COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

AT YALE UNIVERSITY

Box 2125, Yale University
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 885

THE POWER OF COMMITMENT

by

Chien-fu Chou and John Geanakoplos

September 13, 1988

THE POWER OF COMMITMENT*

by

Chien-fu Chou and John Geanakoplos

History has seen many examples of the lone man—like Christ, Luther, Gandhi, or Hitler—who without initial wealth or position, succeeds in changing the behavior of an entire society, for good or for ill. Whence comes this power? No doubt such leaders have possessed extraordinary ability, and have formulated original ideas with great appeal which others could readily follow. But there is another striking similarity among these leaders; namely their single-minded devotion to their ideals, and their uncompromising attitude toward those who opposed them, no matter what the personal cost. There is hardly any need to document this facet of their personalities, so widely is it known. But we cannot help recalling Gandhi's threat to starve himself to death if the fighting between Hindus and Muslims did not stop. Indeed the whole-hearted commitment of these leaders to their ideals was often reflected in their followers' commitment to them. The purpose of this paper is to show how significant is the power to make commitments, perhaps in the name of some ideal.

Economists have long stressed the idea of commitment (or more precisely conditional commitment) and the related concept of credible threat.

---

*An earlier version (October 1985) of this paper appeared under the title "On Finitely Repeated Games and Psuedo-Nash Equilibria," CFDP #777. Geanakoplos would like to thank Paul Milgrom and Partha Dasgupta for helpful comments.

Thus an incumbent monopolist who can commit himself to flood the market in case a potential rival begins production can thereby protect his monopoly. But if by flooding the market the monopolist lowers price so much he destroys his own profitability, then his threat is not necessarily credible, and a shrewd rival may indeed enter without fear of drastic retaliation. If the monopolist could find some device to guarantee his reputation for retaliation, then he could reap monopoly profits (without ever having to exercise his threat).

It is this familiar economic illustration of the power of commitment that is the basis for this paper, but with one important difference. In the above story the incumbent monopolist had initial position and important wealth, for he had the power, by acting differently, to reverse completely the potential entrant's preferences between entering and not entering. If the incumbent had a very small maximum capacity for production, even a credible threat to increase production to the maximum in case of entry would likely not deter entry. Furthermore the monopolist's threat would seem to be all the more ineffectual the shorter the length of time during which he could commit to the maximum output.

In this paper we focus attention on the circumstances under which an agent with arbitrarily small resources can nonetheless cause drastic changes in the behavior of the rest of his community, through his ability to commit himself in advance to actions he might take much later during an arbitrarily short period of time. Our story requires five ingredients. First, there must be at least two agents that the committed agent can play off against each other. Second, these agents must interact with each other on a great many occasions. Third, these interactions must be observable. Fourth, these interactions must be rapid, or else the agents must not discount the distant future too much, so that identical future and contemporary interactions are regarded as nearly equally valuable. Fifth, the relationship between actions and benefits must be smooth, so that drastic changes in benefits can only be brought about by large changes in actions, and more importantly, small changes in benefits can always be brought about by small changes in actions. For reasons that will become clear, we shall refer to this fifth ingredient as "generic smoothness."

Some of these ingredients have already become standard in the economics literature, and these models form our point of departure. We begin with a one-shot game G among N players, describing their possible actions ("one-shot strategies") and the benefits to each as a function of the joint actions taken by all the players. In such a game with self-interested players, we could not expect any outcome other than a "Nash equilibrium," i.e. a vector of joint actions by the players from which no player can advantageously deviate. Of course this eliminates hypothetical possibilities in which the players could act more cooperatively (or more ruthlessly), perhaps making them all better off (or all worse off), because such joint behavior would allow some individual agent to "cheat," i.e. to deviate advantageously from the hypothetical plan. Two famous examples of such games are the "prisoner's dilemma" and "Cournot oligopoly." We shall consider a game G with N players repeated T times, usually called the supergame G^T. Without any player who can make commitments, it is well-known that when the one-shot game G has a unique Nash equilibrium, the only outcome which is consistent with individual incentives (a "perfect Nash equilibrium") in the supergame G^T is the repeated play of the one-shot Nash equilibrium moves. Our purpose is to show that when generic smoothness holds for G, as it does for Cournot's game but not for the prisoner's dilemma, then if the 1st player, no matter how insignificant, can commit himself just for the last period T, he can arbitrarily alter the equilibrium moves of the other N players during most of the periods of the supergame G^T, if T is large enough.

The paradox that without commitment finite repetitions of a game, no matter how long, may not provide for any more cooperation than the one-shot game itself is a long-standing puzzle which has received a great deal of attention. The so-called "folk theorem," which asserts that if the game is infinitely repeated, then in perfect Nash equilibrium it is possible to observe any distribution of joint moves giving each player more (on average, i.e. per period) than the (Minmax) he could guarantee himself if everyone colluded against him, has only strengthened the sense of paradox, although it is clear why the infinite case is so different. In the infinite case any deviation at date t from a hypothetical path can be punished by joint behavior after date t. Each player (including the original cheater) has an incentive to carry out the punishment, because if he doesn't he can be punished by another, perhaps longer punishment phase etc. (see Rubinstein (1979) for details.) In short, when G is repeated infinitely often, no commitment is necessary to enforce cooperation, since at any stage the short run gains from deviation must be weighed against an infinite future of potential retribution. In the T-fold repetition supergame, players have no incentive not to deviate at the T^{th} period, if there are short run gains to be made, since there is no future. By working backwards one can see that only the repeated one-shot Nash equilibrium can be maintained on the equilibrium path of G^T .

The power of the 1st agent to commit himself to a conditional strategy in the last period breaks the yoke of backward induction. In period T-1 the N agents can be slightly deflected from their one-shot Nash strategies by the small threat the 1st player can generate in period T , if the game G is smooth. What is surprising is that for smooth games this arbitrarily small deflection quickly propagates backward into ever larger deviations from the one-shot Nash strategies in earlier periods. In fact we prove that the ability of the 1st agent to commit himself allows for a "folk theorem" for finitely repeated generic smooth games G^T , even though the commitment is for only one time period, and the 1st agent may be very small compared to the others.

Take as an illustration the Cournot game G with N producers whose one-shot strategies consist of choosing a level of production for the commodity they all produce. Suppose that any subset of N-1 players can produce enough of the good to drive its price down to zero, so that each player's Minmax payoff is zero. Take any vector of quantities $q = (q_1, \ldots, q_n)$ affording each player positive profit. For example, each q_j might be near zero, except for q_1 , which might be near 1's monopoly level of production. If player 1 can commit himself at the beginning of the game to play any conditional strategy for just the period T , depending on what other players have done before time T , then there is a perfect Nash equilibrium during which the players play q in all but the last K periods, no matter how large T is. This proposition holds no matter how small is the maximum capacity for production by agent 1. During the last K periods the production of the N agents along the equilibrium path converges to the one-shot Nash strategies. But for large T the proportion of periods during which q is played approaches 1.

There have been several variations of the supergame G^T in the literature for which "folk theorems" apply when G is repeated a large but finite number of times. All of these variations can be reinterpreted as devices to introduce commitment in a more or less plausible way into the definition of equilibrium. In this sense our paper cuts to the heart of the matter by asking directly what is the minimum amount of commitment necessary to sustain the finitely repeated folk theorem.

Radner (1980) broke the yoke of backward induction by defining an ε-average equilibrium for G^T in which a player is satisfied with his strategy unless he can gain at least ε per period more by deviating, and then showing that repeated full cooperation is an ε-average equilibrium of the T-repeated Cournot game if T is large. If we reinterpret Radner's ε-average equilibrium by supposing that there is an $N+1^{st}$ agent who can commit himself at the beginning of time to give a reward no larger than εT at time T+1 to those who followed a prescribed sequence of moves until period T , then Radner's result is a precursor of ours on the power of commitment. For our purposes, however, his result needs improvement, since it demonstrates that when the size of the reward wielded by the committed player rises proportionately with the number of times he wants others to coöperate, he can indeed guarantee cooperation. For large T this will be an excessive amount of resources that one agent can reasonably be expected to command. Our results show that Radner's theorem can indeed be improved to an ε-(total)-equilibrium theorem, in which players are content so long as they cannot gain more than ε in total by deviating.

Benoit and Krishna (1985) showed that a finitely repeated folk theorem similar to ours applies even without commitment if the one-shot game G has multiple Nash equilibrium payoffs. Their result holds because the agents can jointly credibly conditionally commit themselves in advance to play one or the other Nash equilibrium in period T , and also periods T-1, ..., T-K , depending on the previous moves of the game. Naturally our notion of equilibrium with commitment by the 1st agent for one period can be extended to allow for commitment by any coalition C of agents, including C equal to the set of all agents. With more commitment the "finitely repeated folk theorem" becomes easier to prove. Benoit and Krishna have identified a class of games for which the possibility of commitment (albeit of a special form) is built into the definition of perfect Nash equilibrium, and not just for one player for one period, but for all N players for any number of periods. We shall use our methods to give a brief demonstration of the important Benoit-Krishna theorem.

