International Journal of Game Theory (1992) 20: 211-226

On the Equilibrium Payoffs Set of Two Player Repeated Games with Imperfect Monitoring

By E. Lehrer[1]

Abstract: We show that any payoff, sustainable by a joint strategy of finitely repeated games, from which no player can deviate and gain by a non-detectable deviation, is a uniform equilibrium of the infinite repeated game. This provides a characterization of the uniform equilibrium payoffs in terms of the finitely repeated games.

1 Introduction

The study of Nash equilibrium payoffs of undiscounted infinitely repeated games by means of finitely repeated games has drawn a lot of attention. The folk theorem [A] characterizes the set of all upper Nash equilibrium payoffs (NE) in terms of the one-shot game. In some repeated games with imperfect monitoring (see [L1], [L2]), the set of all the equilibrium payoffs has been characterized by means of the stage game. However, these results could be reached only by relying on the particular information structure. A precise description, which uses only terms of the one-shot game, of NE in the most general case is still unknown.

We provide herewith a subset of NE in two player games and show that it characterizes the set of all the uniform equilibrium payoffs. This subset is defined in terms of all the finitely repeated games and not only in terms of the one-shot game.

We say that a strategy in a finitely repeated game, σ', is indistinguishable from σ if both induce the same probability distribution on the other player's signals, no matter what strategy the opponent plays. A payoff is *sustainable* if, first, it can be supported by a joint strategy (σ, τ) in some finitely repeated game and, second, any indistinguishable strategy from σ (resp. τ), say, σ' (resp. τ'), cannot increase player 1's (resp. players 2's) payoff. In other words, in case (σ, τ) is played, a player can profit by a deviation only if it changes the probability of his opponent's signals. Such a deviation is detectable if playing (σ, τ) is repeated many times.

Roughly speaking, the paper shows that the set of all uniform equilibrium payoffs coincides with the set of all the sustainable payoffs. Although the character-

---

[1] Ehud Lehrer, Raymond and Beverly Sackler, Faculty of Exact Sciences, School of Mathematical Sciences, Tel Aviv University, Ramat Aviv, Tel-Aviv, Israel.
The author acknowledges important suggestions of the two referees and the editor in charge of the *International Journal of Game Theory*. This research was partially supported by Grant No. SES-9022305 of the National Science Foundation, Economics.

0020-7276/92/3/211-226 $ 2.50 © 1992 Physica-Verlag, Heidelberg

ization is in terms of the all finitely repeated games, it provides us with a better understanding of the infinitely repeated games. In some particular cases, such as the standard-trivial case (where the information the player gets is either the pair of actions played or a null signal), it enables one to completely characterize NE and the uniform equilibrium payoffs in terms of the one-shot game (see (L3]).

2 The Model

a The One-Shot Game

The repeated game is an infinite repetition of a one-shot game that consists of:

(i) two finite sets of actions, Σ_1 and Σ_2. Denote $\Sigma = \Sigma_1 \times \Sigma_2$;
(ii) two payoff functions h_1, h_2; $h_i: \Sigma \to \mathbb{R}$. Denote $h = (h_1, h_2)$;
(iii) two information functions ℓ_1, ℓ_2 defined on Σ.

Without loss of generality, we may assume that $0 \le h_i \le 1$, $i = 1, 2$.

b The Repeated Game

(i) *Pure strategies.* Denote by L_i the set of all the possible signals of player i (L_i = the range of ℓ_i). A pure strategy of player i is a sequence $f = (f^1, f^2, \ldots)$, where $f^t: L_i^{t-1} \to \Sigma_i$. L_i^{t-1} is the cartesian product of L_i with itself $t-1$ times and it consists of all player i's possible histories of length $t-1$. If f and g are two pure strategies of player 1 and 2, respectively, then $x_i^t(f, g)$ will denote the payoff of player i at stage t if f and g are the strategies played.
(ii) A *mixed strategy* is a probability distribution over the set of all the pure strategies of the repeated game. Let σ_i be a mixed strategy of player i. Denote by $E_{\sigma_1,\sigma_2}(x_i^t)$ the expected payoff of player i at stage t where the expectation is taken with respect to the measure induced by (σ_1, σ_2).

c Nash Equilibria in the Repeated Game

Let σ_i be a mixed strategy of player i. $H_i^*(\sigma_1, \sigma_2)$ is defined as the limit of the means of player i's expected payoff. Precisely, $H_i^*(\sigma_1, \sigma_2) = \lim_T (1/T) \sum_{t=1}^{T} E_{\sigma_1,\sigma_2}(x_i^t)$ if the limit exists. We will say that (σ_1, σ_2) is an upper Nash equilibrium if $H_i^*(\sigma_1, \sigma_2)$ is defined and if for any other pair of mixed strategies $(\bar{\sigma}_1, \bar{\sigma}_2)$,

$$\text{(i)} \quad H_1^*(\sigma_1, \sigma_2) \geq \limsup (1/T) \sum_{t=1}^{T} E_{\bar{\sigma}_1, \sigma_2}(x_1^t), \quad \text{and}$$

$$\text{(ii)} \quad H_2^*(\sigma_1, \sigma_2) \geq \limsup (1/T) \sum_{t=1}^{T} E_{\sigma_1, \bar{\sigma}_2}(x_2^t).$$

Denote by UEP the set of all the payoffs $(H_1^*(\sigma_1, \sigma_2), H_2^*(\sigma_1, \sigma_2))$, where (σ_1, σ_2) is an upper Nash equilibrium.

d Uniform Equilibrium

$\sigma = (\sigma_1, \sigma_2)$ is a *uniform* equilibrium if $H^*(\sigma)$ is defined and if for any $\varepsilon > 0$ there is N s.t. σ induces an ε-Nash equilibrium in the n-fold repeated game for any $n > N$. UNIF denotes the set of all the uniform equilibrium payoffs. For more extensive study of uniform equilibrium, see [S].

e Banach Equilibrium

Let L be a Banach limit. For a joint mixed strategy $\sigma = (\sigma_1, \sigma_2)$ define $H^{*L}(\sigma) = L\left\{1/T \sum_{t=1}^{T} E_{\sigma_1, \sigma_2}(x_1^t, x_2^t)\right\}_T$. σ is an L-equilibrium if

$$H_1^{*L}(\sigma) \geq L\left\{1/T \sum_{t=1}^{T} E_{\bar{\sigma}_1, \sigma_2}(x_1^t)\right\}_T$$

for all strategies $\bar{\sigma}_1$, of player 1, and a similar inequality for player 2.

