustainable by AS/R punishments without discounting, but, as we show in the text, these punishments fail in the discounting case.

But in order to punish player I, II must himself suffer a payoff of -2 for t_1 periods. To induce him to submit to such self-laceration, he must be threatened with a t_2-period punishment, where

$$-2\frac{(1-\delta^{t_1})}{1-\delta}+\frac{\delta^{t_1}\varepsilon(1-\delta^{t_2-t_1+1})}{1-\delta}>1.$$

That is,

(3) $\qquad t_2>-1+\log\frac{\delta^{t_1}\varepsilon-3+2\delta^{t_1}+\delta}{\varepsilon}\Big/\log\delta.$

Such a t_2 exists as long as

$$\delta^{t_1}\varepsilon-3+2\delta^{t_1}+\delta>0,$$

which requires that

(4) $\qquad \delta>\left(\frac{2}{2+\varepsilon}\right)^{1/t_1}.$

But (4) is a more stringent requirement than (2), since

$$(2/(2+\varepsilon))^{1/t_1}>\frac{1}{1+\varepsilon}.$$

Continuing iteratively, we find that, for successively higher order punishments, δ is bounded below by a sequence of numbers converging to 1. Since δ is itself strictly less than 1, however, this is an impossibility, and so an AS/R equilibrium is impossible.

The problem is that in this example the punisher is hurt more severely by his punishment than is his victim. He must therefore be threatened with an even stronger punishment. Without discounting, this can be arranged by (roughly) taking the t_i's to be a geometric series, as in Rubinstein [20]. With discounting, however, arbitrarily long punishments are not arbitrarily severe, because far-off punishments are relatively unimportant.

These punishment strategies are not "simple" in the sense of Abreu [1] because they are not independent of history, i.e., they depend on the previous sequence of deviations. Abreu's work shows that there is no loss in restricting attention to simple punishments when players discount the future. Indeed, we make use of simple punishments in the proof of the following result, which shows that we can do without arbitrarily severe punishments in the two-player case.

THEOREM 1: *For any* $(v_1,v_2)\in V^*$ *there exists* $\underline{\delta}\in(0,1)$ *such that, for all* $\delta\in(\underline{\delta},1)$, *there exists a subgame perfect equilibrium of the infinitely repeated game in which player i's average payoff is v_i when players have discount factor δ.*

PROOF: Let M_1 be player one's minimax strategy against two, and M_2 a minimax strategy against one. Take $\bar{v}_i=\max_{a_1,a_2}g_i(a_1,a_2)$. For $(v_1,v_2)\in V^*$ choose $\underline{\nu}$ and $\underline{\delta}$ such that for $i=1,2$,

(5) $\qquad v_i>\bar{v}_i(1-\underline{\delta})+\underline{\delta}v_i^{**},$

where

(6) $v_i^{**} = (1 - \underline{\delta}^{\nu}) g_i(M_1, M_2) + \underline{\delta}^{\nu} v_i,$

with

(7) $v_i^{**} > 0.$

To see that $\underline{\nu}$ and $\underline{\delta}$ exist, choose $\underline{\delta}$ close enough to 1 so that

(5a) $v_i > \bar{v}_i(1 - \underline{\delta})$

and (7) holds when $\underline{\nu} = 1$. If with $\underline{\nu} = 1$, (5) is violated, consider raising $\underline{\nu}$. From (5a), (5) will eventually be satisfied. Since, for $\underline{\delta}$ close to 1, (6) declines almost continuously as $\underline{\nu}$ increases, by taking $\underline{\delta}$ near enough to one we can ensure that (7) will be satisfied for the first $\underline{\nu}$ for which (5) holds.

Condition (5) guarantees that player i prefers receiving v_i forever to receiving his maximum possible payoff ($\bar{v}_i$) once, then receiving $g_i(M_1, M_2)$ for $\underline{\nu}$ periods, and receiving v_i thereafter. Condition (7) ensures that being punished for deviating is still better than receiving the reservation value, zero, forever. Clearly, for any $\delta > \underline{\delta}$ there is a corresponding $\nu(\delta)$ such that (5) and (7) hold for $(\delta, \nu(\delta))$.

Let (s_1, s_2) be correlated one-shot strategies corresponding to (v_1, v_2): $g_i(s_1, s_2) = v_i$. Consider the following repeated game strategies for player i:

(A) Play s_i each period as long as (s_1, s_2) was played last period.

After any deviation from (A):

(B) Play $M_i\nu(\delta)$ times and then start again with (A). If there are any deviations while in phase (B), then begin phase (B) again.

These strategies form a subgame-perfect equilibrium. Condition (5) guarantees that deviation is not profitable in phase (A). In phase (B), player i receives an average payoff of at least v_i^{**} by not deviating. If he deviates, he can obtain at most 0 the first period (because his opponent, j, is playing M_j), and thereafter can average at most v_i^{**}. Hence deviation is not profitable in phase (B).   Q.E.D.

The statement of Theorem 1 assumes that correlated strategies are available. To see how the theorem must be modified if they are not, see the remark following Theorem A.