Krep-Wilson (1982), Milgrom-Roberts (1982), and Kreps-Milgrom-Roberts-Wilson (1982) proposed an "ε-crazy equilibrium" in which with probability ε a player will behave in some arbitrarily specified "crazy" manner. They have shown (see also Fudenburg-Maskin (1986)) that for any $\varepsilon > 0$, and any vector of average payoffs per period exceeding each player's Minmax, for large enough T , there is a specification of craziness for each player leading to a sequential equilibrium with approximately the given payoffs. A notable feature of their proof is that the number of periods $K(\varepsilon)$ that the behavior of the "crazy" version of an agent differs from the equilibrium behavior of his rational version is independent of T. However, it is important to note that $K(\varepsilon)$ varies inversely with ε, so that $\varepsilon \cdot K(\varepsilon)$ is approximately constant. The expected number of periods during which the play of an agent is "crazy," i.e. not governed by self-interest but arbitrarily committed in advance, stays bounded from above, but also from below, no matter how small the probability of craziness is taken. Since the "crazy" behavior may have a significant impact on the utility of other players, each time it occurs, we see once again that in the "ε-crazy" formulation of equilibrium a large amount of "resources" in effect are committed in advance as potential rewards for "good" behavior, and that it is this large commitment which provides for the cooperative equilibrium.

In this paper we concentrate on what we believe is the central phenomenon of repeated games: without commitment, repetition often leads to nothing new. Yet with enough repetition, the slightest bit of commitment can make a world of difference. For smooth N-player games the required amount of commitment is vanishingly small. Our proposition implies that for smooth games the ε-crazy theorem of Fudenberg-Maskin etc. could be strengthened to require crazy behavior for just one period.

In our view, smooth games, in which actions can always be slightly modified, have not received the attention in the folk-theorem literature which their structure merits. Reluctance to exploit the differentiability of the payoff functions is surprising in view of the nearly universal assumption in the folk theorem literature that the one-shot game G has convex strategy spaces with continuous pay off functions (see for example

Rubinstein (1979), Benoit-Krishna (1985), and Fudenberg-Maskin (1986)). Once one has admitted continuity, differentiability is no longer such a big step. Matrix games, in which the strategy spaces are finite, are considered as special cases in the literature by taking their mixed extensions, and assuming that players are allowed to observe the randomizing devices, as well as the moves, of their opponents. It should perhaps be admitted that this is an unpalatable assumption, since it involves a curious timing of events: when i randomizes at time t, the other players j discover "after the fact" at time $t+1$ what randomizing device i used. Fudenburg-Maskin (1986) shows that this extra observability hypothesis does not play a significant role in infinite horizon repeated games, but they leave open the question for finite horizon games.

We analyze the role of observability in folk-theorem proofs because, as we have said, observability is one of the crucial elements of our story. We show that the Benoit-Krishna theorem can be extended to dispense with this extra observability of randomizing devices. This extension is possible because with multiple one-shot equilibria the threat to which agents can credibly commit is unboundedly large. But the minimum commitment necessary to derive a finitely repeated folk theorem for matrix games G with a unique one-shot Nash equilibrium does increase when observability is imperfect. For matrix games G with a (completely) mixed one-shot Nash equilibrium, the payoff functions (to the mixed extension of G) are differentiable around the equilibrium (and it is only around Nash equilibrium that we use differentiability). If the randomizing devices of the agents were observable, than these games would fall into the class of smooth games, to which our first theorem applies. By dropping the hypothesis that the randomizing devices are observable, we study the effect of imperfect monitoring on the power of commitment. We find in general that the commitment of just one player for one period is not enough to achieve a folk theorem, but in two player games with nondegenerate mixed strategy one-shot Nash equilibria, if both players can jointly commit themselves for one period, then again a finitely repeated folk theorem obtains.

In Section 2 we introduce the formal model of repeated games. In Section 3 we describe generic smooth games and state our main theorem. In Section 4 we give its proof. Our method of proof consists of two steps. In Theorem A we show that once a sufficient punishment can be threatened at least once, any game (smooth or not) must display the folk theorem property. The method of Theorem A can be applied in any of the folk theorem contexts, and this shows the similarity of all of these results. The crucial idea in our proof of Theorem A is the concept of a reusable reward system, which explains how a single threat against each player can be used over and over to enforce cooperation without ever being used up. In the second step of Section 4 we show how to construct a single large threat for generic smooth games. In section 5 we note how easy it is to construct a large threat for games with multiple Nash equilibria, and hence to derive the Benoit-Krishna theorem from Theorem A. In Section 6 we examine the importance of monitoring in matrix games.

Before introducing the formal model, let us mention two analogies that our proof suggests. Imagine that the hierarchy of time is replaced by a hierarchy of rank, as in an army of selfish soldiers, where later time periods are analogous to higher rank. A despot, no matter how poor, might be able to induce a small change in the behavior of his generals if he could commit himself to reward them for appropriate behavior. In particular he could induce them to take small measures to punish colonels who disobey orders. The selfish colonels, taking into account the influence of both the despot and the generals, might be induced into slightly more noble behavior, including taking the trouble to punish disobedient majors. The majors will be induced to act still more nobly than the colonels. At the lowest level, the privates, taking into account the force of the entire army apparatus above them, will display the most noble behavior. In this way an army of perfectly selfish soldiers could take on the character of their leader; curiously, the soldiers farthest removed from his influence would be the most affected by his presence. Note the role of continuous strategy spaces in this analogy. If the generals, for example, had only two choices, instead of a continuum, then a small incentive could not be expected to change their choice at all. The chain of influence would be broken at the very first step.

Our results on commitment are also directly analogous to the effects of the social contrivance of money in the overlapping generations economy. Imagine a world, as in Samuelson (1958), in which every generation from $t \geq 0$ consists of a single agent who lives for two periods, until a last generation at time T. If there is only one commodity per period, then there can be no trade. This might be especially ruinous if each agent is rich when young and poor when old. If the government can issue paper money, backed by an arbitrarily small amount of real goods at time T, then for almost all t there can be trade in which the young give nearly half of their endowment to the old, receiving in return half of the youthful endowment of the next generation. Each generation $t \leq T-K$ will consume approximately half its endowment when young, and save the rest of its youthful income in paper money. When old, the generation will sell off its paper money to the next generation, consuming half that generation's youthful endowment. In the last K periods there will be a rapid inflation which will drive the real value of the paper money nearly to zero, so that in period T the government can buy back its paper money at a small cost in real goods. In the last K periods the equilibrium allocations converge toward autarchy (modulo the arbitrarily small amount of goods provided by the government to back the money). A government that can credibly commit itself to future actions, however small, can profoundly affect the allocation of resources for almost every generation.

2. GENERAL DEFINITIONS

2.1. Nash Equilibrium with Commitment

Definition 1. An N-person game G is defined by $G = [\Sigma_i, \Pi_i, i = 1, \ldots, N]$, where the Σ_i are the strategy spaces and the Π_i are the payoff functions for players $i = 1, \ldots, N$. We assume that each Σ_i is compact and convex. Let $\Sigma = \times_{i=1}^{N} \Sigma_i$. We also assume that $\Pi_i : \Sigma \to R$ is continuous, and concave in the i^{th} coordinate.

We have thus restricted our attention to "one-shot" games which have Nash equilibria. The standard matrix games, where each player has a finite number of pure strategies, can be regarded as a special case if we include all the "randomized" strategies for each player. We shall discuss these games in Section 6. A canonical example of a game G satisfying our definition is the Cournot game, where each player must choose a quantity $q_i \in [0, Q_i] = \Sigma_i$, and the payoffs are given by the profits (equal to revenue minus costs) that are obtained in some market.[1] Typically in a Cournot game the Nash payoffs to the sellers are less than the monopoly profits that could be shared if they agreed to cooperate and produce less. (A still worse situation for player i occurs when all the other players play $q_j = Q_j$, the "minmax strategies" against player i.)

Definition 2. The repeated game G^T is defined to be the game that repeats G T times, and whose payoff is the sum of the individual period payoffs. A strategy $\sigma_i^T \in \Sigma_i^T$ for player i in G^T can be represented by

$$\sigma_i^T = (\sigma_i^{(1)}, \sigma_i^{(2)}, \ldots, \sigma_i^{(T)})$$

where $\sigma_i^{(1)} \in \Sigma_i$, and $\sigma_i^{(t)}$ is a function from $\mathop{\times}\limits_{\tau=1}^{t-1} \Sigma$ to Σ_i, $t = 2, \ldots, T$. Given the N-tuple $\sigma^T = (\sigma_1^T, \ldots, \sigma_N^T)$, let $\bar{\sigma}_i^T = (\bar{\sigma}_i^{(1)}, \bar{\sigma}_i^{(2)}, \ldots, \bar{\sigma}_i^{(T)})$ be the realized sequence of actions for player i. The payoff function for G^T is

$$\Pi_i^T = \Pi_i^T(\sigma^T) = \sum_{t=1}^{T} \Pi_i^{(t)} = \sum_{t=1}^{T} \Pi_i(\bar{\sigma}_1^{(t)}, \ldots, \bar{\sigma}_N^{(t)}) .$$

Notice that if we regard Σ_i as the randomized mixtures over a finite set of pure strategies, then the definition we have just given for G^T allows the t-period choice of player i to depend on the randomizing

[1]Of course the inverse demand function $P = P(q_1 + q_2 + \ldots + q_N)$ must be well-behaved and similarly for the cost functions $c_i(q_i)$ so that the profit function for firm i, $\Pi_i(q_1, \ldots, q_N) = q_i P(q_1, \ldots, q_N) - c_i(q_i)$ will be concave in q_i, for each $i = 1, \ldots, N$.

device used by player $j \neq i$ in period $t-1$. We change our definition when we consider matrix games in Section 6.