Denote by BEP_L the set of all $H^{*L}(\sigma)$, where σ is an L-equilibrium.

The problem of characterizing UEP and BEP_L is still open. We are able to identify a set of payoffs which is included in all of these equilibrium payoffs sets. However, UNIF is fully characterized here.

Remark 1: It is clear that UNIF is contained in UEP and in BEP_L for every Banach limit L.

3 The Main Theorem

Some notations are needed in order to state our main result.

a Extending the Domain of ℓ_i

The information functions ℓ_i where defined on Σ and ranged to L_i. We can extend in a natural way the domain of ℓ_i to be the set of all probability distributions on Σ, denoted by $\Delta(\Sigma)$, so as to attain values in $\Delta(L_i)$, the set of all probability distributions on L_i.

b Indistinguishable Actions

Let p and p' be two mixed actions of player 1. We say that p and p' are *indistinguishable* (denoted $p \sim p'$) if $\ell_2(p, q) = \ell_2(p', q)$ for all mixed actions q of player 2. The same relation is defined also for player 2. Sometimes we say that p is indistinguishable from p'.

c More Informative Actions

We define a partial order $\succ$ on the set of mixed actions as follows. Let p and p' be two mixed actions of player 1. p and p' are thought of as probability distributions over Σ_1. We say that p is *more informative* than p' (denoted $p \succ p'$) if

$$p\{a \in \Sigma_1 \mid \ell_1(a, b) \neq \ell_1(a, b')\} \geq p'\{a \in \Sigma_1 \mid \ell_1(a, b) \neq \ell_1(a, b')\}$$

for every two pure actions, b and b' of player 2.

In a similar way, the partial order $\succ$ is defined on mixed actions of player 2.

The interpretation of the previous definition is the following. p is more informative than p' if the probability of distinguishing between two pure actions of player 2 is greater by playing p than by playing p'.

d The Finitely Repeated Game, G_n

The n-fold repeated game G_n is defined by the sets of strategies Σ_1^n and Σ_2^n, the average payoff functions h_1^n and h_2^n, and the information functions ℓ_1^n and ℓ_2^n. We can treat G_n as we formerly treated the one-shot game and define the relations $\sim$ and $\succ$ on Σ_i^n, $i=1,2$. Using these relations, we will define two subsets of $\Delta(\Sigma_1^n)\times\Delta(\Sigma_2^n)$, the set of all the pairs of n-fold repeated game mixed strategies.

e The Set D_ε^n

Let $\varepsilon>0$. Define

$$D_\varepsilon^n=\left\{(\sigma,\tau)\in\Delta(\Sigma_1^n)\times\Delta(\Sigma_2^n)\,\middle|\,\begin{array}{l}1.\ h_1^n(\sigma,\tau)\geq h_1^n(\sigma',\tau)-\varepsilon\text{ for all }\sigma'\sim\sigma\\ 2.\ h_2^n(\sigma,\tau)\geq h_2^n(\sigma,\tau')-\varepsilon\text{ for all }\tau'\sim\tau\end{array}\right\}.$$

In words, (σ,τ) a pair of mixed strategies is an element of D_ε^n if σ is an ε-best response to τ, among all $\sigma'\sim\sigma$, and a similar condition for player 2.

Remark 2: From continuity and compactness we deduce that if $(\sigma,\tau)\in D_\varepsilon^n$ and if $\|\ell_1^n(a,\tau)-\ell_1^n(a,\tau')\|_\infty<\eta$ for every pure strategy a of player 1, then $h_1^n(\sigma,\tau')\leq h_1^n(\sigma,\tau)+\varepsilon+\delta(\eta,n)$, where $\delta(\eta,n)$ goes to zero as $\eta\to 0$ for every fixed n.

f The Set C_ε^n

The set C_ε^n is defined in a similar way to D_ε^n requiring an additional $\sigma'\succ\sigma$ and $\tau'\succ\tau$. In other words, (σ,τ) is in C_ε^n if σ is an ε-best response versus τ among all σ' that are indistinguishable from σ and at the same time more informative than σ, and a similar condition for player 2. Clearly, $D_\varepsilon^n\subseteq C_\varepsilon^n$.

Remark 3: Define $N_\varepsilon^n=\{(\sigma,\tau)\,|\,(\sigma,\tau)$ is a ε-Nash equilibrium of $G_n\}$.

Clearly, $N_\varepsilon^n\subseteq D_\varepsilon^n$. Moreover, $\text{UNIF}\subseteq\bigcap_{\varepsilon>0} c\ell\bigcup_{n=1}^{\infty} h^n(N_\varepsilon^n)$ where $c\ell$ denotes the closure operator. Therefore, $\text{UNIF}\subseteq\bigcap_{\varepsilon>0} c\ell\bigcup_{n=1}^{\infty} h^n(D_\varepsilon^n)\bigcap\text{IR}$, where IR is the set of individually rational payoffs. The inverse inclusion is a part of the main theorem's contents.

g UNIF is Characterized in Terms of All the G_n

It is well-known that in some zero-sum repeated games the sets of Nash equilibrium payoffs in G_n (the value) tend to the set of the Nash equilibrium payoffs in the infinite repeated game. For instance, this happens in repeated games with incomplete information, lack of information on one side and in stochastic games. This is no longer true in non zero-sum games.

In the present case, we will describe UNIF in general repeated games with imperfect monitoring using the sets D_ε^n, which correspond to the finitely repeated games. Namely, in terms of ε-best response, among equivalent strategies (as opposed to using the term "best response among *all* the mixed strategies" as in Nash equilibrium).

A game is with *non-completely trivial information* if each of the players has two distinguishable actions. Our goal in this paper is to prove the following theorem.