The strategies in the proof of Theorem 1 are easily summarized. After a deviation by either player, each player minimaxes the other for a certain number of periods, after which they return to the original path. If a further deviation occurs during the punishment phase, the phase is begun again.

Notice that in the proof of Theorem 1 the only place where we invoked our assumption that past mixed strategies can be observed was in supposing that deviations from the minimax strategies, M_1 and M_2, can be detected. This assumption is dropped in Section 6.

3B. *Three or More Players*

The method we used to establish Theorem 1—"mutual minimaxing"—does not extend to three or more players. This is because with, say, three players there

may exist no triple of alternatives (M_1, M_2, M_3) such that M_2 and M_3 minimax player one, M_1 and M_3 minimax two, and M_1 and M_2 minimax three; that is, the "mutual minimax" property may fail. However, the situation is even worse: not only does the method of proving Theorem 1 fail to extend, but the result itself does not generalize. To see this, consider the following example.

EXAMPLE 3:

1, 1, 1	0, 0, 0	0, 0, 0	0, 0, 0
0, 0, 0	0, 0, 0	0, 0, 0	1, 1, 1

In this game, player one chooses rows, player two chooses columns, and three, matrices. Note that whatever one player gets, the others get too.

CLAIM: For any $\delta < 1$ there does *not* exist a perfect equilibrium of the super-game in which the average payoff ε is less than $1/4$ (the one-shot mixed strategy equilibrium payoff).

PROOF: For fixed $\delta < 1$, let $\alpha = \inf \{\varepsilon | \varepsilon$ sustainable as an average payoff of a player in a perfect equilibrium$\}$. We must show that $\alpha \geq 1/4$. Let

$$\beta = \min_{\sigma_1^*, \sigma_2^*, \sigma_3^*} \max_{\sigma_1, \sigma_2, \sigma_3} \{g_1(\sigma_1, \sigma_2^*, \sigma_3^*), g_2(\sigma_1^*, \sigma_2, \sigma_3^*),$$

$$g_3(\sigma_1^*, \sigma_2^*, \sigma_3)\}.$$

That is, β is the minimum that the most fortunate defector can obtain in an optimal (one-shot) deviation from an arbitrary configuration of strategies. We claim that $\beta \geq 1/4$. Hence, the mutual minimax property does not hold.

To see this, let a_i be the probability that player i plays the "first" pure strategy, i.e., the first column, row, or matrix as appropriate. For some player i it must be the case that, for $j \neq i \neq k$, either $a_j \geq 1/2$ and $a_k \geq 1/2$ or $a_j \leq 1/2$ and $a_k \leq 1/2$. But since player i can obtain any convex combination of $a_j a_k$ and $(1 - a_j)(1 - a_k)$ as a payoff, he can get a payoff of at least $1/4$.

Thus in any equilibrium, the payoff to deviating for some player is at least $1/4 + \delta\alpha/(1 - \delta)$. Let $\{\varepsilon_m\}$ be a sequence of possible average payoffs in perfect equilibria, where $\varepsilon_m \to \alpha$. For all m we have $1/4 + \delta\alpha/(1 - \delta) \leq \varepsilon_m/(1 - \delta)$. Hence, $1/4 + \delta\alpha/(1 - \delta) \leq \alpha/(1 - \delta)$, and so $\alpha \geq 1/4$.   Q.E.D.

The game of Example 3 is degenerate in the sense that V^*, the individually rational set, is one-dimensional. This degeneracy is responsible for the discontinuity in $V(\delta)$ as the next result demonstrates.

THEOREM 2: *Assume that the dimensionality of V^* equals n, the number of players, i.e., that the interior of V (relative to n-dimensional space) is nonempty. Then, for any $(v_1, \ldots, v_n)$ in V^*, there exists $\underline{\delta} \in (0, 1)$ such that for all $\delta \in (\underline{\delta}, 1)$ there exists a subgame-perfect equilibrium of the infinitely repeated game with discount factor δ in which player i's average payoff is v_i.*

The idea behind the proof of Theorem 2 is simple. If a player deviates, he is minimaxed by the other players long enough to wipe out any gain from his deviation. To induce the other players to go through with minimaxing him, they are ultimately given a "reward" in the form of an additional "ε" in their average payoff. The possibility of providing such a reward relies on the full dimensionality of the payoff set.

PROOF: Choose $s = (s_1, \ldots, s_n)$ so that $g(s_1, \ldots, s_n) = (v_1, \ldots, v_n)$ (again we allow correlated strategies). Also choose $(v_1', \ldots, v_n')$ in the interior of V^* such that $v_i > v_i'$ for all i. Since $(v_1', \ldots, v_n')$ is in the interior of V^* and V^* has full dimension, there exists $\varepsilon > 0$ so that, for each j,

$$(v_1' + \varepsilon, \ldots, v_{j-1}' + \varepsilon, v_j', v_{j+1}' + \varepsilon, \ldots, v_n' + \varepsilon)$$

is in V^*. Let $T^j = (T_1^j, \ldots, T_n^j)$ be a joint strategy that realizes these payoffs. Let $M^j = (M_1^j, \ldots, M_n^j)$ be an n-tuple of strategies such that the strategies for players other than j together minimize player j's maximum payoff, and such that $g_j(M^j) = 0$. Let $w_i^j = g_i(M^j)$ be player i's per-period payoff when minimaxing player j. For each i choose an integer ν_i such that

$$(8) \qquad \frac{\bar{v}_i}{v_i'} < 1 + \nu_i,$$

where, as before, $\bar{v}_i$ is player i's greatest one-shot payoff.