Definition 3. An N-tuple $\sigma^* = (\sigma_1^*, \ldots, \sigma_N^*)$ of strategies for the repeated game G^T is a Nash Equilibrium (NE) if and only if

$$\Pi_i^T(\sigma_1^*, \ldots, \sigma_i^*, \ldots, \sigma_N^*) \geq \Pi^T(\sigma_1^*, \ldots, \sigma_i, \ldots, \sigma_N^*)$$

for all $\sigma_i \in \Sigma_i^T$, and all $i = 1, \ldots, N$.

We are now ready for our main definition:

Definition 4. The N-tuple σ^* of strategies for the repeated game G^T is a Nash equilibrium with commitment by players $i \in C \subset N$, denoted NE_C , if and only if

$$\Pi_i^T(\sigma_1^*, \ldots, \sigma_i^*, \ldots, \sigma_N^*) \geq \Pi_i^T(\sigma_1^*, \ldots, \sigma_i, \ldots, \sigma_N^*)$$

for all $\sigma_i \in \Sigma_i^T$ such that

$$\sigma_i^{(T)} = \sigma_i^{*(T)} \; , \quad i \in C \; .$$

Notice that any NE is a NE_C where $C = \phi$. If $C \subset D$, then any NE_C is a NE_D . For most of this paper we shall be concerned with the case NE_1 , i.e. when $C = \{1\}$ and only player 1 can commit himself in the last period.

It is also possible to define a perfect Nash equilibrium (perfect NE) and a perfect NE_C of the repeated game G^T . Let $\bar{\sigma}^\tau = (\bar{\sigma}^{(1)}, \ldots, \bar{\sigma}^{(\tau)})$ be any sequence of joint moves in $(\Sigma \times \ldots \times \Sigma)/\tau$, where $\tau < T$. Then the N-tuple of strategies $*\sigma^T = (*\sigma_1^T, \ldots, *\sigma_N^T)$ for the game G^T , together with the given τ-period history $\bar{\sigma}^{\tau}$, defines an N-tuple of strategies $(\bar{\sigma}^{\tau}, *\sigma^{T})$ for the game $G^{T-\tau}$ in the obvious way. If this is an NE (NE_C) for $G^{T-\tau}$, no matter what is τ or $\bar{\sigma}^{\tau}$, then we say that $\sigma*$ is a perfect NE (perfect NE_C) of G^T.

<u>Definition 5</u>. Let (perfect) $NE_C(G^T) = \{x = (x_1, \ldots, x_N) \mid \exists$ a (perfect) NE_C $*\sigma^T$ of G^T such that $\frac{1}{T}\Pi^T(G^T) = x\}$.

<u>Example 1</u>. Prisoners' Dilemma Game

The payoff matrix of a prisoners' dilemma game can be represented by

	Left	Right
Top	(2/3,2/3)	(0,1)
Bottom	(1,0)	(1/3,1/3)

The pure strategy pair $(\bar{\sigma}_1^T, \bar{\sigma}_2^T)$ defined below is a perfect $NE_{\{1,2\}}$ but not a NE of G repeated T times:

$$\bar{\sigma}_1^{(t)} = \begin{cases} \text{Top} & \text{if } \bar{\sigma}_2^{(\tau)} = \text{Left and } \bar{\sigma}_1^{(\tau)} = \text{Top for all } \tau < t, \\ \text{Bottom} & \text{otherwise} \end{cases}$$

and

$$\bar{\sigma}_2^{(t)} = \begin{cases} \text{Left} & \text{if } \bar{\sigma}_1^{(\tau)} = \text{Top and } \bar{\sigma}_2^{(\tau)} = \text{Right for all } \tau < t, \\ \text{Right} & \text{otherwise} \end{cases}$$

Let us now describe a perfect NE_1 of the above prisoners' dilemma game. Rather than explicitly describe the entire strategies formally, we define the intended path and what happens in case of deviations. The reader can easily extend this description to full-fledged strategies. The strategies are as above for $t \leq T-4$. But for $t = T-3$ and $t = T-1$, the players intend to play (Bottom,Left). For $t = T-2$ and $t = T$, the players intend to play (Top,Right). If there is any deviation from the intended path, then both players switch to (Bottom,Right) for the rest of the periods. All subsequent deviations are ignored. Hence, $(2/3,2/3) \in \text{perfect } NE_{\{1,2\}}(G^T)$, and $(2/3,2/3) \in \lim_{T\to\infty} \text{perfect } NE_1(G^T)$, whereas $NE(G^T) = \{(1/3,1/3)\}$ for all T .

2.2. <u>The Folk Theorem</u>

Let $\Sigma_{-j} = \times_{i=1, i\neq j}^{N} \Sigma_i$, for $j = 1, \ldots, N$. The <u>minmax value</u> v_i for player i in the (one-shot) game G is $\min_{\sigma_{-i}\in\Sigma_{-i}} \max_{\sigma_i\in\Sigma_i} \Pi(\sigma_i, \sigma_{-i})$.

Let $C(G)$ be the convex hull of all the payoff N-vectors $(\Pi_1(\sigma), \ldots, \Pi_N(\sigma))$ as σ varies over Σ .

<u>Definition 6</u>. We let $I(G) = C(G) \cap \{(x_1,\ldots,x_N) | x_i \geq v_i, i = 1,\ldots,N\}$.

The <u>folk theorem</u> asserts that the Nash equilibrium average payoffs $NE(G^\infty)$ (and also the perfect NE payoffs) to the <u>infinite</u> repeated game G^∞ are given by the set $I(G)$. When the strategy spaces of G are not convex, then $I(G)$ must be modified to take into account randomization.

<u>Definition 7</u>. $x = (x_1, \ldots, x_N)$ is called a (perfect) NE_C (average) payoff, and we let $x \in \lim_{T\to\infty} \text{perfect } NE_C(G^T)$, if and only if there exists a sequence x^T where x^T is a (perfect) NE_C average payoff for G^T , and $\lim_{T\to\infty} x^T = x$.

3. SMOOTH GAMES

We shall now specialize our analysis to smooth games G . Here the strategy spaces Σ_i are taken to be compact manifolds (perhaps with boundary), of dimension ℓ_i . For simplicity the reader can think of them as rectangles in R^{ℓ_i} , respectively. The payoff functions $\Pi_i : \Sigma_1 \times \ldots \times \Sigma_N \to R$ are continuously differentiable, arbitrarily many times; thus we shall write $\partial\Pi_i/\partial s_j$ and $\partial^2\Pi_i/\partial s_i\partial s_j$ and so on. We shall always suppose that G has a one shot Nash equilibrium $\bar{s}$ that lies in the interior of $\Sigma_1 \times \ldots \times \Sigma_N$. The classic example of such a game is the Cournot game, where Π_i is the profit function of firm i , depending on the outputs $s_j \in \Sigma_j = [0, \bar{Q}_j]$ of all the players $j \in N$, and the costs of production, $C_i(s_i) = a_i s_i + \frac{1}{2} b_i s_i^2$ of player i , where a_i and b_i are arbitrary parameters.

We shall always impose three restrictions on the games G that we analyze. It is very important to note, as we shall state more formally in a moment, that these properties hold for nearly any smooth game G taken at random (i.e. they hold generically).

We shall focus attention on one-shot games G that have an interior one-shot Nash equilibrium $\bar{s}$ satisfying

$$(1) \qquad F = \left[\frac{\partial\Pi_i}{\partial s_j} \right] = \begin{bmatrix} \frac{\partial\Pi_1}{\partial s_1}, & \ldots, & \frac{\partial\Pi_1}{\partial s_N} \\ \\ \frac{\partial\Pi_N}{\partial s_1}, & \ldots, & \frac{\partial\Pi_N}{\partial s_N} \end{bmatrix}$$

has full row rank, when evaluated at $\bar{s}$. (Note that since Σ_j may have dimension ℓ_j greater than one, we are interpreting $\partial\Pi_i/\partial s_j$ to be a row vector of dimension ℓ_j .)

Restriction 1 implies that the Nash equilibrium $\bar{s}$ is not Pareto optimal. There are always small deviations $(ds_1, \ldots, ds_N)$ that will make all players better off. In the Cournot game if each player agreed to produce a little less, they would all be better off. Restriction 1 also implies the stronger property, which in the next section we formally define as full dimensionality, that for each player i there is some deviation $(ds_1, \ldots, ds_N)$ that makes only player i better off. In the Cournot game that could happen if all players $j \neq i$ produced a little less, while i produced more.

Let

$$(2) \qquad H = \left[\frac{\partial^2 \Pi_i}{\partial s_i \partial s_j} \right] = \begin{bmatrix} \dfrac{\partial^2 \Pi_1}{\partial s_1 \partial s_1}, & \ldots, & \dfrac{\partial^2 \Pi_1}{\partial s_1 \partial s_N} \\ \dfrac{\partial^2 \Pi_N}{\partial s_N \partial s_1}, & \ldots, & \dfrac{\partial^2 \Pi_N}{\partial s_N \partial s_N} \end{bmatrix} ,$$

evaluated at $\bar{s}$. Then H_{-j}^{-j} has full rank for each $j = 1, \ldots, N$, where by H_{-j}^{-j} we mean the matrix obtained from H by deleting all the rows corresponding to Π_j , and all the columns corresponding to s_j .