The Main Theorem: In repeated games with imperfect monitoring and with non-completely trivial information:

$$\mathrm{UNIF} = \bigcap_{e>0} c\ell \bigcup_{n=1}^{\infty} h^n(D_\varepsilon^n) \bigcap \mathrm{IR} = \bigcap_{\varepsilon>0} c\ell \bigcup_{n=1}^{\infty} h^n(C_\varepsilon^n) \bigcap \mathrm{IR}.$$

Conjecture: $\bigcap c\ell \bigcup h^n(C_\varepsilon^n) \bigcap \mathrm{IR}$ covers all UEP, namely, $\mathrm{UEP} = \bigcap c\ell \bigcup h^n(D_\varepsilon^n) \bigcap \mathrm{IR} = \bigcap c\ell \bigcup h^n(C_\varepsilon^n) \bigcap \mathrm{IR}$. Thus, UEP=UNIF. In words, the set of upper equilibrium payoffs *coincides* with the set of uniform payoffs.

4 Some Properties of D_ε^n and C_ε^n

(i) If $\varepsilon < \varepsilon'$ then $D_\varepsilon^n \subseteq D_{\varepsilon'}^n$.
(ii) For every n and $k\,h^n(D_\varepsilon^n) \subseteq h^{nk}(D_\varepsilon^{nk})$ because if $(\sigma, \tau) \in D_\varepsilon^n$ then the k repetitions of (σ, τ) is a pair in D_ε^{nk}. And for a similar reason:
(iii) $n\,h^n(D_\varepsilon^n) + k\,h^k(D_\varepsilon^k) \subseteq (n+k)\,h^{n+k}(D_\varepsilon^{n+k})$.

The properties (ii) and (iii) hold for C_ε^n as well.

(iv) $c\ell \bigcup_{n=1}^{\infty} h^n(D_\varepsilon^n)$ is a convex set. It is sufficient to show that $\operatorname{conv} \bigcup_n h^n(D_\varepsilon^n) \subseteq c\ell \bigcup_n h^n(D_\varepsilon^n)$. This result is implied by property (iii) and the following claim.

Claim: Let $\{F_n\}$ be a sequence of sets in a norm space. Assume that for every $k, n \in \mathbb{N}$ $nF_n + kF_k \subseteq (n+k)F_{n+k}$, then $\operatorname{conv} \bigcup F_n \subseteq c\ell \bigcup F_n$.

Proof: Let $\alpha \in \operatorname{conv} \bigcup F_n$. So α is a convex combination $\alpha = \sum_{i=1}^{\ell} \gamma_i \alpha_{n_i}$, where $\gamma_i \geq 0$, $\sum \gamma_i = 1$ and $\alpha_{n_i} \in F_{n_i}$. By induction we deduce from the assumption of the claim that for any convex and rational combination $(r_i/q)_{i=1}^{\ell}$ (i.e., $r_i, q \in \mathbb{N}$, and $\sum_{i=1}^{\ell} r_i = q$) one gets $\sum_{i=1}^{\ell} r_i n_{-i} n_i F_{n_i} \subseteq \bar{q} F_{\bar{q}}$, where $n_{-i} = \prod_{j \neq i} n_j$, $n = \prod n_i$ and $\bar{q} = n \sum r_i$. Therefore, any rational convex combination of elements from $\bigcup_n F_n$ is included in $\bigcup_n F_n$. In order to complete the proof of the claim take for every $\varepsilon > 0$ a rational convex combination which satisfies $|r_i/q - \gamma_i| < \varepsilon/\ell$ for every $1 \leq i \leq \ell$. So,

$$\|\sum (r_i/q)\alpha_{n_i} - \sum \gamma_i \alpha_{n_i}\| = \|\sum (r_i/q - \gamma_i)\alpha_{n_i}\| \leq \varepsilon \sum \|\alpha_{n_i}\| \underset{\varepsilon \to 0}{\longrightarrow} 0.$$

Since each of the rational combinations $\sum (r_i/q)\alpha_{n_i}$ is in $\bigcup_n F_n$, $\sum \gamma_i \alpha_{n_i} \in c\ell \bigcup_n F_n$. //

(v) $\bigcap_\varepsilon c\ell \bigcup_n h^n(D_\varepsilon^n)$ is a closed and convex set as an intersection of closed and convex sets.

5 The Proof of the Main Theorem

We will divide the proof into two propositions. The first one states that the infinite intersection defined by the sets C_ε^n is not bigger than the one defined by the smaller sets D_ε^n. The second proposition provides a formula to the set of all the uniform payoffs.

Proposition 1:

$$\bigcap_{\varepsilon > 0} c\ell \bigcup_{n=1}^{\infty} h^n(D_\varepsilon^n) = \bigcap_{\varepsilon > 0} c\ell \bigcup_{n=1}^{\infty} h^n(C_\varepsilon^n)$$

provided that the game is with non-completely trivial information.

In fact, we have to show only that $\bigcap_\varepsilon c\ell \bigcup_n h^n(D_\varepsilon^n) \supseteq \bigcap_\varepsilon c\ell \bigcup_n h^n(C_\varepsilon^n)$, because the inverse direction is obvious.

Proposition 2: Under the conditions of Proposition 1:

$$\text{UNIF} = \bigcap_{\varepsilon > 0} c\ell \bigcup_{n=1}^{\infty} h^n(D_\varepsilon^n) \bigcap \text{IR}.$$

In view of Remark 3 we have to show that UNIF includes $\bigcap_{\varepsilon>0} c\ell \bigcup_{n=1}^{\infty} h^n(D_\varepsilon^n) \bigcap \text{IR}$.

Proof of Proposition 1 (using an idea of S. Sorin [S]): The idea of the proof is to use $(\sigma, \tau) \in C_\varepsilon^n$ in the construction of another pair of strategies, (σ', τ'), in a much longer game, say, G_r. The pair (σ', τ') has two properties: (i) its payoff is close to the payoff of (σ, τ) and (ii) (σ', τ') is included in $D_{3\varepsilon}^r$. Denote by k^n the number of pure strategies in G_n.