Consider the following repeated game strategy for player i:

(A) play s_i each period as long as s was played last period. If player j deviates from (A),[15] then:

(B) play M_i^j for ν_j periods, and then:

(C) play T_i^j thereafter.

If player k deviates in phase (B) or (C), then begin phase (B) again with $j = k$.[16]

If player i deviates in phase (A) and then conforms, he receives at most $\bar{v}_i$ the period he deviates, zero for ν_i periods, and v_i' each period thereafter. His total payoff, therefore, is no greater than

$$(9) \qquad \bar{v}_i + \frac{\delta^{\nu_i + 1}}{1 - \delta} v_i'.$$

---

[15] If several players deviate from (A) simultaneously, then we can just as well suppose that everyone ignores the deviation and continues to play s.

[16] As in footnote 15, we can suppose that simultaneous deviation by several players is ignored.

If he conforms throughout, he obtains $v_i/(1-\delta)$, so that the gain to deviating is less than

$$(10) \qquad \bar{v}_i - \frac{1-\delta^{(\nu_i+1)}}{1-\delta}\, v_i'.$$

Because $(1-\delta^{\nu_i+1})/(1-\delta)$ converges to ν_i+1 as δ tends to 1, condition (8) ensures that (10) is negative for all δ larger than some $\underline{\delta} < 1$. If player i deviates in phase (B) when *he* is being punished, he obtains at most zero the period in which he deviates, and then only lengthens his punishment, postponing the positive payoff v_i'. If player i deviates in phase (B) when player j is being punished, and then conforms, he receives at most

$$\bar{v}_i + \frac{\delta^{\nu_i+1}}{(1-\delta)}\, v_i',$$

which is less than $\bar{v}_i + v_i'/(1-\delta)$. If, however, he does not deviate, he receives at least

$$w_i^j \frac{(1-\delta^\nu)}{1-\delta} + \frac{\delta^{\nu+1}}{1-\delta}(v_i'+\varepsilon),$$

for some ν between 1 and ν_j. Thus the gain to deviating is at most

$$(11) \qquad \bar{v}_i + \frac{(1-\delta^{\nu+1})}{1-\delta}(v_i' - w_i^j) - \frac{\delta^{\nu+1}}{1-\delta}\varepsilon - \delta^\nu w_i^j.$$

As $\delta \to 1$, the second term in (11) remains finite because $(1-\delta^{\nu+1})/(1-\delta)$ converges to $\nu+1$. But, because δ^ν converges to 1, the third converges to negative infinity. Thus there exists $\underline{\delta}_i < 1$ such that for all $\delta > \underline{\delta}_i$, player i will not deviate in phase (B) if the discount factor is δ.

Finally, the argument for why players do not deviate in phase (C) is practically the same as that for phase (A).   *Q.E.D.*

4. INCOMPLETE INFORMATION WITH NASH THREATS

Suppose that a game is repeated finitely many times, ν, that players maximize the (expected) sum of their one-shot payoffs, and that players can observe all past one-shot strategies (including mixed strategies). This repeated game can be embedded in a ν-period sequential game of incomplete information. Suppose that players' payoffs and, perhaps, even their action spaces A_i depend on their types (although we shall not explicitly consider this latter type of incomplete information). With probability, say, $1-\varepsilon$, a given player i is described by g_i. We call a player of this type "sane" or "rational." However with probability ε his payoffs and action spaces may be different and might even be period-dependent. Such a player we call "crazy." The motivation for suggesting this possibility is that often one cannot be sure what kind of player one is up against. One might

be *almost* sure, but even if ε is nearly zero, one may nevertheless wish to take into account other possibilities. Indeed, as the following result shows any vector of payoffs Pareto dominating a Nash equilibrium of the constituent game, g, can arise approximately[17] as the average payoffs of a perfect equilibrium of a game of incomplete information[18] that, with high probability is just a finitely repeated version of g. The result, therefore, is the counterpart for finitely repeated games of incomplete information of Friedman's Theorem C above.

THEOREM 3: *Let* $(e_1, \ldots, e_n)$ *be a Nash equilibrium of the game, g, and let* $(y_1, \ldots, y_n) = g(e_1, \ldots, e_n)$. *For any* $\varepsilon > 0$ *and any* $(v_1, \ldots, v_n) \in V^*$ *such that* $v_i > y_i$ *for all i, there exists $\underline{v}$ such that for any $\nu > \underline{\nu}$ there exists a ν-period sequential game where, with probability $1 - \varepsilon$, player i is described in each period by g_i and in which there exists a sequential equilibrium where player i's average payoff is within ε of* v_i.