Restriction 2 implies that the first order conditions $\partial\Pi_i/\partial s_i = 0$ defining the one-shot Nash equilibrium $\bar{s}$ are non-degenerate. If we fix the moves of any player j at $\hat{s}_j$ near $\bar{s}_j$, then we can solve for the Nash moves of the remaining players in a one-shot play of the game in which they all take for granted that $s_j = \hat{s}_j$.

$$(3) \qquad \text{Let } \frac{d\Pi_i}{ds_j} = -(F_i^{-j})' \left[H_{-j}^{-j}\right]^{-1} H_{-j}^{j} + \frac{\partial \Pi_i}{\partial s_j} , \quad \text{evaluated at } \bar{s}_1 ,$$

where by F_i^{-j} we mean the i^{th} row of F, deleting its j^{th} entry, and by H_{-j}^{j} we mean the j^{th} column of H, deleting its j^{th} entry. Then for any player i we can find players $i_1, \ldots, i_k$ such that

$$\frac{d\Pi_{i_1}}{ds_1} \neq 0,\ \frac{d\Pi_{i_2}}{ds_{i_1}} \neq 0,\ \ldots,\ \frac{d\Pi_i}{ds_{i_k}} \neq 0 .$$

The expression $(d\Pi_i/ds_j)\cdot ds_j$ represents the change from $\Pi_i(\bar{s})$ in player i's one-shot payoff, when player j commits himself to $\hat{s}_j = \bar{s}_j + ds_j$, and then all other players adjust optimally to their one-shot Nash moves. From restriction 2 we know that the matrix H_{-j}^{-j} is indeed invertible, so that these adjustments are well-defined (for small ds_j).

Restriction 3 means that player 1 can directly, or indirectly, affect the utility of every player (including himself). In the Cournot game player 1 can directly hurt every other player by producing more. In the army a general might only be able to discipline his direct subordinates, but they in turn might be able to punish their subordinates, etc. Restriction 3 is clearly indispensable to our theorem concerning the power of player 1's commitment.[2]

---

[2]Although the naturalness (and necessity for our theorem) of restrictions 1-3 is apparent, it is worth noting formally that these restrictions are generic in a precise sense. Let us index the set of possible games G by a set of parameters A. We now suppose that the payoff functions Π_i are smooth on the larger domain: $\Pi_i : \Sigma_1 \times \ldots \times \Sigma_N \times A \to R$. Suppose furthermore that for any choice of joint moves and parameters, $(s_1, \ldots, s_N, a)$, and for each player i, there is a direction a_i in A such that $(\partial/\partial a_i)[\Pi_i(s,a)/\partial s_i] \neq 0$ while the derivatives of the payoffs of the other players are unaffected by changes in a_i. We saw an example of this in the Cournot game when a_i was the linear parameter of

We are now ready to state our main theorem. By generic smooth game we mean any smooth game G that has a one-shot Nash equilibrium $\bar{s}$ satisfying restrictions 1-3.

Theorem 1. For any generic smooth game G,

$$\lim_{T\to\infty} (\text{perfect})NE_1(G^T) = NE(G^\infty) .$$

We can describe the same result more sharply as:

Corollary 1 (of proof). Let $s \in \Sigma_1 \times \ldots \times \Sigma_N$ be a joint one-shot strategy such that $\Pi_i(s) > v_i$ for $i = 1, \ldots, N$, in the generic smooth game G. Then there is a K such that for any $T > K$, there is σ^T, a (perfect) NE_1 of G^T such that along the equilibrium path s is played in each of the first $T-K$ periods. For $t > K$, $\hat{s}_t$ is played, where each $\hat{s}_t$ is near a one-shot Nash equilibrium $\bar{s}$ of G.

Again, we restate the proposition to show its relationship to Radner's ε-average deviation theorem.

---

firm i's cost function. Finally, let us suppose that there is a direction b_i with $(\partial/\partial b_i)[\partial^2\Pi_i(s,a)/\partial^2 s_i] \neq 0$, while all other payoff derivatives are unaffected by changes in b_i, except possibly for player i himself. Again we saw an example of this in the Cournot game where b_i is the quadratic parameter of cost function. Finally, we suppose that for any player i we can find players $i_1, \ldots, i_k$ such that $\partial\Pi_{i_1}(s,a)/\partial s_1 \neq 0$, $\partial\Pi_{i_2}(s,a)/\partial s_{i_1} \neq 0, \ldots, \partial\Pi_i(s,a)/\partial s_{i_k} \neq 0$. Under these assumptions on the set of parameters A indexing our collection of games, a routine application of the transversality theorem [see for example Dubey-Rogawski, 1983] shows that for an open, dense set $\tilde{A} \subset A$, of full Lebesgue measure, any game G_a indexed by parameters $a \in \tilde{A}$ satisfies restrictions 1, 2, and 3 at each of its interior Nash equilibria.

Corollary 2 (of proof). Let G be a generic smooth game with Nash equilibrium $\bar{s}$, and Minmax payoffs $(v_1, \ldots, v_N)$. Let $s \in \Sigma_1 \times \ldots \times \Sigma_N$ satisfy $\Pi_i(s) > v_i$, for $i = 1, \ldots, N$. Let $\varepsilon > 0$ be given, along with an open set $0 \subset \Sigma_1 \times \ldots \times \Sigma_N$ containing $\bar{s}$. Then there is a K such that for all $T > K$, we can find joint strategies σ^{T-1} for G^{T-1} such that for each $1 \leq t \leq T-K$, the players play s along the intended path of σ^{T-1}, while for $T-K < t \leq T-1$ they play $\hat{s}_t$ with $\hat{s}_t \in 0$. Furthermore, the total that any players can gain by deviating from the intended path of σ^{T-1}, or from any other point of the intended path, is less than ε.

Thus for generic smooth games, Radner's ε-average-equilibrium theorem can be strengthened to an ε-total equilibrium theorem.

4. THE MAIN ARGUMENT AND SOME MORE USEFUL DEFINITIONS

4.1. Enforceability and Consecutive Deviations

We continue with the definitions of Section 2, then derive some lemmas leading to the proof of Theorem 1. Let $(r_1, \ldots, r_N)$ be a vector of nonnegative numbers which we call rewards. Clearly we shall use rewards to enforce cooperation. However, if we wish to enforce "perfect" cooperation, then we must be prepared to punish deviations off the equilibrium path, and to punish deviations from deviations, etc. With this in mind, we give:

Definition 8. The strategies $*\sigma^T$ for the game G^T are said to be enforceable by the reward structure $(r_1, \ldots, r_N)$ if
$\Pi_i^T(\sigma_i^T, *\sigma_{-i}^T) - \Pi_i(*\sigma^T) \leq r_i$ for all $\sigma_i^T \in \Sigma_i^T$, and all $i = 1, \ldots, N$.
Similarly, if for any τ period history $\bar{\sigma}^\tau$, with $\tau \leq T$, if

$(\bar{\sigma}^{\tau}, *\sigma^{T})$ is enforced in $G^{T-\tau}$ by the reward structure $(r_1, \ldots, r_N)$, then we say that $(r_1, \ldots, r_N)$ <u>perfectly enforces</u> $*\sigma^{T}$.

Note that $*\sigma^{T}$ is a (perfect) NE of G^{T} if and only if it is (perfectly) enforceable by the reward structure (0, 0, ..., 0).

Let $*\sigma^{T}$ be an N-vector of strategies for G^{T}. Given a history $\bar{\sigma}^{\tau}$, up until $\tau \leq T$, we shall say that player i <u>deviated</u> from $*\sigma_i^{T}$ at time τ if $*\sigma_i^{(\tau)}(\bar{\sigma}^{\tau-1}) \neq \bar{\sigma}_i^{(\tau)}$ and $\sigma_j^{(\tau)}(\bar{\sigma}^{\tau-1}) = \bar{\sigma}_j^{(\tau)}$ for $j < i$. Furthermore, we shall say that player i was the <u>last deviator</u> from $*\sigma^{T}$ up until time τ if there is a time $t \leq \tau$ at which he was the deviator, and if there is no time t', $t < t' \leq \tau$ at which some other player $j \neq i$ was the deviator.

<u>Definition 9</u>. The strategies $*\sigma^{T}$ for the game G^{T} are said to <u>prevent consecutive deviations</u> if for any $\tau \leq T$ period history $\bar{\sigma}^{\tau}$, if i was the last deviator from $*\sigma^{T}$ up until τ, then $\Pi_i^{T-\tau}(\bar{\sigma}^{\tau}, (\sigma_i^{T}, *\sigma_{-i}^{T})) \leq \Pi_i^{T-\tau}(\bar{\sigma}\tau, *\sigma^{T})$ for all $\sigma_i^{T} \in \Sigma_i^{T}$.

"Trigger strategies" are the most famous example of such strategies. Let $(*s_1, \ldots, *s_N) = *s$ be a one-shot Nash equilibrium for G, and let $\Pi_i(s_1, \ldots, s_N) > \Pi_i(*s)$. The trigger strategy $*\sigma_i$ for player i in G^{T} is $*\sigma_i^{(\tau)}(\bar{\sigma}^{\tau-1}) = s_i$ if $\bar{\sigma}^{(t)} = s$ for all $t \leq \tau-1$, and $*\sigma_i^{(\tau)}(\bar{\sigma}^{\tau-1}) = *s_i$ otherwise, for $\tau = 1, \ldots, T$. Trigger strategies have the special property that once one player has deviated, then <u>no</u> player can subsequently deviate advantageously.