For an arbitrary integer t define (σ', τ') as follows. The players will play t time according to a perturbation of (σ, τ) (at the tn first stages). Then, at the coming $[\ell g_2 t]+1$ stages player 2 will choose randomly one of the t repetitions and report the choice to player 1. Then, at the following $[\ell g_2 k^n]+1$ player 1 will report on his signal that he had got at the repetition chosen by player 2. Afterwards the players exchange roles. First player 1 picks randomly a repetition (or a number from 1 to t) transmits it to player 2 and then player 2 reports on his signal at the repetition that was chosen by player 1.

Denote $r(t, n) = tn + 2[\ell g_2 k^n] + 2[\ell g_2 t] + 4$. We will prove that $(\sigma', \tau') \in D_{3\varepsilon}^{r(t,n)}$, provided that t is sufficiently large.

How to choose a random stage and to report it. Since the information is not completely trivial there are $a, c_1, c_2 \in \Sigma_1$ and $b_1, b_2, d \in \Sigma_2$ s.t.

$$\ell_1(a, b_1) \neq \ell_1(a, b_2) \quad \text{and} \quad \ell_2(c_1, d) \neq \ell_2(c_2, d).$$

Player 2 picks randomly a number from $\{1, \ldots, t\}$ and simultaneously reports it to player 1 by the following procedure. Player 1 will play a (regardless of the history) and player 2 will play with probability 1/2 each one of the actions b_1 and b_2. If this procedure is repeated $[\ell g_2 t]+1$ times, player 1 ends up with a random string of length $[\ell g_2 t]+1$ consisting of two symbols ($\ell_1(a, b_1)$ and $\ell_1(a, b_2)$). These strings encode the numbers $1, \ldots, t$. So, getting one of these strings, player 1 is informed about a number between 1 and t which is interpreted as the repetition on which he has to report later.

How to report the signals. The set of player 1's signals is finite and it contains less than k^n characters. Thus, by answering on less than $\ell g_2 k^n + 1$ "Yes-No" questions player 1 can report on his signal to player 2. (Player 2 will play d in order to receive the answers and player 1 will play c_1 for "Yes" and c_2 for "No".)

Denote by $\sigma(\delta)$ (resp. $\tau(\delta)$) a strategy of player 1 (resp. player 2) that assigns probability $1-\delta$ to σ and a positive probability to each one of his pure strategies in G_n. Notice that from continuity arguments one can deduce that for every $\eta > 0$ there is $\delta > 0$ s.t. if $(\sigma, \tau) \in C_\eta^n$ then $(\sigma(\delta), \tau(\delta)) \in C_{2\eta}^n$. Let δ be the one corresponding to the ε under consideration.

The strategies (σ', τ'). The strategies are consisting of five phases. At the first one (tn stages) the players play according to $(\sigma(\delta), \tau(\delta))$. At the second phase ($[\ell g_2 t]+1$

stages) player 2 chooses a number t' from $\{1, \ldots, t\}$ and reports it. At the third phase ($[\ell g_2 k^n]+1$) – a reporting phase – player 1 reports to player 2 about the signal he had got at the repetition t' that was chosen at the second phase. At the fourth phase player 1 chooses a number from $\{1, \ldots, t\}$ and at the last phase, again a reporting phase, player 2 reports on his signal.

According to the construction of (σ', τ') there is a positive probability for any repetition to be chosen (and therefore to be the repetition on which the player will have to report). Thus any other strategy of a player (at the first tn stages) will lead to change in the distribution of the opponent's signals at the reporting phases. To make it clear let us concentrate on player 1. One of the two (one for each player) possibilities that (σ', τ') is not in $D_{3\varepsilon}^{r(t,n)}$ is that there is a strategy $\bar{\sigma}$ which retains the distribution of player 2's signals (i.e., $\bar{\sigma} \sim \sigma'$) and increases player 1's payoffs by more than 3ε.

We think of the first phase as t repetitions of G_n. In playing (σ', τ') each player ignores his memory after each repetition. Since player 2 disregards his memory, the strategy $\bar{\sigma}$ of player 1 has a similar strategy, say, $\hat{\sigma}$ (similar in the sense that both induce the same probability on the set of histories of length n in each one of the repetitions) and that in playing according to $\hat{\sigma}$ player 1 ignores his memory after every repetition. Thus, if $\bar{\sigma}$ is indistinguishable from σ', $\hat{\sigma}$ is also indistinguishable from σ'. Since t is large enough it means that there is t', $1 \le t' \le t$, s.t. the strategy $\sigma_{t'}$, the one induced by $\hat{\sigma}$ at the t'-th repetition, increases the payoff at the t'-th repetition (of G_n) by at least 2ε. However, $(\sigma(\delta), \tau(\delta)) \in C_{2\varepsilon}^n$ and therefore $\sigma_{t'} \not> \sigma(\delta)$. In other words, there is a repetition $t' \in \{1, \ldots, t\}$ in which player 1, by playing $\sigma_{t'}$, (rather than $\sigma(\delta)$), either (i) alters the distribution of player 2's signals in the t'-th repetition or (ii) loses with a positive probability a possibility to collect information.

The first case, (i), immediately implies that $\hat{\sigma} \not\sim \sigma'$ and therefore $\bar{\sigma} \not\sim \sigma'$. If, however, (ii) is the case, then according to (σ', τ') there is a positive probability that player 1 will have to report (at the third phase) on his signal at the t'-th repetition. Hence, by playing $\sigma_{t'}$, player 1 changes the distribution of player 2's signals at the third stage, and therefore $\hat{\sigma} \not\sim \sigma'$. Thus, $\bar{\sigma} \not\sim \sigma'$. Both cases exclude the possibility that by playing a strategy indistinguishable from σ', player 1 can gain on average by at least 2ε at the first phase.