REMARK: Notice that the theorem asserts the existence of a *game* as well as of an equilibrium. This enables us to choose the form of the incomplete information.

PROOF: As above, let $\bar{v}_i = \max_{a_1, \ldots, a_n} g_i(a_1, \ldots, a_n)$. Also define $\underline{v}_i = \min_{a_1, \ldots, a_n} g_i(a_1, \ldots, a_n)$. Choose $s = (s_1, \ldots, s_n)$ so that $g(s_1, \ldots, s_n) = (v_1, \ldots, v_n)$.

We will consider a sequential game where each player i can be of two types: "sane," in which case his payoffs are described by g_i, and "crazy," in which case he plays s_i each period as long as s has always been played previously and otherwise plays e_i. Players initially attach probability ε to player i's being crazy and probability $1 - \varepsilon$ to i's being sane. We shall see that early enough in the game, both types of player i play s_i if there have been no deviations from s. Hence, a deviation from s_i constitutes an "impossible" event, one for which we cannot apply Bayes' rule, and so we must specify players' beliefs about i in such an event. We shall suppose that then all players attach probability one to player i's being sane.

Now starting at any point of this sequential game where there has already been a deviation from s, it is clear that one sequential equilibrium of the continuation game consists of all players playing Nash strategies (the e_i's) until the end of the game. We shall always select this equilibrium.

---

[17] The qualification "approximately" is necessary because the game is repeated only finitely more times.

[18] Because the game is one of incomplete information, we must use some sort of Bayesian perfect equilibrium concept. We shall adopt the sequential equilibrium of Kreps and Wilson [15]. According to this concept a player has probabilistic beliefs about other players' types that are updated in Bayesian fashion according to what other players do. An equilibrium is a configuration of strategies as functions of players' types such that, at every point of the game, each player's strategy is optimal for him, given others strategies and his beliefs about their types (actually the concept is a bit more refined than this, but, given the simple structure of our games, this description will do).

Choose ν so that

$$(12) \qquad \underline{\nu} > \max_i \left[\frac{\bar{v}_i - (1 - \varepsilon^{n-1}) \underline{v}_i}{\varepsilon^{n-1}(v_i - y_i)} \right].$$

We will show that in a period with ν periods remaining in the game, where $\nu \geq \underline{\nu}$, a sane player of type i will play s_i if there have been no deviations from s to that point. If that period he plays something other than s_i, his maximum payoff is $\bar{v}_i$. Subsequently his payoff is y_i every period, since, starting from any point after a deviation from s, we always select the "Nash" sequential equilibrium. Thus, if he deviates from s_i with ν periods remaining, an upper bound to i's payoff for the rest of the game is

$$(13) \qquad \bar{v}_i + (\nu - 1) y_i.$$

Suppose, on the other hand, he uses the sequential strategy of playing s_i each period until someone deviates from s and thereafter playing e_i. In that case, his payoff is v_i each period for the rest of the game if the other players are all crazy. If at least one of the other players is not crazy, the worst that could happen to i is that his payoff is $\underline{v}_i$ in the first period and y_i in each subsequent period. Now, assuming that there have been no previous deviations from s, the probability that all the others are crazy is ε^{n-1}. Hence, a lower bound to i's payoff if he uses this sequential strategy is

$$(14) \qquad \varepsilon^{n-1} \nu v_i + (1 - \varepsilon^{n-1})(\underline{v}_i + (\nu - 1) y_i).$$

From (12), (14) is bigger than (13). Hence all players i will play s_i in any period at least $\underline{\nu}$ periods from the end. Thus, for any $\varepsilon > 0$, we can choose ν big enough so that player i's average payoff of the ν-period sequential game is within ε of v_i. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ Q.E.D.

5. THE FOLK THEOREM IN FINITELY REPEATED GAMES OF INCOMPLETE INFORMATION

In this section we strengthen the result of Section 4 by showing roughly that any individually rational point can be sustained (approximately) as the average equilibrium payoffs of a finitely repeated game if the number of repetitions is large enough. This assertion is not quite true for the same reason that the perfect equilibrium counterpart to Theorem A does not hold for three or more players: a discontinuity in $V(\delta)$ can occur if the payoff set is degenerate. For this reason we confine attention to two-player games.[19]

THEOREM 4: *For any $(v_1, v_2) \in V^*$ and any $\varepsilon > 0$ there exists $\underline{\nu}$ such that for any $\nu > \underline{\nu}$ there exists a ν-period sequential game such that, with probability $1 - \varepsilon$,*

---

[19] If we posited full dimension we could also establish the result for three or more players; i.e., we could establish the analog of Theorem 3.

player i is described in each period by g_i and there exists a sequential equilibrium where player i's average payoff is within ε of v_i.

The proof we provide in this section assumes the existence of a one-shot Nash equilibrium that yields both players strictly more than their minimax values. We have established the theorem in general using a similar but more complex argument that is presented in our 1985 working paper.