The first ingredient in our deviation of a "finitely repeated folk theorem" is a general lemma showing that the same finite reward structure $(r_1, \ldots, r_N)$ can be used to perfectly enforce cooperation for an arbitrarily long time. Please note that the lemma does <u>not</u> depend on any properties of the game G (like convexity or smoothness), except that it assumes G has a (perhaps mixed strategy) Nash equilibrium and it assumes that the minmax strategies μ_j^i that all players $j \neq i$ play against i are observable, i.e. $\mu_j^i \in \Sigma_j$.

Consider the one-shot game G with minmax payoffs $v = (v_1, \ldots, v_N)$, $v_i = \Pi_i(\mu_i^i, \mu_{-i}^i) \geq \Pi_i(\sigma_i, \mu_{-i}^i) \ \forall \sigma_i \in \Sigma_i$, and a Nash equilibrium $*s$ with payoffs $\Pi_i(s*)$. Without loss in generality we take $\underset{\sigma\in\Sigma}{\text{Min}}\, \Pi_i(\sigma) = 0$ and $\underset{\sigma\in\Sigma}{\text{Max}}\, \Pi_i(\sigma) = 1$, for all $i = 1, \ldots, N$.

<u>Lemma 1</u>. Let $\vec{s}$ be an N-tuple of one-shot strategies for G with payoffs $\Pi_i(\vec{s}) = x_i$ satisfying $x_i > v_i$ for all $i = 1, \ldots, N$. Let $W = \{i \in N | \Pi_i(*s) > v_i\}$, let $\bar{x}_i = \text{Min}(x_i, \Pi_i(*s))$, and let $K = \underset{i\in W}{\text{Max}}[1/(\bar{x}_i - v_i)] + 1$. Then for any period t it is possible to devise strategies σ^T that (1) do not permit consecutive deviations and (2) are perfectly enforced by the reward structure (K, ..., K) and (3) yield a history $\bar{\sigma}^T$ with realization $\vec{s}$ in at least T–K periods.

<u>Proof</u>. We shall give a sketch--the proof is obvious except for notation.

Let $T > K$ (otherwise there is nothing to prove). The <u>intended path</u> means each player i plays $\vec{s}_i$ at each date $\tau \leq T-K$, and plays $*s_i$ for $T-K < \tau \leq T$.

Let W be the set of players with $v_i < \Pi_i(*s)$, and let B be the rest, all with $v_i = \Pi_i(*s)$.

If a player $i \in B$ deviates from the intended path, or from any other phase we shall subsequently define, then all players should play $*s$ until the end. Deviations from this path by any player are ignored.

Clearly once player $i \in B$ has deviated, no player can again advantageously deviate. Player $i \in B$ can gain at most 1 by deviating from the intended path.

If a player $i \in W$ deviates from the intended path at time $\tau \leq T-K$, then from $\tau+1$ to $\tau+K$ all other players should play μ^i. After $\tau+K$, all players should return to the intended path. If during period $\tau+1$ to $\tau+K$ a player $j \in W$, $j \neq i$ deviates from the intended punishment of i (playing anything different from that specified by μ_j^i) then play returns immediately to the intended path. If i himself deviates from his own punishment phase, then the punishment continues as before. It is easy to see that no player $i \in W$ can advantageously deviate from the intended path, or from his own punishment. Once i deviates from the punishment of $j \in W$, play immediately returns to the intended path. Hence $i \in W$ cannot deviate consecutively, advantageously, since he also has no opportunity to deviate following a deviation of $k \in B$.

Since the most that any player i can gain by deviating from the punishment phase of $j \in W$ is K (and is zero if $j \in B$), it follows that the behavior of $i \in W$ is perfectly enforced by the reward K.

Q.E.D.

Lemma 1 can easily be extended to average payoffs $(x_1, \ldots, x_N) \gg (v_1, \ldots, v_N)$ that are rational combinations of payoffs from strategy N-tuples $\vec{s}$ and $\vec{\vec{s}}$ for the one-shot game G. Let $x = (n_1/C)\Pi(\vec{s}) + (n_2/C)\Pi(\vec{\vec{s}})$. The intended path must now be a sequence of L-cycles, in any of which $\vec{s}$ occurs n_1 times and $\vec{\vec{s}}$ n_2 times. The punishment phase shall last $K = [\underset{i \in W}{\text{Max}}[1/(\bar{x}_i - v_i)] + 1]$ cycles, i.e. be KL periods long, and r_i must be at least KL.

As T gets large, the K (or KL) periods during which x is not realized becomes negligible.

4.2. Reusable Reward Systems and the Continuation Property

We now turn to the question of how the reward structure $(r_1, \ldots, r_N)$ can be established. When we deal with perfect equilibria, this is by no means a simple idea. In a perfect equilibrium it is necessary always to be able to threaten the last deviator. It may be very simple to arrange a punishment phase for i if he deviates from the status quo. But if j deviates from that punishment phase, then it may be necessary to arrange an even longer phase to punish j. The problem is eliminated if we can establish the existence of a reusable reward system.

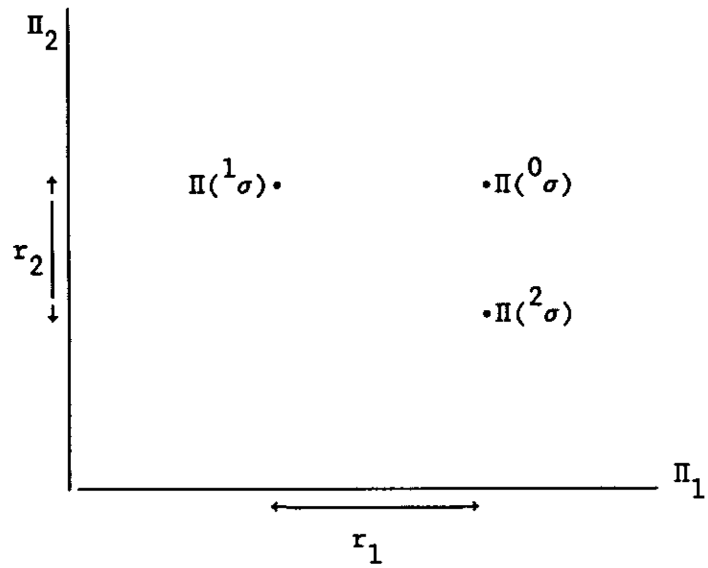
Definition 10. A usable, perfect, T-period, NE_C reward system with reward structure $(r_1, \ldots, r_N)$ is a set of $N+1$ perfect NE_C's $({}^0\sigma, {}^i\sigma, i = 1, \ldots, N)$ for the game G^T satisfying $\Pi_i^T({}^0\sigma) - \Pi_i^T({}^i\sigma) \geq r_i$, for all $i = 1, \ldots, N$.

Definition 11. A reusable, (perfect), T-period NE_C reward system ρ^T with reward structure $(r_1, \ldots, r_N)$ is a set of $N+1$ (perfect) NE_C's $({}^0\sigma, {}^i\sigma, i = 1, \ldots, N)$ for the game G^T which satisfy $\Pi_i^T({}^j\sigma) - \Pi_i^T({}^i\sigma) \geq r_i$ for all $i = 1, \ldots, N$, $j = 0, 1, \ldots, N$, $i \neq j$.[3]

---

[3] Our reusable reward system is in some sense the finite-time horizon analogue of Abreu's [1982] simple penal code for infinite horizon games. A similar device for infinitely repeated games was also used by Fudenberg-Maskin [1986].

The following diagram displays the payoffs to two players of 3 NE_C's forming a reusable reward system: Note that each ${}^i\sigma = ({}^i\sigma_1, \ldots, {}^i\sigma_N)$ is itself an N-vector of strategies, for $i = 0, \ldots, N$.



Notice that as long as the intended perfect NE_C is $\sigma \in \{{}^0\sigma, {}^1\sigma, {}^2\sigma\}$ but $\sigma \neq {}^i\sigma$, then player i can be threatened with a loss of r_i if play switches to ${}^i\sigma$. Since ${}^i\sigma$ is itself a (perfect) NE_C, this threat is credible.

Suppose that the strategies $\sigma^{T'}$ for the game $G^{T'}$ do not allow consecutive deviations, and suppose furthermore that they can be perfectly enforced by the reward structure $(r_1, \ldots, r_N)$. Let $\rho^{T''}$ be a reusable, perfect, T"-period NE_C reward system with reward structure $(r_1, \ldots, r_N)$. Define the strategies $*\sigma^{T'+T''}$ in the game $G^{T'+T''}$ by $*\sigma_i^{(\tau)} = \sigma_i^{(\tau)}$ for all players i and $\tau \leq T'$. Furthermore, from T'+1 to T'+T" the players follow ${}^i\sigma$ if i was the last player to deviate from $\sigma^{T'}$ up until time T' . Otherwise they follow ${}^0\sigma$.

Definition 12. The continuation property asserts that if $*\sigma^{T'}$ prevents consecutive deviations in the game $G^{T'}$ and is (perfectly) enforceable by the reward structure $(r_1, \ldots, r_N)$, and if $\{{}^0\sigma, {}^1\sigma, \ldots, {}^N\sigma\}$ is a (perfect) reusable NE_C reward system for the game $G^{T''}$ with reward structure $(r_1, \ldots, r_N)$, then their combination as above in the game $G^T = G^{T'+T''}$ is a (perfect) NE_C. Moreover, if the ${}^j\sigma$ are all (perfect) NE's, then the combination is a (perfect) NE in the game G^T. We denote the combined strategy for the game $G^{T'+T''}$ by $(*\sigma^{T'}, \rho^{T''})$.