Notice that (σ', τ') do not describe best responses at the last $2[\ell g_2 t] + 2[\ell g_2 k^n] + 4$ stages. However, by playing another strategy, indistinguishable from his prescribed strategy, a player can gain at most $2[\ell g_2 t] + 2[\ell g_2 k^n] + 4$ times the maximal payoff. When this number is divided by $r(t, n)$ it is smaller than ε, whenever t is sufficiently large. We conclude that $(\sigma', \tau') \in D_{3\varepsilon}^{r(t,n)}$, and that if $\delta > 0$ small enough, then $\| h^n(\sigma, \tau) - h^{r(t,n)}(\sigma', \tau') \| < \varepsilon$. Thus, $\bigcup_n h^n(C_\varepsilon^n)$ is contained in the ε-neighborhood of $\bigcup_n h^{r(t,n)}(D_{3\varepsilon}^{r(t,n)}) \subseteq \bigcup_n h^n(D_{3\varepsilon}^n)$. Hence, $\bigcap_\varepsilon c\ell \bigcup_n h^n(C_\varepsilon^n) \subseteq \bigcap_\varepsilon c\ell \bigcup_n h^n(D_\varepsilon^n)$. Q.E.D.

Proof of Proposition 2: We will show that any payoff (γ_1, γ_2) in $\bigcap_\varepsilon c\ell \bigcup_n h^n(D_\varepsilon^n) \bigcap$ IR can be sustained by an equilibrium (f, g). Since (γ_1, γ_2) is in the closure of

$\bigcup_n h^n(D^n_\varepsilon)$ for every $\varepsilon > 0$, we can find a sequence (α^i, β^i) which converges to (γ_1, γ_2) and $(\alpha^i, \beta^i) \in h^{n^i}(D^{n^i}_{1/i})$. Let (σ^i, τ^i) be a pair in $D^{n^i}_{1/i}$ satisfying $h^{n^i}(\sigma^i, \tau^i) = (\alpha^i, \beta^i)$.

The strategies (f, g) will be defined along the following lines. We will divide the set of stages $\mathbb{N}$, into blocks $B^1, B^2, \ldots$ and each of these blocks is divided into sub-blocks of increasing lengths, say $B^i_1, B^i_2, \ldots, B^i_{u_i}$.

At the block B^i a lot of repetition of a perturbation of (σ^i, τ^i) will be played, and after each sub-block the players make statistical tests (based on the relative frequence of the signals) to check whether or not their opponent has deviated at that sub-block. In case where no deviation has been found the game proceeds to the next sub-block, in which the players play repeatedly perturbations of (σ^i, τ^i) as was described above. In the other case, where an alleged deviation has occurred, a punishment will take place for a long time: up to the end of the block. Then, at the new block the game proceeds as if the players forgot the history, by playing repetitions of the perturbed $(\sigma^{i+1}, \tau^{i+1})$, and so on.

A few points should be made clear before defining precisely the strategies.

(i) If (σ^i, τ^i) is repeated many times, the empirical distribution of any signal s (of the game G_{n^i}) should be, with high probability, close to its probability according to (σ^i, τ^i). If not, the player who gets that signal arrives at the conclusion that his opponent has deviated and therefore he should punish him. However, there is a positive probability for the punisher to arrive at the wrong conclusion, and thus to punish the other player while the latter did not deviate at all. For this reason any punishment cannot take place forever. It should be a temporary punishment (till the end of the block) after which the player will return to the master plan.

(ii) Since both players check after deviations, it might occur that for instance, player 1 punishes player 2, who might interpret the punishment as a deviation and thus to punish player 1 who himself can interpret the punishment to the punishment as a deviation and so on ad infinitum. To avoid this possibility, the players ignore their memory at the beginning of each block and the game proceeds as if it is a new game.

(iii) The payoff (γ_1, γ_2) is also in IR, and therefore a player can punish his opponent and push his payoff down to the individually rational level.

In order to define (f, g), let
$\{\varepsilon^i\}$ be a sequence of small positive numbers to be determined later;
k^i be the number of pair of pure strategies in the game G_{n^i};
k^i_j be the number of player j's pure strategies in the game G_{n^i}.

$$s^i \quad = n^i k^i / (\varepsilon^i)^{10}$$

$$u_i \quad = s^{i+1}$$

$$\# B^i_\nu = s^i / (\varepsilon^i)^{\nu+15}, \quad \nu = 1, \ldots, u_i.$$

The *master plan* of f, the strategy of player 1 (a similar description for g) is to play at B^i_ν the following mixed strategy of G_{n^i}: with probability $1 - k^i_j \varepsilon^i$ play σ^i, and with probability ε^i play each one of the k^i_j pure strategies. This perturbation of σ^i will be denoted by $\sigma^i(\varepsilon^i)$, and the similar perturbation of τ^i will be denoted by $\tau^i(\varepsilon^i)$.

The *punishment plan* of f is the following. For each signal $s \in L_l^{n^i}$ of the game G_{n^i}, and for each pure strategy σ of player 1 (in G_{n^i}) denote by $O_v^i(\sigma, s)$ the number of times the signal s was observed at B_v^i, while player 1 was playing σ. Thus, the empirical distribution of s (when player 1 plays σ) is the number $O_v^i(\sigma, s)$ divided by the number of times that σ was actually played, say $N_v^i(\sigma)$.

Player 1 will punish player 2 after the sub-block B_v^i if

$$|O_v^i(\sigma, s)/N_v^i(\sigma) - E^i(\sigma, s)| > (\varepsilon^i) \tag{1}$$

for some σ and s, where $E^i(\sigma, s)$ is the probability of the signal s when player 1 plays σ (and player 2 plays $\tau^i(\varepsilon^i)$). Notice that $E^i(\sigma, s)$ does not depend on v, because the same pair; $(\sigma^i(\varepsilon^i), \tau^i(\varepsilon^i))$ is repeated in all B_v^i $(v = 1, \ldots, u_i)$. Player 1 punishes only till the end of the block B^i. After the punishment terminates the players return to play according to the master plan.

We have to show now that $H^*(f, g) = (\gamma_1, \gamma_2)$ and that there is not any profitable deviation. These will be proved at the following two lemmas.

Lemma 1: $H^*(f, g) = (\gamma_1, \gamma_2)$.