Briefly, the proof goes as follows: we know that with an infinite horizon our "mutual minimax" strategies of Theorem 1 will enforce any individually rational outcome. The problem with a finite horizon is to avoid the familiar "backwards unraveling" of these strategies from the end. To do so, we introduce the probability ε that a player is "crazy" and will punish his opponent for deviations that would otherwise be too near the end to be deterred by credible (i.e., sequentially rational) threats. More specifically, we partition the game into three "phases." In the first phase, Phase I, players follow the strategies of Theorem 1. That is, they play strategies enforcing the desired outcome unless someone deviates, which triggers mutual minimaxing for β periods, followed by a return to the original path. Deviations during the punishment period restart the mutual punishment. Phase II is a transitional phase. Punishments begun in Phase I are continued, if necessary, in Phase II, but deviations in Phase II are ignored until Phase III. In Phase III, a crazy type plays a Nash equilibrium strategy unless his opponent deviated in Phase II, in which case he plays his minimax strategy. Phase III is an "endgame" in which the crazy types create punishments that do not unravel, and Phase II simply connects this endgame to the strategies of Phase I. The proof shows that by making the last two phases long enough we indeed have an equilibrium, and, moreover, that the required lengths are independent of the total length of the game. Thus if the game lasts long enough, Phase I constitutes most of the game, and our result follows.[20]

PROOF: Let $x_i = g_i(M_1, M_2)$. Clearly $x_i \leq 0$. Let (y_1, y_2) be the expected payoffs to a Nash equilibrium (e_1, e_2) of the one-shot game g, and assume y_1 and y_2 are strictly positive.

As before, we suppose that players can use correlated mixed strategies. Let (s_1, s_2) be correlated strategies yielding payoffs (v_1, v_2). Let β be an integer such that

(15) $\beta \geq \max_i (\bar{v}_i / v_i),$

and, as before, let $\underline{v}_i = \min_{a_1, a_2} g_i(a_1, a_2)$. For given $\varepsilon > 0$, choose an integer α_i so that

(16) $\beta \bar{v}_i + \alpha_i (1 - \varepsilon) y_i < \alpha_i y_i + \underline{v}_i + \beta x_i$

and take $\alpha = \max_i \alpha_i$.

---

[20] We thank a referee for suggesting this simplified form of our earlier proof.

To describe the equilibrium play and the "crazy" player types, we partition the game into three "phases." We will number the periods so that the game ends in period 1. Phase I runs from period ν to period $\alpha + \beta + 1$, Phase II from $(\alpha + \beta)$ to $\alpha + 1$, and Phase III from α to 1.

We will specify crazy behavior recursively, that is, in each period we specify how the crazy player will behave if play to that point has corresponded to the "crazy" play specified for the previous periods, and also how the crazy player will respond to any deviation from that behavior.

Let us begin with Phase I. We define the index $\Psi(t)$ as follows. Set $\Psi(\nu) = \nu + \beta$. In period t, $\nu \geqslant t > \alpha + \beta$, the crazy type (of player i) plays s_i if $\Psi(t) - t \geqslant \beta$, and M_i otherwise. We set $\Psi(t) = \Psi(t+1)$ if there was no deviation from "crazy" behavior in period $t+1$, and $\Psi(t) = t$ otherwise. Thus the crazy type plays s_i until someone deviates. Deviations trigger β periods of minimaxing followed by a return to s_i if there have been no further deviations. Any deviation restarts the mutual punishment portion of the sequence, which runs for β periods after the deviation.

The crazy type follows the same strategy in Phase II as in Phase I, except that deviations in this phase do not change the index $\Psi(t)$. More specifically, in Phase II, $\Psi(t) = \Psi(t+1)$ regardless of play in period $t+1$, and the crazy type plays s_i if $\Psi(t) - t \geqslant \beta$, and plays M_i otherwise. Deviations in Phase II influence behavior in Phase III through a second index variable, Θ. This index has four possible values: $\Theta = 0$ if there have been no deviations from crazy behavior in Phase II; $\Theta = 1$ if only player one deviated; $\Theta = 2$ if only player two deviated; and $\Theta = b$ if both players have deviated. The index Θ is not changed by deviations in Phase III. In Phase III the crazy type plays e_i if $\Theta = 0$, i, or b, and plays M_i if $\Theta = j$. That is, the crazy type of player i punishes his opponent in Phase III for having deviated in Phase II *unless* player i himself also deviated.

Next we describe the behavior of the "sane" types of each player. For a sequential equilibrium we must specify both a strategy for each player, mapping observations into actions, and a system of beliefs, mapping observations into inferences. In Phase I, each sane type's strategy is the same as the corresponding crazy strategy. If his opponent deviates from crazy behavior, the sane player's beliefs are unchanged—he continues to assign the *ex ante* probabilities of ε and $1 - \varepsilon$, respectively, to his opponent being crazy or sane.

In Phase II, if, in state $\Theta = 0$, a player deviates from crazy behavior, his opponent attaches probability one to his being crazy. The strategy of the sane type (of player i) in Phase II if $\Theta = 0$ or j is to play as a crazy type. We do not specify sane play if $\Theta = i$ or b.