The continuation property hardly needs proof.

4.3. Full-Dimensionality and Reusable Reward Systems

Definition 13. We say that the game G is full-dimensional iff the convex hull of the set of one-shot payoffs to G, $C(G) = C_0\{\Pi(\sigma) = (\Pi_1(\sigma), \ldots, \Pi_N(\sigma)) | \sigma \in \Sigma\}$ has nonempty interior in R^N.

Lemma 2 shows that when G is full-dimensional, usable reward systems can be converted into reusable reward systems.

Lemma 2. Suppose for each T it is possible to construct a usable, perfect, NE_C reward system $\rho^T = ({}^0\sigma^T, {}^1\sigma^T, \ldots, {}^N\sigma^T)$ for G such that $\lim_{T\to\infty} \frac{1}{T}\Pi({}^i\sigma^T) = {}^i\Pi$ exists for each $i = 0, 1, \ldots, N$, and ${}^0\Pi >> \Pi_{min} = ({}^1\Pi_1, {}^2\Pi_2, \ldots, {}^N\Pi_N)$. Then if G is full-dimensional, there is a sequence of T-period reusable, perfect, NE_C reward systems with reward structures $(r_1^T, \ldots, r_N^T)$ satisfying $\lim_{T\to\infty} \frac{1}{T} r_i^T = r_i > 0$.

__Proof__. Note first that since $\frac{1}{T}\Pi({}^0\sigma^T) \in C(G)$, ${}^0\Pi \in \overline{C(G)}$.

From the full-dimensionality hypothesis, we know that there is an open set of vectors in $C(G)$ arbitrarily near ${}^0\Pi$. In particular, they may all be taken to strictly dominate Π_{min} . We can always choose N of them to have the form $x - \varepsilon e_i$, $i = 0, 1, \ldots, N$, where $e_0 = 0$ and e_i is the i^{th} unit vector, $i = 1, \ldots, N$, for some vector $x \in C(G)$, and some small positive ε .

Suppose that for each $i = 0, 1, \ldots, N$, $x - \varepsilon e_i = \Pi({}^i s)$ for some ${}^i s \in \Sigma$. Fix T large enough so that the convergence to each ${}^i\Pi$, $i = 1, \ldots, N$ described in the lemma is within $\varepsilon/4$ in each coordinate, and such that $T\cdot(\varepsilon - \varepsilon/4) \gg (1, \ldots, 1)$. Then for $\tilde{T} > T$ we can denote the NE_C ${}^i\tilde{\sigma}^{\tilde{T}}$ by $({}^i s, \ldots, {}^i s, {}^0\sigma^T)$, by which we mean that all the players are intended to play ${}^i s$ for the first $\tilde{T}-T$ periods. If there is a deviation by player j in one of the first $\tilde{T}-T$ periods, say in period K , then play switches to ${}^j\sigma^{\tilde{T}-K}$ for the remainder of the game. If there has been no deviation until period $\tilde{T}-T$, then the remainder of the game follows strategy ${}^0\sigma^T$. One can verify immediately that this is a perfect NE_C, and hence $\tilde{\rho}^{\tilde{T}} = ({}^0\tilde{\sigma}^{\tilde{T}}, \ldots, {}^N\tilde{\sigma}^{\tilde{T}})$ is a reusable reward system and for all $i = 0, 1, \ldots, N$,

$$\lim_{\tilde{T}\to\infty} \frac{1}{\tilde{T}} \Pi({}^i\tilde{\sigma}^{\tilde{T}}) = x - \varepsilon e_i .$$

More generally, if $x - \varepsilon e_i$ can only be written approximately as a rational combination of payoffs in $C(G)$, $x - \varepsilon e_i \approx (n_1/n)\Pi({}^1 s) + \ldots + (n_K/n)\Pi({}^K s)$ then simply replace each ${}^i s$ in the above construction with a cycle of n_1 periods of ${}^1 s$,

n_2-periods of ^2s, ..., and n_K periods of ^Ks. Take T so large that $(r_1^T, \ldots, r_N^T) \gg (n, n, \ldots, n)$. Q.E.D.

Combining Lemmas 1 and 2 we get the following theorem.

Theorem A: Let $G = [\Sigma_i, \Pi_i, i = 1, \ldots, N]$ be a full-dimensional N-person game. Suppose that for any $(r_1, \ldots, r_N)$ it is possible to construct a usable NE_C reward system for G^T, for some T, with reward structure at least $(r_1, \ldots, r_N)$. Then $\lim_{T\to\infty}(\text{perfect})NE_C(G^T) = NE(G^\infty)$.

4.4. The Proof of Theorem 1

From Lemmas 1 and 2, and the continuation property, it suffices, to prove the "finitely repeated folk theorem" for any game G, to show that it is possible to construct a sequence of usable reward systems ρ^T with arbitrarily large rewards, satisfying Lemma 2.

Lemma 3. Let G be a generic smooth game. Then for each T it is possible to construct a usable, perfect, NE_1 reward system $\rho^T = ({}^0\sigma^T, \ldots, {}^N\sigma^T)$ for G such that $\lim_{T\to\infty} \frac{1}{T} \Pi({}^i\sigma^T) = {}^i\Pi$ exists for each $i = 0, 1, \ldots, N$, and ${}^0\Pi \gg \Pi_{min} = ({}^1\Pi_1, {}^2\Pi_2, \ldots, {}^N\Pi_N)$.

Proof. Let $\bar{\Pi} = \Pi(\bar{s})$ be the N-vector of payoffs to all the players at a Nash equilibrium $\bar{s}$ in the one-shot generic smooth game G. Since F has full row rank by restriction 1, it follows that for all small ε it is possible to find small finite changes $d\bar{s} = (d\bar{s}_1, \ldots, d\bar{s}_N)$ such that $\Pi(\bar{s} + d\bar{s}) = \bar{\Pi} + \varepsilon e$, where $e = (1, \ldots, 1)$. Since $\bar{s}$ is a Nash equilibrium, and the strategy spaces Σ_i are compact, and the payoffs Π_i

are continuous, the maximum gain any player can obtain by deviating from $\bar{s} + d\bar{s}$ is an amount $f(\varepsilon)$ that tends to zero as ε tends to zero.

Similarly, for any pair of players i and j with $d\Pi_i/ds_j \neq 0$, there is some $d\bar{s}_j(i,j)$ and $d\bar{s}_k(i,j)$, $k \neq j$, such that $\hat{s}(i,j) = \bar{s} + d\bar{s}(i,j) = \bar{s} + (d\bar{s}_1(i,j), \ldots, d\bar{s}_N(i,j))$ is a Nash equilibrium for players $k \neq j$, given that j is committed to playing $\hat{s}_j(i,j) = \bar{s}_j + d\bar{s}_j(i,j)$. Furthermore we may suppose that player i's payoff satisfies $\Pi_i(\hat{s}(i,j)) = \bar{\Pi}_i - \varepsilon(i,j)$, and that the most player j can get by deviating is $f_j(\varepsilon(i,j))$, which tends to zero as $\varepsilon(i,j)$ tends to zero.

Fix $T > N$. One perfect NE_1 for G^T is simply $\bar{s}$ repeated T times. A second NE_1 for G^T is defined as follows. For each period $t \leq T-N$, the intended play is $(\bar{s} + d\bar{s})$. For each $t > T-N$, the intended play is $\bar{s}$. Deviations in periods after $T-N$ are ignored. If there is a deviation in period $t \leq T-N$ by player i, then play switches to $\bar{s}$ from period $t+1$ until $T-N$. From period $T-N+1$ until T, the following occurs.

For each player i we know by restriction 3 that there is a sequence of players $i_1, i_2, \ldots, i_K$ through which player 1 indirectly affects player i. To ease notation let us suppose $i = N$, and the sequence of players is $2, \ldots, N-1$. Choose numbers $\varepsilon(2,1), \varepsilon(3,2), \ldots, \varepsilon(N, N-1)$ such that $\varepsilon < \varepsilon(N, N-1)$, $f_{N-1}(\varepsilon(N, N-1)) < \varepsilon(N-1, N-2)$, $f_{N-2}(\varepsilon(N-1, N-2)) < \varepsilon(N-2, N-3), \ldots, f_2(\varepsilon(3,2)) < \varepsilon(2,1)$. By the continuity of the various f_j functions, this can always be arranged. In period $T-(N-1)$, $\hat{s}(N, N-1)$ is played. If player $N-1$ does not deviate, then play continues with $\bar{s}$ for the remaining periods. Further deviations are ignored. If player (N-1) does deviate at time T-(N-1), then in period T-(N-2) play continues with $\hat{s}(N-1, N-2)$. Afterwards intended play is $\bar{s}$, unless N-2 deviates at time T-(N-2), ... Finally, if player 2 deviates at time T-2 from $\hat{s}(3,2)$, then in period T-1 $\bar{s}$ is played, but in the final period T, $\hat{s}(2,1)$ is played. Note that we left room for one extra period in case the deviator at $t \leq T-N$ was player 1 himself. We might have needed a chain 1, 2, ..., N, 1.

It is easy to verify that the second construction is an NE_1, and that for large T the average payoffs are approximately $\bar{\Pi} + \epsilon e$. Let this be the sequence ${}^0\sigma^T$. And for each i = 1, ..., N, let ${}^i\sigma^T$ be the T-fold repetition of $\bar{s}$, giving rise to average payoff $\bar{\Pi}$. Q.E.D.