Proof: Recall that $x_j^t(f, g)$ is denoting the payoff for player j at stage t. We will show that $(1/T) \sum_{t=1}^{T} x_j^t(f, g)$ tends to γ_j a.s. By the bounded convergence theorem this implies that $\lim_T (1/T) \sum_{t=1}^{T} E(x_j^t(f, g)) = \gamma_j$, $j = 1, 2$.

We will prove first that with probability 1 there is an integer I s.t. if $i \geq I$ then the average of $x_j^t(f, g)$ over B_v^i is close to γ_j up to $2k^i\varepsilon^i$ for every $1 \leq v \leq u_i$.

The pair $(\sigma^i(\varepsilon^i), \tau^i(\varepsilon^i))$ (recall, the perturbation of (σ^i, τ^i)) is repeated $\# B_v^i/n^i$ times in B_v^i, provided that both players follow the master plan. From the definitions of B_v^i and s^i one gets,

$$\begin{aligned} \# B_v^i/n^i &\geq s^i/(\varepsilon^i)^{v+15}(n^i) \\ &= n^i k^i/(\varepsilon^i)^{25+v}(n^i) = k^i/(\varepsilon^i)^{25+v}. \end{aligned} \tag{2}$$

Thus, $(\sigma^i(\varepsilon^i), \tau^i(\varepsilon^i))$ is repeated at least $k^i/(\varepsilon^i)^{25+v}$ times in B_v^i. By Chebyshev inequality the probability that the relative frequency of any pair, (σ, τ), of pure strategies of G_{n^i} to be far from its probability by more than ε^i, is by (2) less than $(\varepsilon^i)^{23+v}/k^i$. Thus, the probability of the event, say A, that *all* the pairs of pure strategies of G_{n^i} will have empirical distribution within ε^i of their probability is bounded from below by

$$1 - k^i(\varepsilon^i)^{23+v}/k^i \geq 1 - (\varepsilon^i)^{23+v} \tag{3}$$

(recall, provided that the master plan is played).

However, there is a positive probability to the event, say, A', that the players will *not* play according to the master plan. Our task now is to estimate this probability, namely, to estimate prob(A'). Suppose that in $B^i_{\bar{\nu}}$ the players play according to the master plan. Let us first convince ourselves that the probability that player 1 will find a deviation in $B^i_{\bar{\nu}}$, i.e., that (1) holds for some pure strategy σ and a signal s is bounded by

$$k^i_1/(\varepsilon^i/2)^2(\#B^i_{\bar{\nu}}/n^i)+k^i_1 k^i_2/(\varepsilon^i)^2(\varepsilon^i/2)(\#B^i_{\bar{\nu}}/n^i). \tag{4}$$

Recall that each pure strategy σ is played (according to $\sigma^i(\varepsilon^i)$) at least with probability ε^i. By Chebyshev inequality, each σ is played in $B^i_{\bar{\nu}}$ at least $(\varepsilon^i/2)(\#B^i_{\bar{\nu}}/n^i)$ times with probability of at least $1-1/(\varepsilon^i/2)^2(\#B^i_{\bar{\nu}}/n^i)$. Given that σ is played at least $(\varepsilon^i/2)(\#B^i_{\bar{\nu}}/n^i)$ times, the probability for a particular signal s to satisfy (1) alongside with σ is bounded by $1/(\varepsilon^i)^2(\varepsilon^i/2)(\#B^i_{\bar{\nu}}/n^i)$. Thus, the probability of finding *some* signal that satisfies (1) is less than $k^i_2/(\varepsilon^i)^2(\varepsilon^i/2)(\#B^i_{\bar{\nu}}/n^i)$. Therefore, the probability that (1) holds true for a particular σ and some s (now taking into account also the possibility that the relative frequency of σ is *less* than $(\varepsilon^i/2)(\#B^i_{\bar{\nu}}/n^i)$) is less than $1/(\varepsilon^i/2)^2(\#B^i_{\bar{\nu}}/n^i)+k^i_2/(\varepsilon^i)^2(\varepsilon^i/2)(\#B^i_{\bar{\nu}}/n^i)$. In order to get the probability of (1) for any σ and s one should multiply it by k^i_1 and get (4).

For getting an upper bound to the probability that the punishment plan is being played at B^i_ν we should sum up (4) over all $\bar{\nu}$ $(1\le\bar{\nu}<\nu)$ and add a similar number for player 2. The outcome is $16(\varepsilon^i)^2$ by the following calculation. The summation of (4) over $\bar{\nu}$ is:

$$\sum_{\bar{\nu}=1}^{\nu}[k^i_1/(\varepsilon^i/2)^2(\#B^i_{\bar{\nu}}/n^i)+k^i_1 k^i_2/(\varepsilon^i)^2(\varepsilon^i/2)(\#B^i_{\bar{\nu}}/n^i)]$$

$$\le[(4k^i)/(\varepsilon^i)^3]\sum_{\bar{\nu}=1}^{\nu} n^i/\#B^i_{\bar{\nu}}=4k^i\sum_{\bar{\nu}=1}^{\nu} n^i(\varepsilon^i)^{\bar{\nu}+12}/s^i\le 8(\varepsilon^i)^2. \tag{5}$$

(For two players we get $16(\varepsilon^i)^2$.) In other words, prob$(A')\ge 1-16(\varepsilon^i)^2$.

Now we can combine (3) and (5) and determine that the probability that there will be a pair of pure strategies (σ,τ) with an empirical distribution which is far by more than ε^i from its probability, is less than $(\varepsilon^i)^{23+\nu}+16(\varepsilon^i)^2$. In other words,

$$\text{prob}(A)\le(\varepsilon^i)^{23+\nu}+16(\varepsilon^i)^2. \tag{6}$$

Given the event A the average of $x^t_j(f,g)$ over B^i_ν is far from γ_j by only $2k^i\varepsilon^i$ (one $k^i\varepsilon^i$ for the perturbation and one for the distance between the empirical distribution and the real probability). In other words,

$$\text{prob}(\{|1/\#B^i_\nu\sum_{t\in B^i_\nu}x^t_j(f,g)-\gamma_j|<2k^i\varepsilon^i\}\,|A)=1.$$

If $\varepsilon^i\le 1/i$ the summation of (6) over all i and $1\le\nu\le u^i$ is finite and we conclude, by Borel-Cantelli lemma, that with probability 1 there is an I s.t. if $i\ge I$ then the average of $x^t_j(f,g)$ over $t\in B^i_\nu$ for all $1\le\nu\le u^i$ is close to γ_j up to $2k^i\varepsilon^i$.