In Phase III, the sane type plays e_i if $\Theta = 0$ or j. If player i did not deviate in Phase II, then his beliefs are not changed by play in Phase III. If player i did deviate in Phase II, and player j plays M_j at the beginning of Phase III, j is revealed to be crazy, while if j plays e_j, j is revealed to be sane. We do not specify sane behavior for Phase III if $\Theta = i$ or b except to require that it depend on past outcomes only through the player's beliefs and Θ. Thus we choose some equilibrium for each set of initial beliefs and Θ. The exact nature of this behavior and

TABLE I

Phase	Θ	State $\Psi(t)$	Strategies crazy	sane	Beliefs (probability i attaches to j's being crazy)
I (periods ν to $\alpha + \beta + 1$)	—	$\geqslant \beta + t$	s_i	s_i	ε
	—	$< \beta + t$	M_i	M_i	ε
II (periods $\alpha + \beta$ to $\alpha + 1$)	0	$\geqslant \beta + t$	s_i	s_i	ε
	0	$< \beta + t$	M_i	M_i	ε
	j	$\geqslant \beta + t$	s_i	s_i	1
	j	$< \beta + t$	M_i	M_i	1
	i	$\geqslant \beta + t$	s_i	?	?
	i	$< \beta + t$	M_i	?	?
	b	$\geqslant \beta + t$	s_i	?	?
	b	$< \beta + t$	M_i	?	?
III (periods α to 1)	0		e_i	e_i	ε
	j		M_i	e_i	1
	i		e_i	?	?
	b		e_i	?	?

the behavior in Phase II if $\Theta = i$ or b is irrelevant for our analysis. We know there must exist an equilibrium for each such subgame,[21] and, by deriving upper bounds on player i's payoffs there, we will show that these subgames are not reached on the equilibrium path. Thus, regardless of the form of this "endplay," there is a sequential equilibrium of the whole game in which sane types play as described in Phase I. The specified behavior and beliefs are summarized in Table I.

Now we must show that the specified strategies form a Nash equilibrium in each subgame, and that the beliefs in each period are consistent with Bayes rule. We shall consider whether player one's specified behavior is optimal given his beliefs and player two's specified behavior.

We begin in Phase III. If $\Theta = 0$ or 2, player one expects his opponent to play the Nash strategy e_2 for the duration of the game (recall that if $\Theta = 2$, player one believes player two is crazy), so that the best player one can do is to play his Nash strategy e_1.

Now consider some period t in Phase II, i.e., $\alpha + \beta \geqslant t > \alpha$. First assume $\Theta = 0$. If player one conforms to his specified strategy in Phase II, his payoff each period is either v_1 (if $\Psi(t) - t \geqslant \beta$) or x_1 (if $\Psi(t) - t < \beta$). Thus his lowest possible expected payoff for the remainder of Phase II is $(t - \alpha)x_1$. If he sticks to specified behavior in Phase III as well, he receives αy_1. Thus if player one conforms from period t in Phase II onwards he receives at least

$$(17) \qquad (t - \alpha)x_1 + \alpha y_1.$$

If however player one deviates in Phase II, his highest payoff in that phase is $(t - \alpha)\bar{v}_1$. Then in Phase III, player two plays M_2 if crazy, and e_2 if sane. Thus

---

[21] To establish this we can appeal to the existence theorem of Kreps-Wilson [15], since g is a finite game.

an upper bound to player one's expectation in Phase III is $(1-\varepsilon)\alpha y_1$, and the total payoff to deviating in period t of Phase II is at most

$$(18) \qquad (t-\alpha)\bar{v}_1+(1-\varepsilon)\alpha y_1.$$

Since t is in Phase II, $t-\alpha \le \beta$, and the equation defining α, (16), ensures that deviation is unprofitable.

If $\Theta = 2$ in Phase II, player one is sure that player two is crazy. Thus if player one follows his specified strategy, his payoff is again bounded by (17), while if he deviates in period t his payoff is at most

$$(19) \qquad (t-\alpha)\bar{v}_1+\alpha \cdot 0.$$

Once more, formula (16) ensures that α is large enough so that deviation is unprofitable.

Finally consider a period t in Phase I. From our specification, deviations in Phase I do not change the players' beliefs or the value of Θ. Thus from our previous analysis, both players will conform in Phases II and III regardless of the play in Phase I, so that any sequence of deviations must end at the start of Phase II.

First assume that $\Psi(t) < \beta + t$, so that t is part of a "punishment sequence." If player one conforms in period t and subsequently, his payoff is

$$(20) \qquad (t-\Psi(t)+\beta)x_1+(\Psi(t)-\alpha-\beta)v_1+\alpha y_1.$$

If player one deviates in period t and thereafter conforms, his maximum payoff in period t is zero, and he endures the "punishment" of x_1 for the next β periods, so his payoff is at most

$$(21) \qquad \beta x_1+(t-\beta-\alpha-1)v_1+\alpha y_1,$$

which is less than (20). In particular, player one would never deviate in the last period of Phase I, and, by backwards induction, will not wish to deviate in period t.