5. PERFECT NASH EQUILIBRIA WHEN G HAS MULTIPLE ONE-SHOT NASH EQUILIBRIA

The method of proof we outlined in Section 4 can be applied in a variety of contexts. In this section we take $C = \phi$ and derive the theorem of Benoit-Krishna that replaces our generic smoothness hypothesis with the assumptions that G is full-dimensional and has at least two Nash equilibria which yield each player different payoffs.

Theorem 2 (Benoit-Krishna). Let G satisfy the conditions of Section 2. Let G be full dimensional, and let G have L one-shot Nash equilibria ${}^1\bar{s}, \ldots, {}^L\bar{s}$ whose payoff N-tuples ${}^\ell\bar{\Pi} = \Pi({}^\ell\bar{s})$, $\ell = 1, \ldots, L$. Suppose that for each player i the Nash payoffs are not all identical;

$${}^*\Pi_i = \frac{1}{L}\sum_{\ell=1}^{L} {}^\ell\bar{\Pi}_i > \min_{\ell=1,\ldots,L} {}^\ell\bar{\Pi}_i = {}^{\ell(i)}\bar{\Pi}_i , \quad i \in N . \text{ Then}$$

$$\lim_{T\to\infty} \text{perfect } NE(G^T) = NE(G^\infty) .$$

Proof. It suffices to construct a sequence of usable, perfect NE reward systems satisfying the conditions of Lemma 2. For any T take as ${}^{0}\sigma^{T}$ the alternation of each one-shot Nash equilibrium in cyclical fashion. If L does not divide T, then some Nash equilibria will appear one more time than others. Take as ${}^{i}\sigma^{T}$ the one-shot Nash equilibrium ${}^{\ell(i)}\bar{s}$ repeated T times. Then the reward $[\Pi_i({}^{0}\sigma^{T}) - \Pi_i({}^{i}\sigma^{T})]$ is approximately $T({}^{*}\Pi_i - {}^{\ell(i)}\bar{\Pi}_i) \to \infty$. Q.E.D.

6. Matrix Games

Until now we have assumed that strategy spaces are convex. We said that the usual matrix games, in which each player has access to a finite number of pure strategies, could be considered convex games provided that we supposed each player i has access to a randomizing device which is observable by all the other players, but only after i has moved. We call such devices personal-1 randomizing devices. If i flips a personal-1 coin to decide his move in period t, then j finds out the outcome of the flip at the beginning of period $t+1$. A more satisfying theory would not allow j to check the result of i's coin flip "after the fact," but this personal-1 observability hypothesis is common in the repeated games literature; it has been used for example by Aumann-Shapley (1976), Rubinstein (1979), Benoit-Krishna (1985), and Fudenberg-Maskin (1986). Fudenburg-Maskin show that for some purposes the personal-1 hypothesis can be dropped for infinitely repeated games.

Theorem A applies to finitely repeated matrix games even if we drop the personal-1 observability hypothesis. In particular the Benoit-Krishna theorem can be significantly generalized. However it is now no longer the case that an arbitrarily large threat can be created if only one player can commit himself for just the last period T. Our main theorem must accordingly be weakened.

If i's coin flip is observable simultaneously to all the other players and i, then we call it public randomization. Public randomization is less at odds with common sense than personal-1 randomization, and for matrix games we shall always suppose it is possible. In games without side payments public randomization can be used to serve an important function: i and j might agree to publicly flip a coin, making moves that are advantageous to i if heads, and advantageous to j if tails.

<u>Theorem 3</u>: Theorem A and Theorem 2 are valid for matrix games, even when players do not have access to personal-1 randomizing devices.

<u>Proof</u>: Specializing our various proofs to the case of matrix games, we see that personal-1 randomization is used in only one place: when the players are called upon to use (possibly mixed) minmax strategies in the punishment phase of some player j who has just cheated. If player i is not indifferent between all the pure strategies over which he is required to mix, then he will certainly cheat if cheating cannot be detected. Note, however, that Lemma 2 does not depend on any minmax strategies, hence it holds automatically for matrix games. We can modify the proof of Theorem A as follows.

The reusable reward system constructed in the proof derives from Lemma 2, and so does not depend on personal-1 randomization. Moreover the payoffs to the $N+1$ perfect NE_1's have convex hull with nonempty interior in R^N. By one public randomization, new equilibria can be created whose payoffs are arbitrary convex combinations of these N+1 payoffs. In particular, we can always substitute N+1 new equilibria with intended payoffs which are strictly interior to this convex set, but still far enough apart that their difference is much larger than K . With this in mind, suppose that player j cheats, and this cheating is followed by K periods of punishment by players i different from j . For the time being j has also earned himself the title of the last deviator. If there is no subsequent cheating by any player, then in the last phase we can always randomize over the original N+1 reusable equilibria so that all players i different from j are compensated so that the sum of the K period payoffs and the final phase compensation is exactly what their expected payoff would be if every player followed his prescribed (mixed) strategy for the K periods of the punishment phase and in the final phase the new j^{th} reusable equilibrium were followed. Player j is not compensated, receiving in the final phase exactly the new jth reusable equilibrium payoff.

Under this scheme it is clear that all players i different from j might as well follow the prescribed (mixed) punishment of j during the K period punsishment phase after j cheats. If that is how they are playing, then clearly j cannot gain during his own punishment phase since his strategy already prescribes playing as best he can against the minmax play of his opponents. The important point is that the final reusable reward phase now depends not only on who the last deviator was, but also on what happened during his last punishment phase. (If j cheats consecutively than in the final phase players i different from j are compensated only for their behavior during the last punishment phase of j).

Theorem A thus holds. Furthermore, since the creation of arbitrarily large reward systems in matrix games with multiple one shot Nash equilibria does not depend on personal-1 randomization, theorem 2 also holds for matrix games. Q.E.D.

Even when we allow for personal-1 randomizing devices, matrix games do not necessarily have the NE_1 folk theorem property because the payoffs are not necessarily differentiable at the vertices where pure strategies are played. When there is a mixed strategy Nash equilibrium, then (generically) the NE_1 finitely repeated folk theorem does hold because there are differentiable variations around equilibrium which permit the methods of Section 4 to be applied. For these games it is especially interesting to drop the personal-1 hypothesis in order to see how much the possibilities for cooperation deteriorate when monitoring is imperfect.

Consider for example the two person mxn matrix game $G = (A,B)$ given by

$$G = \begin{bmatrix} (1,0) & (0,1) \\ (0,.8) & (.9,0) \end{bmatrix} .$$

It is impossible to support the alternation of payoffs (0,1) and (1,0) in a perfect NE_1 . What is surprising is that such payoffs can be supported in perfect $NE_{\{1,2\}} = NE_N$, where $N = \{1,2\}$.

<u>Theorem 4</u>: Let $G = [A,B]$ be any $m \times n$ matrix game with a nondegenerate mixed strategy Nash equilibrium. Let G also be such that either there is a unique (i,j) with $A_{ij} = \max_{k,\ell} A_{k\ell}$ or else a unique (i',j') with $B_{i',j'} = \max_{k,\ell} B_{k\ell}$. Then $\lim_{T\to\infty}$ perfect $NE_{(1,2)}(G^T) = NE(G^\infty)$.

We defer the proof to the appendix.

BIBLIOGRAPHY

[1] Abreu, D., "Infinitely repeated games with discounting: A general theory," Harvard mimeo, 1984, Princeton thesis, 1982.

[2] Aumann, R., "Lecture notes on game theory," 1977.

[3] ________ and L. Shapley, "Long term competition: A game theoretic analysis," mimeo, Hebrew University, 1976.

[4] Benoit, J. and V. J. Krishna, "Finitely repeated games," Econometrica, 53 (1985), 905-922.

[5] Chou, C., Two Essays on Game Theory, Ph.D. Dissertation, Department of Economics, Yale University, 1984.

[6] Dubey, P. and A. Rogawski, "Inefficiency of Nash equilibria," CFDP No. 622, March 1982.

[7] Fudenberg, D. and D. Levine, "Subgame-perfect equilibria of finite- and infinite-horizon games," Journal of Economic Theory, 31 (1983), 251-268.

[8] ________ and E. Maskin, "The folk theorem in repeated games with discounting or with incomplete information," Econometrica, 54 (1986), 533-554.

[9] Geanakoplos, J. and D. Brown, "Understanding overlapping generations economies as lack of market clearing at infinity," mimeo, 1983.

[10] Kreps, D. M., P. Milgrom, J. Roberts, and R. Wilson, "Rational cooperation in the finitely repeated prisoners' dilemma," Journal of Economic Theory, 27 (1982), 245-252.

[11] Kreps, D. M. and R. Wilson, "Sequential equilibria," Econometrica, 50 (1982), 863-894.

[12] Luce, R. P. and H. Raiffa, Games and Decisions, Wiley, New York, 1957.

[13] Milgrom, P. and J. Roberts, "Limit pricing and entry under incomplete information: An equilibrium analysis," Econometrica, 50 (1982), 443-459.

[14] Radner, R., "Collusive behavior in noncooperative epsilon-equilibria of oligopolies with long but finite lives," Journal of Economic Theory, 22 (1980), 136-154.

[15] Rubinstein, A., "Equilibrium in supergames with the overtaking criterion," Journal of Economic Theory, 21 (1979), 1-9.

[16] __________, "Perfect equilibrium in a bargaining model," Econometrica, 50 (1982), 97-110.