Since the length of B^i_ν compared to all its predecessors tends to zero as i goes to infinity we conclude that if $k^i\varepsilon^i \to 0$ the average of $x_j^t(f, g)$ over all t's tends to γ_j. This concludes the proof of Lemma 1. Q.E.D.

Remark 4: The proof of the previous lemma actually shows more than desired. It shows that the average payoffs converge *almost surely* to (γ_1, γ_2).

Lemma 2: (f, g) is a uniform equilibrium.

Proof: Suppose that instead of playing g player 2 plays $\bar{g}$. (A similar argument goes through for any $\bar{f}$.)

(i) We will estimate first the probability for player 2 to gain by more than (see Remark 2 for the definition of $\delta(\cdot, \cdot)$) $1/i + \delta(3\varepsilon^i, n^i)$ at B_ν^i without being detected. Recall that in B_ν^i the perturbation $(\sigma^i(\varepsilon^i), \tau^i(\varepsilon^i))$ is played $b = \# B_\nu^i / n^i$ times.

For every pure strategy, σ, of player 1 in G_{n^i} define the random variable R_σ^t to be 1 if player 1 played σ at the t-th repetition and 0 otherwise. Similarly, let Y_τ^t be 1 if player 2 played the pure strategy τ at the t-th repetition and 0 otherwise. Define $B_\sigma = \{t \mid R_\sigma^t = 1\}$ and let $P^i(\sigma)$ be the probability of σ according to $\sigma^i(\varepsilon^i)$.

In the following computation we make use of Lemma 5.5 of [L1]. (We will elaborate on this point in Appendix 1.)

The average payoff of player 2 in B_ν^i

$$= 1/b \sum_{t=1}^{b} \sum_{\sigma} \sum_{\tau} R_\sigma^t Y_\tau^t h_2^{n^i}(\sigma, \tau)$$

(with probability of at least $d_1 = 1 - 1/(\varepsilon^i)^2 b$).

$$\leq 1/b \sum_{\sigma} \sum_{\tau} \sum_{t=1}^{b} (P^i(\sigma) Y_\tau^t + \varepsilon^i) h_2^{n^i}(\sigma, \tau)$$

(defne $q(\tau) = 1/b \sum_{t=1}^{b} Y_\tau^t$)

$$= \sum_{\sigma} \sum_{\tau} (P^i(\sigma) q(\tau) + \varepsilon^i) h_2^{n^i}(\sigma, \tau)$$

(q is the strategy of player 2 defined by $q(\tau)$)

$$= h_2^{n^i}(\sigma^i, q) + k^i \varepsilon^i.$$

Using once again Lemma 5.5 of [L1] (see Appendix 2) one obtains that, with probability of at least $d_2 = 1 - ((\varepsilon^i)^4/b + (\varepsilon^i)^2/b) k_2^i$,

$$\| \ell_1^{n^i}(\sigma, q) - O_\nu^i(\sigma, \cdot)/N_\nu^i(\sigma) \| < (\varepsilon^i)^2 / P^i(\sigma) + 1/N_\nu^i(\sigma). \tag{7}$$

By Chebyshev's inequality, $N_v^i(\sigma)$ is greater than (recall that $P^i(\sigma)\geq\varepsilon^i$) $(\varepsilon^i)b/2$ with probability of at least $d_3=1-4/(\varepsilon^i)^2b$.

Thus, with probability of at least d_3, (7) can be rewritten as:

$$\|\ell_i^{n^i}(\sigma, q)-O_v^i(\sigma, \bullet)/N_v^i(\sigma)\|<\varepsilon^i+2/(\varepsilon^i)b\leq 2\varepsilon^i. \tag{8}$$

Since (1) does not hold, (8) implies

$$\|\ell_i^{n^i}(\sigma, q)-\ell_i^{n^i}(\sigma, \tau^i)\|\leq 2\varepsilon^i+\varepsilon^i=3\varepsilon^i. \tag{9}$$

In view of Remark 2, (9) yields

$$h_2^{n^i}(\sigma^i, q)\leq h_2^{n^i}(\sigma^i, \tau^i)+1/i+\delta(3\varepsilon^i, n^i). \tag{10}$$

If ε^i is appropriately chosen, $\delta(3\varepsilon^i, n^i)$ goes to zero as i tends to infinity.

To recapitulate, (10) holds true without player 2 being detected in B_v^i with probability of at most $c_v^i=(1-d_1)+(1-d_2)+(1-d_3)$. Notice that all the above estimates do not depend on the particular deviation, $\bar{g}$, played by player 2.

(ii) After a detection comes a punishment phase which pushes, with high probability (to be calculated like the calculations above), the punished player's payoff toward his individually rational level.

(iii) The sub-blocks are designed in such a way that a length of each sub-block B_v^i compared to its past tends to zero with i. Thus, there is no harsh punishing: the punisher can start punishing at the end of the sub-block, because the influence of the stage payoffs at a particular sub-block on the average payoff is negligible. Moreover, even in case where a deviation occurs at a sub-block located at the end of the block (which means that the punishment phase will be short – till the end of the block), the deviator cannot gain by much.

(i)–(iii) show that for any $\varepsilon>0$ there is a time T s.t. $t>T$ implies $h_2^t(f, \bar{g})\leq\gamma_2+\varepsilon$ for all $\bar{g}$. I.e., (f, g) is a uniform equilibrium.

As we have already the machinery we are able to show more. We will prove that $\limsup(1/T)\sum_{t=1}^{T} x_2^t(f, \bar{g})$ is smaller than γ_2 *almost surely*. Recall from (i) that player 2 can gain by more than $1/i+\delta(3\varepsilon^i, n^i)$ in B_v^i without being detected with probability that does not exceed c_v^i.