Last assume $\Psi(t) \ge \beta + t$, so that player two plays s_2 in period t. If player one deviates in period t but conforms thereafter, he receives at most

$$(22) \qquad \bar{v}_1+\beta x_1+(t-\alpha-\beta-1)v_1+\alpha y_1.$$

If player one conforms to his prescribed strategy, he receives

$$(23) \qquad (t-\alpha)v_1+\alpha y_1.$$

The gain to deviating, the difference between (22) and (23), is thus

$$(24) \qquad \bar{v}_1+\beta x_1-(\beta+1)v_1.$$

Since x_1 is nonpositive, formula (15) defining β ensures that (24) is negative, so player one will not deviate. Thus the specified strategies are indeed in equilibrium. This equilibrium will yield the payoff (v_1, v_2) for $v-\alpha-\beta$ periods, so that by taking ν sufficiently large we can make each player i's average payoff arbitrarily near v_i. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ Q.E.D

Notice that in the proof of Theorem 3 we not only chose the form of "crazy" behavior to suit our needs, but also selected particular conjectures for sane players

when Bayes' rule is inapplicable. We should emphasize that our choice of conjectures was not arbitrary; the theorem is not true if, for example, a player believes his opponent to be *sane* with probability one after a deviation.

Kreps [12], moreover, has pointed out that, because of our choice of conjectures, our equilibrium may not be stable in the sense of Kohlberg–Mertens [11].[22] In response, we offer the following modified version of our construction. This version has no zero probability events, so that the issue of the "reasonableness" of the conjectures and the stability of the equilibrium do not arise. Specifically, assume that at each period in Phase II a crazy player plays as before with probability $(1 - \mu)$, while assigning strictly positive probability to every other pure strategy. If μ is sufficiently near zero, the expected payoffs in every subgame are essentially unchanged, and our strategies are still in equilibrium. Given that the crazy player "trembles" with positive probability in Phase II, any deviation in that phase must reveal that the deviator is crazy, as we specified.

6. UNOBSERVABLE MIXED STRATEGIES

The arguments in Sections 2–5 rely on mixed strategies' being observable. Although this assumption is often used, at least implicitly, in the Folk Theorem literature and can be justified in some circumstances, the more natural hypothesis is that only the moves that players actually make are observed by their opponents. In this section we argue that our results continue to hold with unobservable mixed strategies.

We suggested earlier that the only significant use that our proofs make of the assumption that mixed strategies are observable is in supposing that *minimax* strategies are observable. The heart of the argument, in Theorem 5, therefore, is to show that it suffices for other players to observe the realization of a punisher's random mixed strategy.

Although we rule out observation of private mixed strategies, we continue to assume, for convenience, that strategies can depend on the outcome of publicly observed random variables. We also impose the nondegeneracy assumption of Theorem 2.

THEOREM 5: *Theorem 2 continues to hold when we assume that players can observe only the past actions of other players rather than their mixed strategies.*

PROOF: Choose s, $(v_1', \ldots, v_n')$, $(\nu_1, \ldots, \nu_n)$, $(M^1, \ldots, M^n)$, and w_i^j, $i, j = 1, \ldots, n$ as in the proof of Theorem 2. For each i and j, consider M_i^j, player i's minimax strategy against j. This strategy is, in general a randomization among the m_i^j pure strategies $\{a_i^j(k)\}_{k=1}^{m_i^j}$, where we have chosen the indexation so that,

---

[22] The intuitive basis for Krep's observation is that since the crazy types prefer crazy play, the sane types are "more likely" to deviate from it. Of course, in the games as specified this is not strictly true, but in the "perturbed" versions of the game considered when testing for stability, there would be some deviations that did not increase the opponents's belief that the deviator is crazy.

for each $k = 1, \ldots, m_i^j - 1$,

$$g_i(a_i(k), M_{-i}^j) \leqslant g_i(a_i^j(k+1), M_{-i}^j).$$

For each k, let

$$p_i^j(k) = g_i(a_i^j(k), M_{-i}^j) - g_i(a_i^j(1), M_{-i}^j).$$

The repeated game strategies we shall consider closely resemble those in the proof of Theorem 2. Player i:

(A) plays s_i each period as long as s was played the previous period. If player j deviates from (A), then player i:

(B) plays M_i^j for ν_j periods.