[17] __________, "Strong perfect equilibrium in supergames," International Journal of Game Theory, 9 (1979), 1-12.

[18] Selten, R., "Re-examination of the perfectness concept of equilibrium points in extensive games," International Journal of Game Theory, 1 (1975), 25-55.

APPENDIX

Let us give a proof of Theorem 2. Let $G = [A,B]$ be a matrix game with a mixed strategy Nash equilibrium p^* , q^* . Let A and B be normalized so $\text{Max } A_{ij} = 1 = \text{Max } B_{k\ell}$, and $\text{Min } A_{ij} = 0 = \text{Min } B_{k\ell}$. Let $w_1^* = p^*Aq^*$, and let $w_2^* = p^*Bq^*$. We shall prove in Part I that for any $\delta > 0$ it is possible to construct for some T , a T-period perfect NE_N $*\mu^T$ with residue $\Pi_i^T(\mu^*) - (T-1)w_i^* > 1-\delta$, for $i = 1, 2$. In Part II we use this perfect NE_N to construct the usual reusable reward system.

Proof of Part I

Let us assume for now that $p^* \gg 0$, and $q^* \gg 0$.

Since (p^*, q^*) is nondegenerate, let us suppose that by perturbing player two's strategy we can improve player one, hence, there exists a $q = (q_1, q_2, \ldots, q_n)$ such that $\sum_{j=1}^{n} q_j = 0$ (hence, $p^*Bq = 0$) and $p^*Aq > 0$. For q sufficiently small, $\bar{q} = q^* + q$ is still a completely mixed strategy. Later we shall make q even smaller by multiplying by $1/k$. Consider the one-shot strategy pair $(p^*, \bar{q}) = (p, q^* + q)$. The one-shot payoffs are

$$\Pi_1 = p^*A(q^* + q) > p^*Aq^* = w_1^*$$

and

$$\Pi_2 = p^*B(q^* + q) = p^*Bq^* + p^*Bq = w_2^* + 0 = w_2^* .$$

Since player one still plays the NE strategy, p^* , player two is indifferent between any two strategies in the one-shot game. But p^* is not an optimal strategy for player one when player two plays $\bar{q}$.

Consider the repeated game G^T. Suppose player one plays p^* and player two plays $\tilde{q}$ all the time except the last move. If we use the payoff of the last move T to compensate player one such that he is indifferent between all strategies, and at the same time guarantee player two an expected payoff arbitrarily close to 1 on the last move, then we can derive a NE_N with a sufficiently large residue for both players. Against p^* player two has no incentive to cheat, and player one is steadily gaining his residue. One problem is that after paying two's residue in period T we have only a very limited resource, that is, 1-two's residue, to compensate player one for not cheating in the earlier periods. Another problem is that we can observe only the realization of one's randomization at each move, not how he randomizes.

Let

$$c_j = (\text{expected one-shot payoff for player 1 from strategy } j \text{ against } \tilde{q}) - w_1^*$$

$$(c_1, c_2, \ldots, c_m) = Aq ,$$

and let $Y^j(t)$ be the total number of times player one plays strategy j in rounds $1, 2, \ldots, t$. Define $X(0) = 0$ and

$$X(t) = \sum_{j=1}^{m} Y^j(t)c_j/k , \quad t = 1, 2, \ldots, T-1 .$$

$X(t) + tw_1^*$ is player one's expected payoff through period t, given his choices for $\tau \le t$, and given that player two randomizes according to $\tilde{q}_k = q^* + q/k$ at each move. Let $s > 0$ be given. For large enough k, $|X(t+1) - X(t)|$ will always be less than s. Let

$\bar{a} = \text{Max}\{a_{ij} | b_{ij} = 1\}$. It follows from continuity that for any c , with $\bar{a} < c < 1$, there are one-shot strategies $p(c)$ and $q(c)$ with $p(c)Aq(c) = c$; moreover, c near $\bar{a}$, $p(c)Bq(c)$ may be taken near 1.

Define a NE_N , (μ_1^*, μ_2^*) as follows

(1) At the t^{th} move $(1 \le t < T)$,

i) if $-s < X(\tau) < 1 - \bar{a} - 3s$ for all $\tau < t$, then (p^*, q_k) is played, where $q_k = q^* + q/k$;

ii) if $-s \ge X(\tau)$ or $X(\tau) \ge 1 - \bar{a} - 3a$ for some $\tau < t$, (p^*, q^*) is played.

Define

$\bar{t}$ = the first time $X(t)$ crosses either one of the boundaries, $-s$ or $1 - \bar{a} - 3s$,

T-1 if $X(t)$ never cross the boundaries.

(2) At T^{th} move, suppose $X(\bar{t}) = 1 - 2s - c$. Then $(p(c), q(c))$ with $p(c)Aq(c) = c$ as defined above is played. Observe that for large enough k , c satisfies $\bar{a} < c < 1$.

The above rules specify strategies for the two players.

Since player one always plays p^* except the last move, and since player two cannot affect the last move, player two is indifferent between all strategies. Hence, to test that (μ_1^*, μ_2^*) is a NE_N , we need only check that player one cannot gain by deviating during $t = 1, 2, \ldots, T-1$. But it is obvious that no matter what strategy player one adopts, if player two plays according to μ_2^* and player one does not vary at time T , then his expected payoff is:

$$(T-1)w_1^* + 1 - 2s \ .$$

Thus, (μ_1^*, μ_2^*) is certainly a NE_N . Hence, for k sufficiently large so that $|X(t+1) - X(t)| < s$,

$$r_1(\mu_1^*, \mu_2^*) = 1 - 2s \ .$$

Notice that when the players stick to the strategies μ_1^* and μ_2^* , the path of $X(t)$ is a random main process with main drift p^*Aq/k and variance proportional to $1/k^2$. Hence, for k sufficiently large, the probability of crossing the top boundary if player one always plays p^* approaches 1. At the top boundary, the period T payoff goes almost entirely to player two. Hence, given any $h > 0$, for k sufficiently large and s sufficiently small, player two's expected payoffs is

$$\Pi_2^N(\mu_1^*, \mu_2^*) \geq (T-1)w_2^* + (1-h)(1-s)$$

and the residue is

$$r_2(\mu_1^*, \mu_2^*) \geq (1-h)(1-s) \ .$$

Finally, we note that dropping the hypothesis that $p^* >> 0$ and $q^* >> 0$ changes almost nothing in the proof. Player two is still indifferent, against μ_1^* , to using any strategy μ_2 which only involves playing one shot strategies j with $q_j^* > 0$. Doing anything else can only make player two worse off. As for player one, we still define $(c_1, c_2, \ldots, c_m) = Aq$ and $X(t)$ as before. If p^* is not strictly positive, then player one's expected payoff through period t is bounded above by $X(t) + tv_1$; equality necessarily holds only when player one uses one shot strategies i with $p_i^* > 0$. Q.E.D.

Proof of Part II

We can now show how the perfect NE_N μ^* can be used to construct the reusable reward system posited in the continuity principle.

Since (p^*, q^*) is a nondegenerate mixed strategy, it is strictly Pareto dominated, and there are strategies σ^A and σ^B and integers $k \le K$ since that $x = \frac{k}{K}\Pi(\sigma^A) + \frac{K-k}{K}\Pi(\sigma^B) >> (w_1^*, w_2^*)$. Let $\Pi(\sigma^A) = (a,b) = ab$, and $\Pi(\sigma^B) = (a', b') = a'b'$. We shall support the average payoff x as a perfect NE_N for a game with arbitrarily large T . Once one has two perfect NE_N's whose total difference in payoffs can be made arbitrarily large, then the proof is finished as in our proof of Benoit-Krishna.

Of course the idea is to alternate the payoffs ab and a'b' over each cycle of K periods so that ab occurs k times and a'b' $K-k$ times. At the end comes NE_N μ^* . In case there is any deviation, both players get (w_1^*, w_2^*) till the very end. Thus by cheating a player loses $1-\delta$ from the final phase.

Lemma: Let (a,b) and (a', b') be given. Let $x = \frac{k}{K}(a,b) + \frac{K-k}{K}(a', b')$. Then there exists a function $f : \{1, \ldots, K\} \to \{(a,b), (a', b')\}$ such that $\#f^{-1}(a,b) = k$, and such that for any $1 \leq t \leq K$,

(1) $$\sum_{\tau=t}^{K} f_1(\tau) \geq (K-t+1)x_1 \quad \text{and}$$

(2) $$\sum_{\tau=t}^{K} f_2(\tau) \geq (K-tx_2 \ .$$

Suppose player one deviates at some time t . He gets at most $1 + (K-t)w_1^*$ until the end of the cycle, instead of $(K-t+1)x_1$. His net gain is $(1 - x_1) - (K-t)(x_1 - w_1)$. Similarly player two's maximum gain is $1 - (K-t)(x_2 - w_2)$. Since $x_i > w_i$, these numbers must be less than $1-\delta$ (for sufficiently small δ), except when $t = K$.

Thus the only time a player can gain as much as 1 is on the last period, if the payoff is supposed to be $a0$ (or $0b'$). We can always reduce $a0$ and $0b'$ with $1\bar{b}$ and $\bar{a}1$. If there is a unique ij with $A_{ij} = 1$, or a unique $k\ell$ with $B_{k\ell} = 1$, then either $1\bar{b}$ or $\bar{a}1$ has the property that neither player can change his strategy alone and gain 1.

Q.E.D.