(iv) The number of repetitions b is so designed that $\sum_i\sum_{v=1}^{u_i} c_v^i<\infty$. Therefore, by the Borel-Cantelli lemma, the event that there are infinitely many B_v^i's in which the payoff of player 2 is greater than $\gamma_2+\varepsilon$, for $\varepsilon>0$ and he is not being detected is zero. And similarly for player 1.

(v) For the limsup of the averages of player 2's payoffs to exceed γ_2 it is necessary that there will be infinitely many sub-blocks B_v^i in which player 2's payoffs exceed $\gamma_2+1/i+\delta(3\varepsilon^i, n^i)$ without being detected. However, this event has, as was shown in (iv), probability zero. Therefore,

$$\limsup(1/T)\sum_{t=1}^{T} x_2^t(f, \bar{g})\leq\gamma_2 \text{ a.s.} \qquad \text{Q.E.D.}$$

6 Concluding Remark

The approach represented by the uniform equilibrium views the infinitely repeated game as an approximation of long finitely repeated games. Not only the payoff (of the infinite game) is defined as the limit of the expected finite average payoffs (i.e., limit of the payoffs in the truncated games) but also the incentive compatibility constraints are defined in terms of the finitely repeated game. One may take another approach which ignores the underlying stage game (and all the corresponding finitely repeated games). This is the "almost surely" approach.

Define (γ_1, γ_2) to be *a.s.-equilibrium payoff* if there is a pair of strategies (f, g) satisfying: (i) the limit of the finite average payoffs exists (f, g)-almost surely and moreover, the expected value of these limits is (γ_1, γ_2), and (ii) for every strategy $\bar{f}$ of player 1 the upper limit (limsup) of player 1's finite average payoffs do not exceed γ_1 $(\bar{f}, g)$-almost surely. And a similar condition for every $\bar{g}$.

There is no inclusion relation that can be derived directly from the two equilibria definitions. However, this paper shows (see Remark 4 and (iv) and (v) of Lemma 2's proof) that every uniform equilibrium payoff is an a.s.-equilibrium payoff. The question whether every a.s.-equilibrium payoff is also a uniform one remains yet to be answered.

Appendix 1

Lemma 5.5 of [L1] is the following:

Lemma: Let $R_1, \ldots, R_n$ be a sequence of identically distributed Bernoulli random variables with parameter p, and let $Y_1, \ldots, Y_n$ be a sequence of Bernoulli random variables, such that for each $1 \le \ell \le n$, R_ℓ is independent of $R_1, \ldots, R_{\ell-1}, Y_1, \ldots, Y_\ell$. Then

$$\text{prob}\left\{\left|\frac{R_1 Y_1 + \ldots + R_n Y_n}{n} - p\,\frac{Y_1 + \ldots + Y_n}{n}\right| \ge \varepsilon\right\} \le \frac{1}{n\varepsilon^2}.$$

To apply the lemma, put $R_t = R^t_\sigma$, $Y_t = Y^t_\tau$, and $n = b$.

Appendix 2

Fix a pure strategy σ of player 1.

Define $Z^t_s = 1$ if player 2 played a strategy which yields, together with σ, the signal s (of player 1) at the t-th repetition. Denote $Z^t = (Z^t_s)_{s \in L_1^{n^i}}$. Thus,

$$\ell_1(\sigma, q) = 1/b \sum_{t=1}^{b} Z^t \quad \text{and} \quad O^i_\nu(\sigma, \bullet)/N^i_\nu(\sigma) = 1/N^i_\nu(\sigma) \sum_{t \in B_\sigma} Z^t.$$

Moreover, for every s the following holds:

$$\left|1/b \sum_{t=1}^{b} Z_s^t - 1/N_v^i(\sigma) \sum_{t\in B_\sigma} Z_s^t\right|$$
$$\leq \left|1/b \sum_{t=1}^{b} Z_s^t - (1/bP^i(\sigma)) \sum_{t=1}^{b} R_\sigma^t Z_s^t\right|$$
$$+ \left|(1/bP^i(\sigma)) \sum_{t=1}^{b} R_\sigma^t Z_s^t - 1/N_v^i(\sigma) \sum_{t\in B_\sigma} Z_s^t\right|.$$

By using Appendix 1 (put $R_t = R_\sigma^t$ and $Y_t = Z_s^t$), the first summand is smaller than $(\varepsilon^i)^2/P^i(\sigma)$ with probability of at least $d_4 = 1-(\varepsilon^i)^4/b$. As for the second summand, notice that $\sum_{t=1}^{b} R_\sigma^t Z_s^t = \sum_{t\in B_\sigma} Z_s^t$. Moreover, $|1/bP^i(\sigma) - 1/N_v^i(\sigma)| < 1/N_v^i(\sigma)$ with probability of at least $d_5 = 1-(\varepsilon^i)^2/b$. This computation holds for a fixed signal s. Since there are at most k_2^i such signals (for a fixed σ) we have that, with probability of at least $1-((1-d_4)+(1-d_5))k_2^i$.

$$\|\ell_1(\sigma, q) - O_v^i(\sigma, \bullet)/N_v^i(\sigma)\| \leq (\varepsilon^i)^2/P^i(\sigma) + 1/N_v^i(\sigma)$$
$$\leq \varepsilon^i + 1/N_v^i(\sigma).$$

References

[A] Aumann RJ (1981) Survey of Repeated Games. In: Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern. Bibliographisches Institut, Mannheim-Wien-Zürich, 11–42.

[L1] Lehrer E (1990) Nash Equilibria of n-Player Repeated Games with Semi-Standard Information. International Journal of Game Theory, 19, 191–217.

[L2] Lehrer E (1986) Two Player Repeated Games with Non-Observable Actions and Observable Payoffs, to appear in Mathematics of Operations Research.

[L3] Lehrer E (1991) Internal Correlation in Repeated Games. International Journal of Game Theory, 19, 431–456.

[S] Sorin S (1990) Supergames. In: Game Theory and Applications, Ichiichi T, Neyman A, Tauman Y (eds.). Academic Press, 46–63.

Received July 1990
Revised version June 1991