If player i plays pure strategy $a_i^j(k)$ in periods $t_1, \ldots, t_m$ of phase (B), define

$$r_i^j(k) = \sum_{h=1}^m \delta^{t_h - 1} p_i^j(k).$$

Thus, $r_i^j(k)$ is the expected "bonus" that player i obtains from playing $a_i^j(k)$ rather than $a_i^j(1)$ in those periods. Take $r_i^j = \sum_k r_i^j(k)$. Then, r_i^j is the total expected bonus from phase (B). Let

$$z_i^j = \frac{r_i^j(1 - \delta)}{\delta^{\nu_j}}.$$

Because $(v_1', \ldots, v_n')$ is the interior of V^* and V^* has full dimension, there exists $\varepsilon > 0$ so that, for each j,

$$(v_1' + \varepsilon, \ldots, v_{j-1}' + \varepsilon, v_j', v_{j+1}' + \varepsilon, \ldots, v_n' + \varepsilon)$$

is in V^*. Since z_i^j tends to zero as δ tends to 1, we can choose δ big enough so that, for all i and j, $z_i^j < \varepsilon/2$. Then

(25) $\quad (v_1' + \varepsilon - z_i^j, \ldots, v_{j-1}' + \varepsilon - z_{j-1}^j, v_j', v_{j+1}' + \varepsilon - z_{j+1}^i, \ldots, v_n' + \varepsilon - z_n^j)$

is in V^*. If player h deviates from the prescribed behavior in phase (B) the phase is begun again with $j = h$. Player i cannot detect whether player h has deviated from M_h^j, but he can observe whether h has deviated from the *support* of M_h^j. Accordingly, if h so deviates, player i begins phase (B) again with $j = h$. Let $T^j(z) = (T_1^j(z), \ldots, T_n^j(z))$ be a vector of strategies that realizes the payoffs (25) (note that $T^j(z)$ depends on the particular realization of pure strategies in phase (B)). Now suppose that at the conclusion of phase (B), player i:

(C) Plays $T_i^j(z)$ thereafter, and, if player h deviates from (C), then i begins phase (B) again with $j = h$.

The strategies T^j are chosen so that player i will be indifferent among all the pure strategies in the support of M_i^j. The idea is that any expected advantage that player i obtains from using $a_i^j(k)$ rather than $a_i^j(i)$ in phase (B) is subsequently removed in phase (C). Player i then may as well randomize as prescribed by M_i^j. He will not deviate from the support of M_i^j since such a deviation will be detected and punished. *Q.E.D.*

We can also show that Theorem 1 continues to hold with unobservable mixed strategies; we omit the details, except to say that our proof relies on "rewarding" a player who uses a "costly" element of his minimax set with a (small) probability that play will switch from mutual minimaxing to a static Nash equilibrium.

University of California, Berkeley
and
Harvard University

REFERENCES

[1] ABREU, D.: "Repeated Games with Discounting," Ph.D. dissertation, Department of Economics, Princeton University, 1983.
[2] AUMANN, R.: "Subjectivity and Correlation in Randomized Strategies," *Journal of Mathematical Economics*, 1(1974), 67–96.
[3] AUMANN, R., AND L. SHAPLEY: "Long Term Competition: A Game Theoretic Analysis," mimeo, Hebrew University, 1976.
[4] BENOIT, J. P., AND V. KRISHA: "Finitely Repeated Games," *Econometrica*, 53(1985), 890–904.
[5] FRIEDMAN, J.: "A Noncooperative Equilibrium For Supergames," *Review of Economic Studies*, 38(1971), 1–12.
[6] ———: *Oligopoly and the Theory of Games.* Amsterdam: North-Holland, 1977.
[7] ———: "Trigger Strategy Equilibria in Finite Horizon Supergames," mimeo, University of North Carolina, Chapel Hill, 1984.
[8] FUDENBERG, D., AND E. MASKIN: "Nash and Perfect Equilibrium Payoffs in Discounted Repeated Games," mimeo, Harvard University, 1986.
[9] ———: "The Folk Theorem in Repeated Games with Discounting and with Incomplete Information," MIT Working Paper, 1985.
[10] HART, S.: "Lecture Notes: Special Topics in Game Theory," Stanford IMSSS Technical Report, 1979.
[11] KOHLBERG, E., AND J.-F. MERTENS: "On the Strategic Stability of Equilibrium," mimeo, Harvard Business School, 1982.
[12] KREPS, D.: "Signalling Games and Stable Equilibria," mimeo, Stanford Business School, 1984.
[13] KREPS, D., P. MILGROM, J. ROBERTS, AND R. WILSON: "Rational Cooperation in the Finitely-Repeated Prisoner's Dilemma," *Journal of Economic Theory*, 27(1982), 245–252.
[14] KREPS, D., AND R. WILSON: "Reputation and Imperfect Information," *Journal of Economic Theory*, 27(1982), 253–279.
[15] ———: "Sequential Equilibria," *Econometrica*, 50(1982), 863–894.
[16] LOCKWOOD, B.: "Perfect Equilibria in Repeated Games with Discounting," mimeo, Birkbeck College, 1983.
[17] MILGROM, P., AND J. ROBERTS: "Limit Pricing and Entry Under Incomplete Information," *Econometrica*, 50(1982), 443–460.
[18] RADNER, R., R. MYERSON, AND E. MASKIN: "An Example of a Repeated Partnership Game with Discounting and with Uniformly Inefficient Equilibria," *Review of Economic Studies*, 53(1986), 59–70.
[19] RUBINSTEIN, A.: "Equilibrium in Supergames," Center for Mathematical Economics and Game Theory, Hebrew University of Jerusalem, 1977.
[20] ———: "Equilibrium in Supergames with the Overtaking Criterion," *Journal of Economic Theory*, 21(1979), 1–9.